AEROSPACE ENGINEERING DESK REFERENCE

Aerospace Engineering Desk Reference

Note from the Publisher

This book has been compiled using extracts from the following books within the range of Aerospace books in the Elsevier collection:

Filippone (2006) *Flight Performance of Fixed and Rotary Aircraft* 9780750668170
Megson (2007) *Aircraft Structures for Engineering Students* 9780750667395
Cook (2007) *Flight Dynamics Principles* 9780750669276
Jenkinson and Marchman (2003) *Aircraft Design Projects* 9780750657723
Tooley (2007) *Aircraft Digital Electronic and Computer Systems* 9780750681384
Tooley and Wyatt (2007) *Aircraft Communications and Navigation Systems* 9780750681377
Watkinson (2003) *Art of the Helicopter* 9780750657150
Curtis (2005) *Orbital Mechanics for Engineering Students* 9780750661690
De Florio (2006) *Airworthiness* 9780750669481

The extracts have been taken directly from the above source books, with some small editorial changes. These changes have entailed the re-numbering of Sections and Figures. In view of the breadth of content and style of the source books, there is some overlap and repetition of material between chapters and significant differences in style, but these features have been left in order to retain the flavour and readability of the individual chapters.

End of chapter questions
Within the book, several chapters end with a set of questions; please note that these questions are for reference only. Solutions are not always provided for these questions.

Units of measure
Units are provided in either SI or IP units. A conversion table for these units is provided at the front of the book.

Upgrade to an Electronic Version
An electronic version of the Desk reference, the *Aerospace Engineering e-Mega Reference,* 9781856175760

- A fully searchable Mega Reference eBook, providing all the essential material needed by Aerospace Engineers on a day-to-day basis.
- Fundamentals, key techniques, engineering best practice and rules-of-thumb at one quick click of a button
- Over 1,500 pages of reference material, including over 1,000 pages not included in the print edition

Go to http://www.elsevierdirect.com/9781856175753 and click on **Ebook Available**

Aerospace Engineering Desk Reference

ELSEVIER

Amsterdam · Boston · Heidelberg · London · New York · Oxford
Paris · San Diego · San Francisco · Sydney · Tokyo
Butterworth-Heinemann is an imprint of Elsevier

Butterworth-Heinemann is an imprint of Elsevier
Linacre House, Jordan Hill, Oxford OX2 8DP, UK
525 B Street, Suite 1900, San Diego, CA 92101-4495, USA

First edition 2009

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalog record for this book is available from the Library of Congress

ISBN: 978-1-85617-575-3

Printed and bound in the United States of America

09 10 11 11 10 9 8 7 6 5 4 3 2 1

Contents

Author Biographies

Michael Cook was a Lecturer for the Aerodynamics Department of the College of Aeronautics, Cranfield. He is a Fellow of the Royal Aeronautical Society and a Chartered Engineer. He has contributed to industrial research projects, including the BAe Fly-By-Wire Jaguar programme, various advanced helicopter concepts and the development of digital flight control technology. He has written or contributed to over 100 technical reports, papers and in 1994 was awarded the Royal Aeronautical Society John Britten prize for a paper on hang glider stability.

Howard D. Curtis is a former Associate Dean at the College of Engineering, ERAU, and Associate Fellow of the American Institute of Aeronautics and Astronautics (AIAA). He is also the former Chair of the Aerospace Engineering Department, ERAU. He has also held various roles at the Jet Propulsion Laboratory, MIT Lincoln Laboratory and Martin-Marietta.

Filippo De Florio is a retired aeronautical engineer and private pilot. He previously worked in the Technical Services at ALITALIA and AERFER and as a researcher in the Institute of Aerodynamics at the University of Naples. He was inspector at the RAI's (now ENAC) regional office in Milan and also director of the RAI/ENAC Type Certification Division in Rome. He was a Member of the JAA Certification Committee and is currently Honorary Member of UVS International and Member of the OSTIV Sailplane Development Panel. He was recently awarded the first UAS Pioneer Award.

Dr. Antonio Filippone is a senior lecturer in aerospace engineering at The University of Manchester, where he teaches subjects on flight mechanics, helicopter flight and high-speed aerodynamics. His research interests are in the area of aero-flight mechanics of fixed and rotary-wing aircraft, and also non-conventional vehicles at the frontier of flight.

Dr. Lloyd Jenkinson has worked for most of his professional life on aircraft project design in industry and at university. He has also acted as a consultant to industrial and government agencies. He has presented several technical papers on aircraft design at RAeS, AIAA and SAE conferences and journals. He is currently working as an engineering consultant, and as a part-time senior lecturer at Loughborough and Southampton Universities.

Dr. Jim Marchman is Professor of Aerospace and Ocean Engineering at Virginia Polytechnic Institute and State University, and consultant to various commercial agencies. At Virginia he has held several positions, including Assistant Department Head of Aerospace and Ocean Engineering and Associate Dean of Engineering for Academic Affairs. His experience elsewhere includes positions as an Aerospace Engineer, U.S. Army Aviation Test Activity, Edwards AFB, CA and Research Assistant, North Carolina State University. He is Associate Fellow of the AIAA and a member of the ASEE.

Dr. THG Megson took early retirement from his post as Senior Lecturer in Civil Engineering at the University of Leeds in 1996 and became a Senior Fellow lecturing part-time for three years until he retired fully in 1999. Since then he has produced the fourth edition of *Aircraft Structures*, been engaged in consultancy for different firms and been a governor of Harrogate Grammar School in North Yorkshire for eight years including a four-year term as Chairman.

Mike Tooley is an ex-Vice Principal and Dean of Engineering at Brooklands College in Surrey, England. He was responsible for the delivery of learning to over 10,000 Further and Higher Education students increasingly by flexible, open and on-line distance learning. Mike is the well-known author of a large number of popular engineering and related text books, including the principal course texts used by many of today's UK engineering students.

John Watkinson is an independent author, journalist and consultant in the broadcasting industry with more than thirty years of experience in research and development. He has held teaching posts at a senior level with The Digital Equipment Corporation, Sony Broadcasting and Ampex Ltd., before forming his own consultancy. He presents lectures, seminars, conference papers and training courses worldwide, in audio, video and data recording. He is a Fellow of the AES, a member of the British Computer Society and a chartered information systems practitioner.

David Wyatt is currently the Design Manager at Lees Avionics for GA and business aircraft. He was previously a Development Engineer (B747 avionic systems) at British Airways; Product Support Engineer (B737 FMCS) with Lear Siegler; Business Development Manager at Kidde Graviner (Eurofighter, Tornado and AVRO RJ fire and overheat protection systems). This was followed by commercial positions with Weston Aerospace (aero-engine transducers) and Ametek (cooling and demist systems). David also lectured EASA Part 66 modules at Brooklands College.

This page intentionally left blank

SECTION 1

Introduction

This page intentionally left blank

Aerospace Engineering Desk Reference; ISBN 9781856175753

1.0 Introduction

Antonio Filippone

> *. . . but the fact remains that a couple of bicycle mechanics from Dayton, Ohio, had designed, constructed, and flown for the first time ever a practical airplane.*
>
> J. Dos Passos, in *The Big Money*, 1932

After an uncertain start at the beginning of the 20th century, aviation has grown to a size on a global scale. By the year 2000, over 100 million passengers traveled through the airports of large metropolitan areas, such as London. In the same year, there have been 35.14 million commercial departures worldwide, for a total of 18.14 million flight hours. Demand for commercial air travel has grown by an estimated 9% a year since the 1960s. The expansion of the aviation services is set to increase strongly. Today, every million passengers contribute about 3,000 jobs (directly and indirectly) to the economy. Therefore, aircraft performance is a substantial subject.

The calculation and optimization of aircraft performance are required to:

- Design a new aircraft;
- Verify that the aircraft achieves its design targets;
- Efficiently operate an existing aircraft or fleet;
- Select a new aircraft;
- Modify, upgrade and extend the flight envelope;
- Upgrade and extend the mission profile;
- Investigate the causes of aircraft accidents;
- Provide data for the aircraft certification (*Certificate of Airworthiness*).

The engineering methods for the evaluation of aircraft performance are based on theoretical analysis and flight testing. The latter method is made possible by accurate measurement techniques, including navigation instruments. Flight testing is essentially an experimental discipline – albeit an expensive one. Performance flight testing involves the calibration of instruments and static tests on the ground, testing at all the important conditions, gathering of data from computers, data analysis, and calibration with simulation models. Wind tunnel testing is only used for the prediction of the aerodynamic characteristics. Graphical methods, such as finding the intersection between two performance curves, belong to past engineering practice. Analytical and numerical methods, including the equations of motion of the aircraft, are the subject of this textbook. Analytical methods yield closed-form solutions to relatively simple problems. Numerical solutions address more complex problems and allow the aircraft engineers to explore "virtually" the complete parametric space of the aircraft flight. This practice avoids expensive and risky flight testing. Methods for flight testing and evaluation of the fixed- and rotary-wing aircraft performance are discussed by Kimberlin, Olson and Cooke and Fitzpatrick, respectively.

Due to the variety of requirements, the subject of aircraft performance intersects several other disciplines, such as aircraft design, scheduling, operational research, systems, stability and controls, navigation, air traffic operations, flight simulation, optimization, in addition to aerodynamics, structures, propulsion systems and integration. Therefore, aircraft performance is essentially a multidisciplinary subject. Among the ones well known to the aerospace engineers there is the flight mechanics approach, the dynamics and aerodynamics of flight.

Personal interests may be involved in selecting the type of flight vehicles, as these include conventional airplanes, high-performance military aircraft, helicopters, V/STOL aircraft, rockets and vehicles for transatmospheric flight. Old and modern books on the subject deal only with some of these flight vehicles – as convenient.

The basic performance of the fixed- and rotary-wing aircraft can be calculated with little mathematical effort, using the one-degree of freedom model. However, a more accurate prediction of any performance parameter, particularly if the aircraft is maneuvering in unsteady mode, is a challenging subject, because it generally involves a number of free parameters in non-linear differential equations. It will be shown how the question of *how fast can an airplane fly* is difficult to answer. In short, it depends on *how* it flies.

Performance prediction is at the base of any concrete aircraft design methodology. The estimation of weights, range and power plant size requires the calculation of basic aircraft performance from a few input data. In this case the

approximation is generally good enough for parameter estimation and design. Input from operational parameters and flight testing is required for detailed analysis.

Performance optimization is at the heart of design and operation of all modern aircraft. From the operational point of view, commercial aviation is driven by fuel prices, and operations at minimum fuel consumption are of great relevance. Performance optimization requires notions of optimal control theory, a subject unfamiliar to aerospace engineers.

Performance efficiency goes beyond the design point and requires that the aircraft produces the best performance over the widest range of its flight envelope. For this reason, the subject of performance optimization is essential in design. The fighter jets Grumman F-14 and McDonnell-Douglas F-15 (1970s) were the first to be designed with the optimization approach, and all the aircraft of later generations were conceived in the same fashion.

In the past 30 years these optimal conditions have been increasingly challenged by environmental concerns, including noise emission, air quality near airports, global climate change and sustainability. Some aspects of the impact of aviation on climate change are the subject of routine review. Publications of relevance include the ones from the Intergovernmental Panel on Climate Change (IPCC).

1.1 Physical Units Used

International units (SI) are used whenever possible. Unfortunately, most data in aviation are still in imperial units. Conversion to international units is not foreseen for the immediate future. In most cases, the flight altitudes will be converted to feet, because of the extensive practice of working out the performance parameters in term of this unit. Speeds will also be given in knots or km/h.

The SI nomenclature notwithstanding, some spurious engineering units have had to be retained in some cases. One of the most confusing units ever devised is the kg. This unit is used for both weight (force) and mass: weight in kgf is equal to mass in kgm. This equivalence can fool any experienced engineer. Unfortunately, there is no way around it, because it is more convenient to denote a weight with kgf, rather than the newton. The mass, instead of the weight, appears in the energy equations, which is the main reason for retaining the kgm. By contrast, the weight appears in the aerodynamic coefficients, and if the other parameters are in international units, then the weight must be converted into newtons. Therefore, the confusion is sometimes overwhelming.

To the student approaching the subject for the first time there is a special word of caution. It is easy to miscalculate an aircraft's performance because of the use of non-conformal units. Some of the most common errors arise from using speeds in km/h instead of m/s, and kN or kW instead of N and W (thrust and power, respectively). The units for specific fuel consumption can also be confusing. With some critical thinking these errors can be avoided. A range of 2,600,000 km, instead of 2,600 km, is achieved by an airplane if one oversees the coherence of units. The former result is a distance from Earth to the Moon and back three times, while the correct result is a medium range flight in many parts of the world.

1.2 Performance Parameters

A performance parameter is a quantitative indicator representing how a vehicle operates in a specific flight condition. Typical performance parameters are weights, speeds, aerodynamic loads, engine thrust and power, range and endurance, accelerations, emission indexes (noise, exhaust gases) and many more. At least 60 different parameters can be taken into account in a full aircraft performance analysis.

In accident investigation, the flight parameters considered are the air speed, the Mach number, the dynamic pressure, the altitude, the air temperature, the rate of climb or descent, the flight path angle, the side-slip angle, the angular velocities and accelerations, the load factor, the rudder position, the control surfaces position, the fuel load and the engine's status.

It is not obvious what distinguishes a performance parameter from a purely aerodynamic, propulsion, operational parameter. The drag coefficient of the wing section is not a performance parameter, but the aircraft's drag coefficient is. The wing's aerodynamic characteristics are not a subject of aircraft performance, but the aerodynamic characteristics of the aircraft as a whole are. In performance analyses the drag coefficients are the known part of the problem, while in aircraft design they are part of the problem. The thrust is by itself an engine performance; the same engine mounted and integrated on the airframe becomes an aircraft parameter. The analysis must take into account that the system engine/airframe is not the same thing as the engine alone (airframe/engine integration). The stealth capabilities of an aircraft (radar signature, thermal signature, noise emission) depend more on the design of the aircraft than its operation. Not all the parameters will appear on the instrument panels in the cockpit, and some of them are not relevant to the pilot.

Many performance indexes cannot simply be expressed by a single value, but are presented with charts, because they are dependent on other parameters. The combination of those parameters is essential in defining the operation of the aircraft. Typical charts are the air speed relationships, the weight/altitude/temperature charts, the flight envelopes, and the payload range.

Some performance data are readily available from the manufacturer; other data can be inferred by appropriate analysis; others are clouded by secrecy or confidentiality; and others are difficult to interpret, because the conditions under which the aircraft performs are not given. Among the most common data covered by secrecy are the drag data, the stability characteristics, the excess power diagrams and the engine performance. Other examples are 1) the aircraft range, when the payload is not supplied together with the range; 2) the altitude at which this range is achieved; and 3) the radius of action of a military interceptor – this radius, in fact, may lie in the favourie field of enemy fire.

The maximum take-off weight (MTOW) and the operating empty weight (OEW) are available for most aircraft. However, these data are not sufficient to calculate the maximum payload, because the difference between MTOW and OEW must include the mission fuel. Therefore, some educated guess is needed. A weight advantage compared to heavier rivals translates into significant revenue-earning advantages, which in a competitive market is the most important factor for choosing and operating an aircraft. It is not uncommon to find manufacturers unhappy that their performance data and charts are published in the public domain. Performance charts allow customers and competitors to look at various options, to select the most competitive aircraft and to discover the flaws of the competitors' technology: sharing information makes everybody better players.

The purpose of this book is to take the reader through some simple performance calculations, to look at the performance data, and to give an introduction to aircraft performance optimization. A large number of data are published by *Jane's Information Systems* and *Flight International*; other valuable data for lesser known aircraft have been provided by Gurton and Loftin. The latter publication differs from the other ones because of its critical analysis. Loftin also points to additional sources of aircraft performance data. Further data have been taken from official documents of the international authorities (Federal Aviation Administration (FAA), International Civil Aviation Organization (ICAO)), and specialized publications such as Advisory Group for Aeronautical Research and Development (AGARD), Engineering Sheets Data Units (ESDU), and by our own research. ESDU provide data and methods on all areas of performance, and are of invaluable value to the practitioner engineer. Of particular interest are landing and take-off performance, drag of airframes, and range and endurance.

Updated data that are not proprietary are published regularly by the magazines *Flight International* and *Flug Revue.* All flight manuals report the essential performance curves of the aircraft and its engines, and include data that may not be available elsewhere. Most flight manuals are now available in electronic form, and represent a great wealth of information for the aircraft performance engineer.

As in any other technology sector, the operator of an aircraft is concerned that the performance parameters quoted by the manufacturer match the actual performance; therefore, accuracy of performance prediction methods is essential.

1.3 Performance Optimization

In the early 1950s, computers made their first appearance in aerospace engineering. Bairstow wrote in 1951 that:

> The use of electric calculators is coming in to reduce manual labor, but there is little hope of doing nearly all that we would like to do.

Computer solutions of aircraft performance are now routine jobs, and have reached a phenomenal level of sophistication, to include the coupling between flight mechanics, aerodynamics, structural dynamics, flight system control and differential game theory. With analog computers first, then with digital computers, the problems solved grew in complexity. See, for example, the 1959 edition of Etkin's book on flight dynamics to gain a perspective. In 1982, Ashley exemplified the problems of optimization in a paper titled "On Making Things the Best". The author argued, among other things, that flight planning ceased to be a matter of hand calculation by the time commercial jet propulsion was introduced (late 1950s).

There are two categories of optimization: optimization of aircraft performance during the design phase and optimization of operational performance for the given airplane. In the former case, one can investigate the alternative changes in configuration that improve one or more performance parameters. This is more appropriately the subject of aircraft design. We will consider some cases of operational optimization. An excellent source for optimization problems with aircraft applications is the classic book of Bryson and Ho on optimal control. Some of these problems, including multistage rocket trajectories, were also reviewed by Ashley.

Today there are programs that plan optimal trajectory routes to minimize Direct Operating Costs (DOCs), while complying with several airline constraints. These programs have several types of input data: weather conditions, route, aerodynamics, aircraft performance, and flight-specific information, such as payload, fuel cost, etc. On output they provide the amount of fuel for optimal cruise altitude, climb and descent points, optimal cruise speed, and flight path.

1.4 Certificate of Airworthiness

The Certificate of Airworthiness is a document that grants authorization to operate an aircraft. It specifies the limits of operation of the vehicle in terms of weights, take-off and landing requirements, and a number of other parameters, such as maintenance records, service history and compliance with safety regulations.

The certificate proves that the aircraft conforms to the type specified and it is safe to fly. The certificate is valid as long as the aircraft meets the type specification (commercial, commuter, utility, etc.); it is operated safely and all the airworthiness directives are met. The aircraft may lose its certificate for a number of reasons, including modifications, upgrades, and new directives approved by the international organizations that make the aircraft obsolete, not just unsafe to operate. Other documents are generally required, such as the type certificate data sheet, the certificate of maintenance, and a list of other papers. These documents seldom contain detailed performance data.

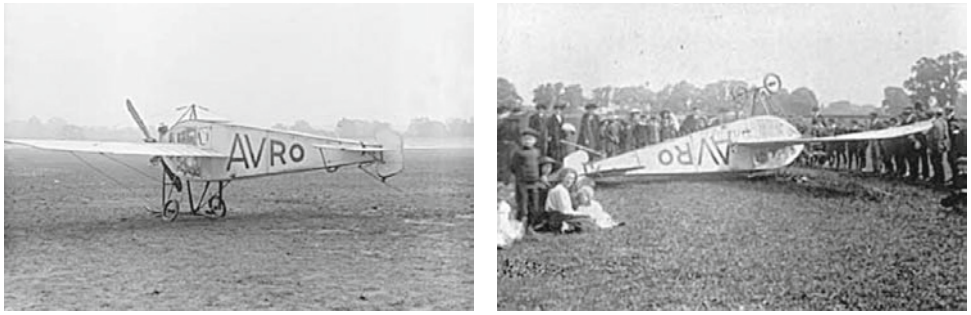
Certificates of Airworthiness are issued by the FAA in the USA, by the European Aviation Safety Agency (EASA), by the Civil Aviation Authority (CAA) in the UK and by other national and international bodies around the world. Certification is a complex legal and technical matter that is beyond the scope of this book.

1.5 Upgrading of Aircraft Performance

The age of bicycle mechanics has long passed. In the current technology situation, most aircraft are likely to be upgraded and modified to fit the changing market and technological advances. The technology that is fitted over the years can be phenomenally different from the first design. The service time of a single aircraft is of the order of 20 to 25 years, and the life of an aircraft family may exceed 50 years. A lifetime career can be devoted to a single airplane. The famed aircraft engineer Reginald J. Mitchell designed 24 aircraft, including the Spitfire, before dying prematurely, aged 42, in 1937.

To be fair, in the early days of aviation, a new aircraft could roll out of the factory in a few months. Indeed, some aircraft were prototypes that logged a few flights and then were scrapped – if they survived a crash. Figure 1.1 shows the Avro Model F (1912) at Manchester. This airplane was the first to have an enclosed cockpit, but it was capable of flying at only one speed, 65 mph. Only two airplanes were built. The picture to the right shows the airplane after it crash-landed in May 1912 due to an engine failure. The photo appears to have been published as a postcard.

It took only 43 days to build the Ryan NYP that made the transatlantic crossing in 1927, which included a total of 850 engineering hours (including performance and flight testing) and 3,000 man-hours for construction. In 1936, it took just one year for the German aircraft designer Kurt Tank to get from concept to first flight of the Focke Wulf Condor Fw-200, the first long-range passenger (and later reconnaissance and bomber) aircraft to fly from Berlin to New York without stopping en-route (1938).



(a) (b)

Figure 1.1 The Avro-F, built by A.V. Roe (1912). (a) Photo first published by the magazine *Flight* on 18 June 1942; (b) photo from the AV Roe Archives. Family outing with airplane crash (25 May 1912).

The wings of the Douglas DC-3 (1935), one of the most successful aircraft ever built, had simple performance improvement devices, a split flap for landing and outer-board ailerons for roll control. A jet aircraft of the first generation, such as the Boeing 727 (1958), had four outer-board leading-edge slats, three inboard leading-edge Kruger flaps, two banks of triple-slotted trailing-edge flaps, an inboard aileron for high-speed roll control, an outer-board aileron for low-speed flight, and seven spoilers (including five flight spoilers and two ground spoilers), also used as air brakes. This airplane is still flying.

By the 1960s, commercial airplane design and testing required thousands of man-years. The Boeing B-747-100, which first flew in 1969, required 15,000 hours of wind tunnel testing, 1,500 hours of flight testing with five aircraft over a period of 10 months, and 75,000 technical drawings. The latest version of this aircraft consists of about 6 million parts, 274 km of wiring and 8 km of tubing!

The B-747-400 incorporates major aerodynamic improvements, including a more slender wing with winglets to reduce drag. A weight saving of approximately 2,270 kg was achieved in the wing by using new aluminum alloys. Finally, the version B-747-400ER has an increased take-off weight of 412,770 kg. This allows operators to fly about 410 nautical miles (760 km) further, or carry an additional 6,800 kg payload, for a range up to 14,200 km.

An even older airplane is the Lockheed Hercules C-130A. Its first model was delivered to the US military in 1956. The design of this aircraft actually started several years earlier. By the early 1960s, a V/STOL variant was designed. Since then, the aircraft has progressed through at least 60 different variants. The current C-130J is actually a new airplane. Compared to the earlier popular version C-130E, the maximum speed is increased by 21%, climb time is reduced by 50%, the cruising altitude is 40% higher, the range is 40% longer, and its Rolls-Royce AE-2100DE engines generate 29% more thrust, while increasing fuel efficiency by 15%. With new engines and new propellers, the C130-J can climb to 9,100 m (28,000 feet) in 14 minutes.

Another example is the military utility helicopter CH-47, which has been in service since 1958. In particular, the MTOW has increased by over 50% and the useful payload has doubled. To the non-expert the aircraft looks the same as it did in the 1960s.

Technological advances in aerodynamics, engines and structures can be applied to existing aircraft to improve their performance. Over time weights grow, power plants become more efficient and are replaced, aerodynamics are improved by optimization, fuselages are stretched to accommodate more payload, and additional fuel tanks are added. This is one of the main reasons why aircraft manufacturers are not challenged to start a brand new design.

The conversion of aircraft for different commercial or military applications, and the development of derivative aircraft from successful aircraft require new performance calculations and a new certification. For example, the KC-10 tanker aircraft was derived from the commercial jet DC-10 (commercial to military conversion), and the Hercules C-130 was converted to the Lockheed L-100 (military to commercial). The conversion practice is more common with helicopters.

1.6 Mission Profiles

A mission profile is a scenario that is required to establish the weight, fuel, payload, range, speed, flight altitude, loiter and any other operations that the aircraft must be able to accomplish. The mission requirements are evidently specific to the type of aircraft. For high-performance aircraft they get fairly complicated and require some statistical forecasting.

Over the years many commercial aircraft operators have specialized in niche markets, which offer prices and services to selected customers. These niches include the executive jet operators between major business centers, operators flying to particular destinations (oil and gas fields), the all-inclusive tour operators to sunny holiday resorts, and the no-frills airlines flying to minor and underused airports. These operators have different schedules and cost structures.

First, let us start with long range passenger operations, which are serviced for the greatest part by subsonic commercial jets. The basic principle is that the airplane takes off from airport A and flies to airport B along a recognized flight corridor, then returns to A. The main parameters of the mission planning are the distance between the airports, the flight time, the downtime at the airport for getting the airplane ready (also called *time-on-station*), the flight speed, the local air traffic, and the departure times at both ends. Back at the airport of origin, the day is not over for the aircraft, and the operator wishes to utilize the airplane for another flight to the same destination, or to another destination – if possible. The key is the departure time, and the minimization of the curfew. Figure 1.2 shows the typical scheduling profile of such an airplane over a transatlantic route from a major airport in Europe to an airport on the East Coast of the United States.

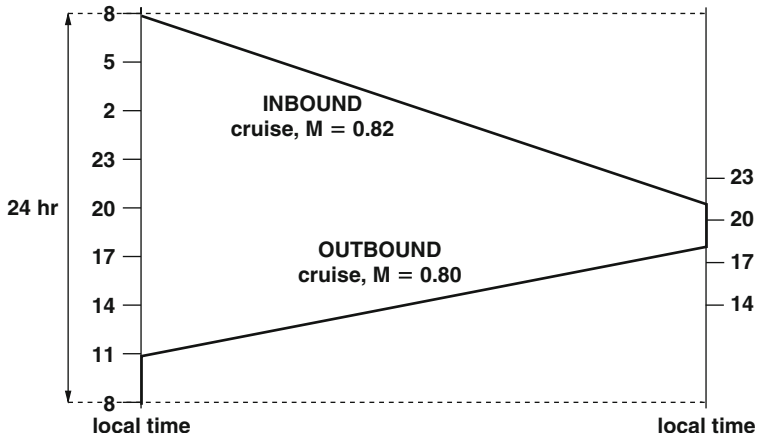


Figure 1.2 Scheduling of transatlantic flight. The numbers on the left and right side are local times.

Due to the time-zone effect, a late morning departure from Europe arrives to the USA in the mid-afternoon. An early evening departure arrives back in Europe in the early hours of the day after. Over the 24-hour period the airplane will have done a return flight and worked about 14 to 16 hours. For a flight arriving late in the evening, a return may not be possible until early morning on the next day. This adds to the operational costs, because of the need of maintaining the crew away from the home port. The time needed to get the aircraft ready for the next intercontinental flight may require up to 3 hours. Boarding of the Boeing-747 requires 50 to 60 minutes.

Scheduling of the type shown in Figure 1.2 leads to a block time of the order of 700 to 900 hours per year (depending on aircraft and service route). An airplane flying a day-time shorter route, and returning in the mid-afternoon, should be able to make another return flight, to the same destination or otherwise. For an airline company operating anything above a dozen airplanes, scheduling and optimal operation of the fleet is a complex problem. Events such as bad weather can lead to dozens of airplanes and flight crews out of position for several days. Scheduling and operation of aircraft is a subject for operations research, and is addressed by specialized publications. It deals with demand forecasting, network design, route planning, airline schedule planning, irregular operations, integrated scheduling, airport traffic simulation and control, and more.

1.6.1 Fighter Aircraft Requirements

The fighter aircraft has evolved from a reconnaissance airplane of the First World War to the most complex aircraft of modern days. Von Kármán reported that fighter aircraft first flew over the battlefields of Europe to spy on enemy lines. Obviously, enemy aircraft wanted to prevent this happening, so their pilots started shooting at enemy aircraft with a pistol. This was the beginning of a dog fight. Toward the end of the war, the Dutchman Anton Fokker, working at the service of the German Army, invented a system that synchronized the shooting of a machine gun through the propeller (*interrupter gear*) – mounted on a single-seater monoplane. With the interrupter pilots had their hands free to maneuver and fight at the same time. This advance was heralded as the birth of the fighter aircraft.

The requirements for fighter aircraft now include multipurpose missions, aircraft with complex flight envelopes, several configurations (changeable in flight), supersonic flight, combat capabilities, delivery of a wide range of weapon systems (all-weather operations), and maneuverability. There are dozens of different mission scenarios, as discussed extensively by Gallagher and co-workers. Typical missions are: basic, assault, combat, retrieval, close support, transport, refuel, and reconnaissance. For each of these missions there is a specific take-off weight, mission fuel, payload, range, maximum rate of climb, and service ceiling. This field is now so advanced that engineers use differential game theory and artificial intelligence to study the effectiveness of a given aircraft, and the tactical maneuverability to incoming threats.

One example of mission profile for this type of aircraft is shown in Figure 1.3. Such a profile must include warm-up, acceleration, take-off, climb, cruise, dash, combat, decelerate/climb, descent, and landing (with allowance for loiter and fuel reserve). In a more detailed breakdown, a typical plan may look like that in Table 1.1. The analysis of the various flight sections is essential in predicting the mission fuel; the mission fuel is essential for returning to base.

An alternative graphic method for indicating a mission profile is shown in Figure 1.4, representing an interdiction operation. The numbers indicate each flight segment. The vertical axis is an arbitrary flight altitude. The graph

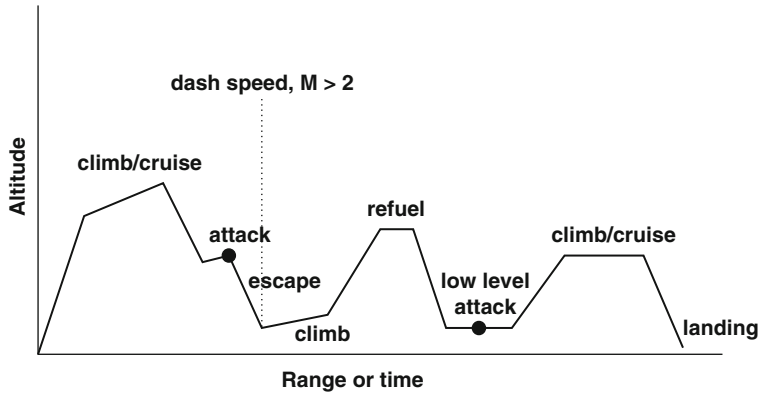


Figure 1.3 Generic mission profile for fighter aircraft.

Table 1.1 Summary of flight segments of a supersonic jet fighter.

Performance	Description
Warm-up, acceleration, take-off	2 minutes; 0.5 minutes at maximum power
Climb	at maximum power
Dash speed	$M = 1.8$, at 12,000 m (39,370 ft)
Combat	2 minutes, maximum power, dash speed
Subsonic cruise	30 minutes, $M = 0.85$
Decelerate and climb	to $M = 0.8$, at 9,500 m (31,168 ft)
Descent and landing	to sea level; no fuel credit
Loiter	10 minutes at sea level, minimum fuel
Landing	45 s at minimum power
Reserve fuel	5% of total mission fuel

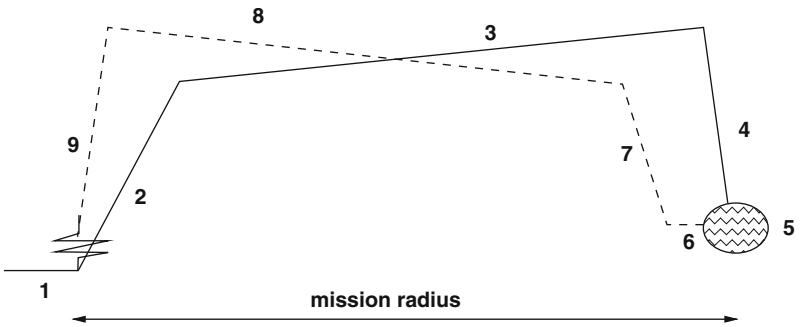


Figure 1.4 Generic mission profile for interdiction operation.

shows the mission radius (in arbitrary scale) and the point of engagement. Each segment is further specified by requirements such as those in Table 1.1.

The *dash speed* is a supersonic speed that the aircraft can maintain for a limited amount of time, or flight distance. It is usually the maximum speed at the best flight altitude. As indicated in the table, this altitude is around 12,000 m (39,370 feet).

Loiter is the operation around an airport. It usually consists of various turns along prescribed flight corridors, before the aircraft is permitted to land. Delay in landing (and longer loiter) may be due to air traffic control and weather conditions.

A performance index characteristic only of fighter jet aircraft is the *effectiveness*. Effectiveness is defined as the product of ordnance transport rate, availability in war time and kill effectiveness. The ordnance transport weight is the product of the ordnance mass and the number of sorties per day. The availability in war time is the time the aircraft is available for operations (compared to downtime for maintenance, service, loading, etc.). The killing effectiveness is the knocking-out success rate.

1.6.2 Supersonic Commercial Aircraft Requirements

After the Concorde era came to a close, no serious attempts have been made to replace the aircraft and operate a commercial flight at supersonic speeds. Nevertheless, the theoretical analyses regarding the feasibility of such an airplane under modern environmental and financial constraints abound. A new generation of supersonic civil transport aircraft that would replace the Concorde, should be able to fly longer routes, possibly at higher speeds. A Los Angeles–Tokyo route would require a cruise speed of $M = 2.4$ in order to be able to schedule two round trips over a 24-hour period. A replacement for Concorde, operating on the North Atlantic routes, should be able to fly at $M = 2.0$, or possibly lower, if the turn-around time can be reduced. This speed is important in the cycle because it allows the airplane to be serviced at both ends, avoiding long curfews. The operators of Concorde could make a profit (once the mortgage for the acquisition of the aircraft was taken out of the spreadsheets) by having two return flights per day.

Problems

1. Discuss the possible mission profiles for a V/STOL aircraft, and extract a set of performance criteria that can be applied to all operational conditions.
2. Make a list of all the performance segments of a supersonic jet fighter, and provide a critical discussion. Provide a scenario to deal with a fuel shortage at the end of a scheduled operation, before returning to base.
3. You are asked to plan a flight timetable between London and Berlin. Provide a plan for a subsonic jet transport that maximizes the block time for the operation between the two cities. Analyze the alternative, consisting in operating a turboprop aircraft. Do the necessary research of the data needed for the solution of this problem (flight corridor, distance in nautical miles, estimated flight time, flight speeds, etc.).
4. The Boeing B-52 is one of the oldest aircraft still in service. It has progressed from the first version in 1954, B-52A, to the version B-52G. Do the necessary research to investigate how propulsion, aerodynamics and general performance parameters have changed from the first to the latest version. List the most important quantities in a spreadsheet and draw a conclusion.
5. Analyze the ground operations required to get a Boeing B-747-400 ready for an intercontinental flight (refueling, systems checks, food supplies, water, boarding of passengers). Produce a spreadsheet that indicates the time of each operation and which operations can be performed in parallel (*problem-based learning: additional research is required*).

SECTION 2

Flight Performance

This page intentionally left blank

Aerospace Engineering Desk Reference; ISBN 9781856175753

2.1 The Aircraft and its Environment

Antonio Filippone

Contrary to the British and German military authorities, both of whom believed in nose armament, the Italians took the view that speeds of 250 mph made frontal attacks by fighters unlikely.

J.H. Stevens, 1953

This chapter discusses the general aircraft model for the rigid-body approximation, the reference systems, the nomenclature of the aircraft, forces, moments and angles. The aircraft operates in the atmosphere; therefore, the standard air model is reviewed, along with relevant approximating functions for performance calculations. Some atmospheric effects in non-standard conditions are briefly reviewed.

2.1.1 General Aircraft Model

Since the late 1940s, the accepted aircraft model for performance calculations consists of a point mass concentrated at the center of gravity. The engines are assumed to operate at the aircraft symmetry plane. There have been attempts to improve on the point-movement prediction methods, to include the fact that thrust, aerodynamic center and weight operate at different points. Aircraft flexibility is important at supersonic speed. Although a subsonic aircraft has a nearly "square" dimension, a viable supersonic transport aircraft has a length about double its wing span.

Variable geometry produces changes in the aerodynamic coefficients and in the handling quality. Longitudinal flexibility can be controlled by combined wing, tail-plane and canards. Allowance for wing flexibility is essential in special flight conditions, in order to avoid flutter effects. Wing flexibility at take-off may also contribute to the ground performance. Current research focuses on the aerodynamics, structural dynamics and flight mechanics coupling. However, the point model is quite accurate for most purposes. Therefore, only forces applied to this point and moments calculated around this point must be considered.

The model is shown in Figure 2.1-1: FL denotes the fuselage longitudinal axis. This is assumed to run through the center of gravity; α is the nominal angle of attack of the aircraft – it is defined as the angle between the FL and the true air speed vector. The angle of attack of the aircraft and the angle of attack of the wing are two different quantities, because the reference line of the wing is a chord-line. The angle of the thrust on the FL is called ϵ. This angle is generally quite small, and for the purposes of this discussion can also be considered zero (e.g. $\alpha + \epsilon \sim \alpha$). The angle of climb γ is the angle between the true air speed vector U and the horizon.

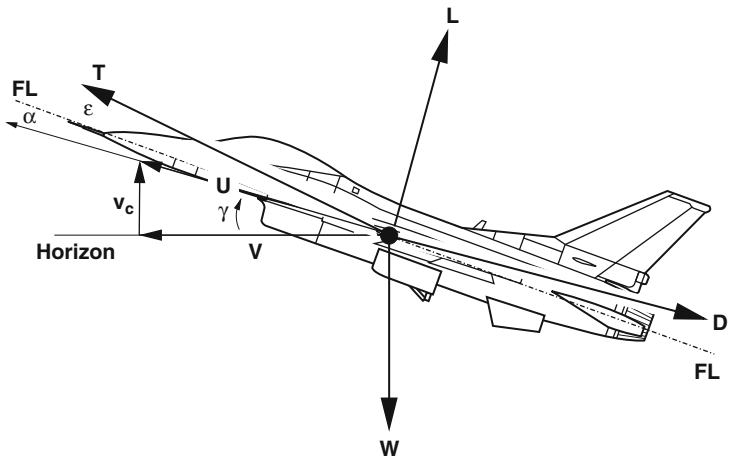


Figure 2.1-1 Aircraft model in the vertical plane, with forces concentrated at the CG.

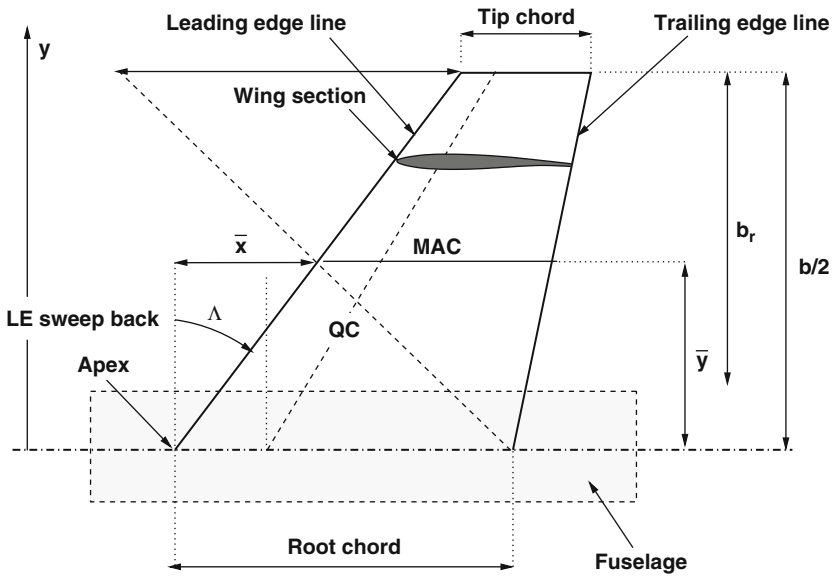


Figure 2.1-2 Planform view of wing geometry.

One shape factor, mostly used in supersonic aerodynamics, is the fineness ratio l/d', where

$$d' = \sqrt{\frac{4A_b}{\pi}} \tag{2.1.1}$$

is the equivalent body diameter, A_b is the maximum cross-sectional area of the aircraft, and l is the aircraft's length. The main wing of the aircraft is characterized by a number of essential parameters, as shown in Figure 2.1-2. The aspect-ratio is

$$\mathcal{AR} = \frac{b^2}{A}, \tag{2.1.2}$$

where b is the wing span and A is the wing area. Then there is the leading edge and the quarter chord sweep angle (Λ_{le}, Λ_{qc}); the taper ratio (c_{tip}/c_{root}); the root-mean-square thickness ratio

$$t/c = \left[\frac{1}{b/2 - b_r}\int_{b_r}^{b/2} (t/c)^2\, dy\right]^{1/2}, \tag{2.1.3}$$

with b_r = spanwise location of the wing root. The mean aerodynamic chord is

$$\text{MAC} = \frac{2}{A}\int_o^{b/2} c^2(y)\, dy. \tag{2.1.4}$$

The dihedral angle ($\varphi > 0$) is the angle with respect to the ground plane made by the leading edge or quarter-chord line. If $\varphi < 0$, the wing is said to have an anhedral. A number of analytical relationships exist to correlate various wing parameters. These are reported by some reference manuals, such as AIAA.

2.1.2 Reference Systems

There are three essential reference systems: the Earth system, the body system, and the wind system. Local reference systems may be required for specific reasons (for example, wing aerodynamics).

The aircraft is supposed to fly with respect to a Cartesian system fixed on the ground (*Earth axes*), which for our purposes is considered flat. In fact, most of the performance calculations will be done for relatively short flight times and at relatively low altitudes. The Earth's curvature and rotation are important for inertial navigation systems and to take into account the Coriolis effects (accelerations) over a rotating Earth. The Coriolis acceleration is estimated at less $10^{-3}\, g$ in atmospheric flight mechanics. The gravitational field is characterized by a constant acceleration of gravity, equal to the standard value of $g = 9.807$ m/s^2. The Earth system has the x axis pointing North, the z axis normal to the ground and pointing downward; the y axis pointing East, and making a right-hand system with x and z.

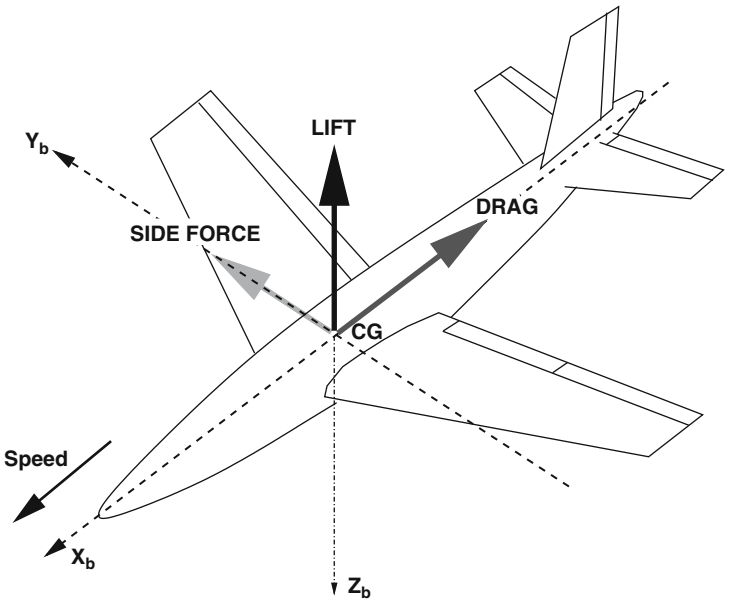


Figure 2.1-3 Body reference system and forces on the aircraft.

There are several ways to define reference axes on the airplane; the choice will be limited by the fact that there is always one plane of symmetry. This is not always the case. In the early days of aviation, symmetric wings were relatively unstable, and some aircraft designers used asymmetric concepts to alleviate the rolling problem. For example, the Ansaldo SVA (1917) had unequal wing spans, the Messerschmidt Bf-109/Mf-109 (1935) had an asymmetric fin airfoil, the Republic P-47 (1941) had an offset fin, and the Spitfire I (1934) had asymmetric radiators under the wings to compensate for the engine-propeller torque*.

Back to the concept of aircraft coordinates, there will be a body-conformal orthogonal reference system, centered at the center of gravity (CG) of the airplane. The subscript "b" will be used to denote body axes. The position of the CG is a known parameter, although its estimation is not straightforward. Also, its position changes with the aircraft loading. This fact is important for the analysis of stability and control. The reference system is shown in the three-dimensional view of Figure 2.1-3. The longitudinal axis x is oriented in the direction of the forward speed (wind axis); the z axis is vertical (along the acceleration of gravity g) and the y axis makes a right-hand Cartesian system with x and z. The positive y axis is at the starboard side of the airplane.

The main forces on the aircraft (propulsive, aerodynamic, inertial) are applied at the CG. The drag is the opposite direction of the air speed; the lift is at 90 degrees with the drag; the weight is vertical and pointing downward. The thrust generated by the engines on both sides of the aircraft is replaced by a single thrust force.

The correlation between body and Earth reference system is done by three attitude angles. The *pitch attitude* of the aircraft is the angle θ between the longitudinal axis and the horizontal plane (positive with the nose up). The *yaw attitude* is the angle between the aircraft's speed and the North–South direction. It is positive clockwise, e.g. when the aircraft is heading eastwards. This is sometimes called *heading* or *azimuth angle* ψ. The *bank attitude* ϕ is the angle between the aircraft spanwise axis y_b and the horizontal plane.

The side force would not normally be present on the aircraft (and its occurrence should be avoided). It is mostly due to atmospheric effects (lateral gusts), asymmetric thrust, and center of mass off the symmetry line (due, for example, to a differential use of the fuel in the wing tanks). The presence of such forces may lead to a yawed flight condition. The yaw angle β is the angle between the longitudinal axis and the true air speed vector.

The velocity (wind) axis reference system indicates the direction of the flight path with respect to the Earth system. At any given point on the trajectory the aircraft will have a track and a gradient. The track is the angle on the horizontal plane between the flight direction and the North–South axis. The gradient is the angle of the velocity on the horizon, which we have called γ in Figure 2.1-1.

*Asymmetric wings can be employed to reduce the wave drag at supersonic speeds, but this is an aerodynamic problem on its own right.

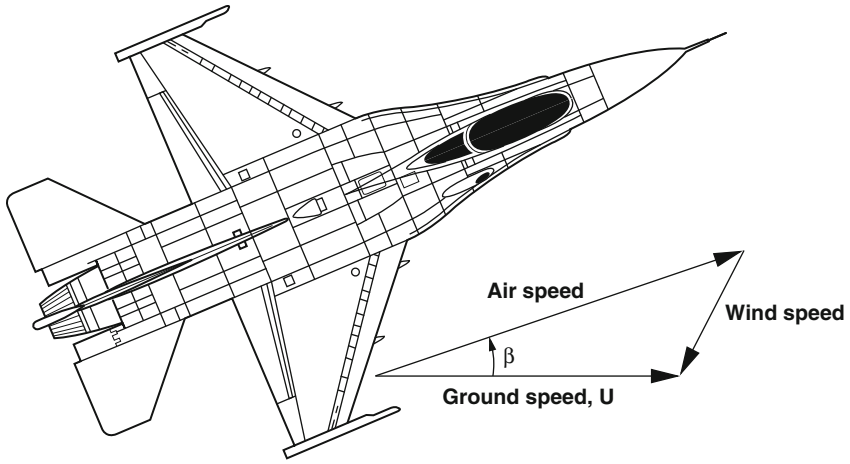


Figure 2.1-4 Vector relationship between ground speed and air speed.

If $\boldsymbol{V}$ is the ground speed, $\boldsymbol{V}_w$ is the wind speed, the air speed is found from

$$\boldsymbol{V}_a = \boldsymbol{V} + \boldsymbol{V}_w. \tag{2.1.5}$$

The corresponding side-slip angle β is indicated in Figure 2.1-4.

The transfer of forces between one reference and the other is done through rotation matrices. The order of these rotations is important for the correct development of the flight mechanics equations. The full derivation of these equations may require several pages. A modern presentation is given by Yechout and co-workers. A detailed discussion of reference systems and flight paths on a curved surface has been provided by Miele.

2.1.2.1 Angular Relationships

We consider some simple flight cases to provide a correlation between angles in the different reference systems, Figure 2.1-5. First, consider the pitch angles. From the definitions, the pitch attitude θ is related to the angle of attack α and the flight path gradient γ by

$$\theta = \alpha + \gamma. \tag{2.1.6}$$

Next, consider a yaw. The heading ψ is related to the track ξ and the side-slip by the equation

$$\psi = \xi - \beta. \tag{2.1.7}$$

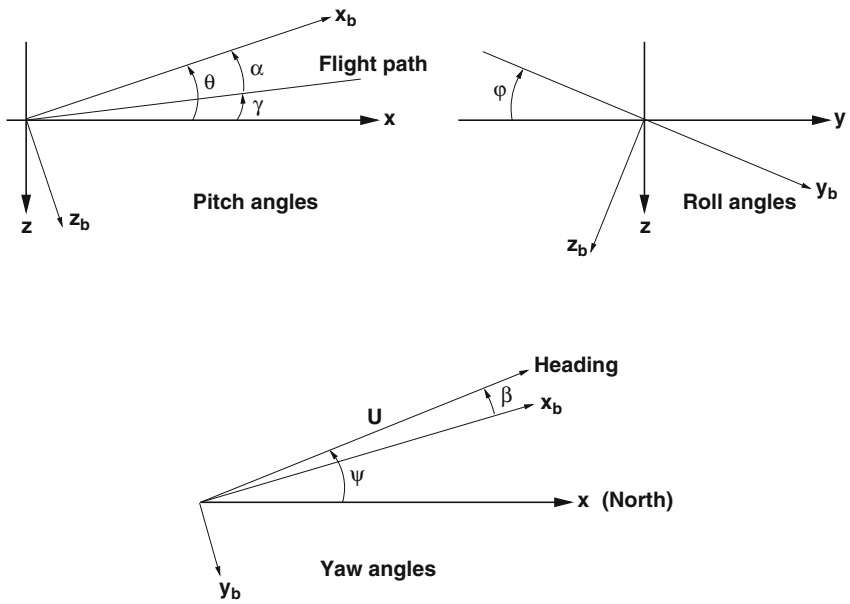


Figure 2.1-5 Relationships between angles and reference systems.

If there is no side-slip (normal flight condition) the heading and the track are the same angle. Finally, consider a roll problem. The bank angle ϕ is the inclination of the spanwise body axis on the horizontal plane. The bank attitude is the same as the bank angle on a level flight path. However, if the aircraft climbs or descends the correlation becomes complicated.

2.1.3 Forces on the Aircraft

For the fixed-wing aircraft the balance of forces on the aircraft, after assuming that these are applied at the CG, is the following

$$\boldsymbol{F} = m\frac{\partial \boldsymbol{V}}{\partial t}, \tag{2.1.8}$$

where the right-hand side denotes the acceleration of the center of gravity. There will be essentially three types of force: aerodynamic ($\boldsymbol{A}$), propulsion ($\boldsymbol{T}$) and gravitational ($m\boldsymbol{g}$). If we specialize these actions, Eq. 2.1.8 becomes

$$\boldsymbol{A} + \boldsymbol{T} + m\boldsymbol{g} = m\frac{\partial \boldsymbol{V}}{\partial t}. \tag{2.1.9}$$

An essential concept is the *trim*. An aircraft is said to be trimmed if the sum of all forces is zero, and the sum of all moments is zero. Trimming the aircraft may have the undesirable effect of creating other forces, such as additional drag. Trim is required in most cases, because the resulting forces on the aircraft do not operate at the center of gravity, as assumed in this book.

2.1.4 Moments of Inertia

The moment of inertia is representative of the inertia of a body to rotational accelerations, just as the mass is the inertia of a body to a linear acceleration. The moment of inertia about a generic axis is the volume integral

$$I = \int_V r^2\, dm. \tag{2.1.10}$$

The term $r^2\, dm$ is the moment of inertia of the mass dm with respect to the axis at a distance. The units of the moment of inertia are [kg m^2]. For analyses with respect to a Cartesian reference system, there are three principal moments of inertia. For example, the moment of inertia about the x axis is

$$I_x = \int_V x^2 dm = \int_V x^2 \rho dV = mr_x^2, \tag{2.1.11}$$

where r_x is the *radius of gyration* with respect to the principal axis x. The radius of gyration expresses the distance at which the aircraft mass should be concentrated to give the same moment of inertia.

The calculation of the moment of inertia requires the exact knowledge of the mass distribution around the axis and the geometry of the aircraft. Applications of the moments of inertia to aircraft performance are discussed elsewhere. A first-order approximation for the moments of inertia around the x axis and y axis (roll and pitch) of a generic aircraft is

$$I_x = \frac{1}{4}\frac{b^2 W r_x^2}{g}, \quad I_y = \frac{1}{4}\frac{l^2 W r_y^2}{g}, \tag{2.1.12}$$

where l is the aircraft's length, and r_x and r_y are radii of gyration. These can be derived from the calculation of existing aircraft. The moments of inertia are proportional to the square of the reference length and proportional to the weight.

2.1.5 Flight Dynamics Equations

The rigid aircraft is defined by six degrees of freedom. These are the parameters needed to identify completely the position and orientation of the aircraft in the Earth axes. The parameters are the coordinates of the center of gravity (x, y, z), and the orientation of the body axes on the Earth axes, (θ, ψ, ϕ).

In the specialized literature dealing with stability and control the angular velocities are called p, q, r, respectively. Unfortunately, these symbols generate confusion with other quantities (pressure p, dynamic pressure q, radius r, etc.). In stability and control the moments are called $\mathcal{L}, \mathcal{M}, \mathcal{N}$. A summary of symbols is given in Table 2.1-1.

Table 2.1-1 Nomenclature and symbols in the body-conformal reference system.

Body-conformal axes	x_b	y_b	z_b
Velocity components	u	v	w
Principal moments of inertia	I_x	I_y	I_z
Rotation angles	θ	ϕ	ψ
Angular velocities	$\dot{\theta}$	$\dot{\phi}$	$\dot{\psi}$
(Alternative symbols)	p	q	r
Moments	M_x	M_y	M_z
(Alternative symbols)	$\mathcal{L}$	$\mathcal{M}$	$\mathcal{N}$
Moment coefficients	C_{Mx}	C_{My}	C_{Mz}

The rigid-body velocity with respect to a reference system on the ground is

$$\boldsymbol{V} = \boldsymbol{V}_\infty + \boldsymbol{\Omega} \times \boldsymbol{r}, \tag{2.1.13}$$

where $\boldsymbol{V}_\infty$ is the velocity of the center of gravity; $\boldsymbol{\Omega} = (\dot{\theta}, \dot{\phi}; \dot{\psi})$ is the rotation vector centered at the CG; and (x, y, z) is the vector distance between the reference system on the ground and the CG. The linear acceleration of the aircraft is found from the derivation of Eq. 2.1.13,

$$\boldsymbol{a} = \frac{\partial \boldsymbol{V}_\infty}{\partial t} + \frac{\partial}{\partial t}(\boldsymbol{\Omega} \times \boldsymbol{r}). \tag{2.1.14}$$

These are general equations. However, for practical reasons, in the following chapters we will consider U as the total velocity in the flight path, V the total velocity parallel to the ground and v_c the vertical velocity. If the aircraft accelerates around its axis, the scalar form of Eq. 2.1.14 is quite elaborate. It is, in fact, of little practical interest in performance calculations, since only particular flight conditions are considered.

However, for a number of problems it is important to calculate the relationship between relative and absolute velocities and accelerations. Consider a reference on the ground $\{O, x, y, z\}$, and a non-inertial reference $\{O_1, x_1, y_1, z_1\}$. The speed of the aircraft in the Earth axes is

$$\boldsymbol{V} = \boldsymbol{V}_\infty + \omega \times \boldsymbol{r} + \boldsymbol{V}_r, \tag{2.1.15}$$

where $\boldsymbol{V}_\infty$ is the speed of the non-inertial system, ω is its rotational speed, $\boldsymbol{r}$ is the position vector of the CG in this system, and $\boldsymbol{V}_r$ is the relative velocity. The acceleration is found from the time derivative of Eq. 2.1.15. The results of classical mechanics give

$$\boldsymbol{a} = \boldsymbol{a}_\infty + \omega \times (\omega \times \boldsymbol{r}) + \dot{\omega} \times \boldsymbol{r} + \boldsymbol{a}_r + 2\omega \times \boldsymbol{V}_r. \tag{2.1.16}$$

In this equation $\boldsymbol{a}_\infty$ is the acceleration of O_1, and $\boldsymbol{a}_r$ is the relative acceleration of the CG in O_1. The term

$$\boldsymbol{a}_t = \boldsymbol{a}_\infty + \omega \times (\omega \times \boldsymbol{r}) + \dot{\omega} \times \boldsymbol{r} \tag{2.1.17}$$

is the transport acceleration, and

$$\boldsymbol{a}_{Cor} = 2\omega \times \boldsymbol{V}_r \tag{2.1.18}$$

is the Coriolis acceleration. Consequently, we have

$$\boldsymbol{a} = \boldsymbol{a}_t + \boldsymbol{a}_r + \boldsymbol{a}_{Cor}. \tag{2.1.19}$$

In conclusion, the absolute acceleration is the sum of the relative, transport and Coriolis accelerations.

2.1.6 The International Standard Atmosphere

Nearly all the basic calculations of aircraft performance are done in International Standard Atmosphere (ISA) conditions, whose parameters at sea level are given in Table 2.1-2. The standard humidity is zero, which is far from true. Sometimes the reference to ISA conditions is replaced with a reference to *standard day*. For temperatures above or below the standard day, there is a reference to *hot* or *cold* day.

Observations on the state of the atmosphere at sea level go back hundreds of years, but they have become systematic in the last century with aviation, rocket and satellite data and perfect gas theory. A number of *standard* versions exist: NACA's atmosphere (1955), the ARDC (1959), the US standard (1962, amended in 1976) and the ICAO standard.

Table 2.1-2 Sea level data of the International Standard Atmosphere.

Parameter	Symbol	Sea level value
Temperature	$\mathcal{T}_o$	15.15°C
Pressure	p_o	$1.01325 \cdot 10^5$ Pa
Density	ρ_o	1.225 kg/m^3
Viscosity	μ_o	$1.7894 \cdot 10^{-5}$ Ns/m^2
Humidity		0%

These tables are basically equivalent to each other up to about 20 km (65,000 ft), that covers most of the atmospheric flight mechanics. We shall be concerned with altitudes below 31 km, or about 100,000 ft.

Most books of performance and flight mechanics (as well as books of aerodynamics) report these tables. These are made obsolete, from an engineering point of view, by the availability of analytical correlations of great accuracy*.

The atmosphere is divided into a number of layers. These layers are identified for the sole purpose of our discussion into atmospheric aircraft performance.

The atmosphere below 11,000 m (36,089 ft) is called *troposphere*. It is characterized by a decreasing temperature from sea level, and reaches a standard value of −56.2°C. The altitude of 11,000 m is called *tropopause*. The level above is called *lower stratosphere* and covers an altitude up to 20,000 m (65,627 ft), in which the temperature remains constant. The air density keeps decreasing with the increasing altitude. The upper limit of this layer includes most of the atmospheric flight vehicles powered by air-breathing engines. The *middle stratosphere* reaches up to an altitude $h = 32{,}000$ m (104,987 ft). In this layer the atmospheric temperature increases almost linearly from the value of −56.2°C. The edge of space is generally considered to be at an altitude of 100.5 km, where the gravity is considerably lower than at sea level.

A number of functions are sometimes used to approximate the ICAO data. For the temperature a linear expression is used:

$$\mathcal{T} = \mathcal{T}_o - 0.0065h, \tag{2.1.20}$$

where $\mathcal{T}_o$ is the standard sea level temperature and h is the altitude in meters. If we use the equation of ideal gases to describe the atmosphere,

$$\frac{p}{\rho} = \mathcal{R}\mathcal{T}, \tag{2.1.21}$$

then

$$\frac{p}{\rho\mathcal{T}} = \frac{p_o}{\rho_o\mathcal{T}_o} = \mathcal{R}, \tag{2.1.22}$$

or

$$\frac{p}{p_o} = \frac{\rho}{\rho_o}\frac{\mathcal{T}}{\mathcal{T}_o}. \tag{2.1.23}$$

The value of the gas constant $\mathcal{R} = 287$ J/kg K. The relative density is called σ, the relative pressure is δ and the relative temperature is θ. Therefore,

$$\delta = \sigma\theta. \tag{2.1.24}$$

We call the altitude corresponding to a given air density *density altitude*. If, instead, we relate the altitude to the local air pressure, then we have a *pressure altitude*. In order to find the pressure/altitude and the density/altitude relationships we use the buoyancy law for still air along with Eq. 2.1.20, to find the rate of change of the pressure with altitude. The buoyancy law is

$$\frac{\partial p}{\partial h} = -\rho g. \tag{2.1.25}$$

If we insert the differential form of Eq. 2.1.20, this law becomes

$$\frac{\partial p}{\partial \mathcal{T}} = -\frac{\rho g}{\lambda}. \tag{2.1.26}$$

* Programs that perform ISA calculations can be found from the public domain.

The next step is to use Eq. 2.1.21 to eliminate the density from Eq. 2.1.26

$$\frac{\partial p}{\partial \mathcal{T}} = -\frac{pg}{\lambda \mathcal{R} \mathcal{T}}. \tag{2.1.27}$$

Rearranging of this equation leads to

$$\frac{dp}{p} = -\frac{g}{\lambda \mathcal{R}} \frac{d\mathcal{T}}{\mathcal{T}}. \tag{2.1.28}$$

Integration of Eq. 2.1.28 yields

$$\ln p = -\frac{g}{\lambda \mathcal{R}} \ln \mathcal{T} + c, \tag{2.1.29}$$

where c is a constant of integration that is found from the sea level conditions (Table 2.1-2). The final result is

$$\ln\left(\frac{p}{p_o}\right) = \frac{g}{\lambda \mathcal{R}} \ln\left(\frac{\mathcal{T}}{\mathcal{T}_o}\right), \tag{2.1.30}$$

$$\delta = \frac{p}{p_o} = \left(\frac{\mathcal{T}}{\mathcal{T}_o}\right)^{g/\lambda \mathcal{R}}. \tag{2.1.31}$$

The value of the power coefficient $g/\lambda \mathcal{R} = 5.25864$. If we insert Eq. 2.1.20 in Eq. 2.1.31, we have a *pressure/altitude* correlation:

$$\delta = \frac{p}{p_o} = \left(1 - \frac{0.0065}{\mathcal{T}_o} h\right)^{5.25864} = (1 - 2.2558 \cdot 10^{-5}\, h)^{5.25864}, \tag{2.1.32}$$

with h expressed in meters. Equation 2.1.32 is in good agreement with the ICAO data. A more approximate expression is

$$\delta = \frac{p}{p_o} = (1 - 2.2558 \cdot 10^{-5}\, h)^{5.25588}, \tag{2.1.33}$$

The advantage of Eq. 2.1.33 is that the altitude is related to the pressure ratio, and therefore it can be read directly from the altimeter that is calibrated with the ISA reference value of p_o. To find a *density/altitude* correlation we use Eq. 2.1.24, with the relative pressure from Eq. 2.1.33:

$$\sigma = \frac{(1 - 2.2558 \cdot 10^{-5}\, h)^{5.25588}}{1 - 0.0065\, h/\mathcal{T}_o}. \tag{2.1.34}$$

The accuracy of the approximating function is shown in Figure 2.1-6 for the air density up to an altitude of 22,000 m. This accuracy is good enough for all the type of calculations shown in this book (see Problem 1).

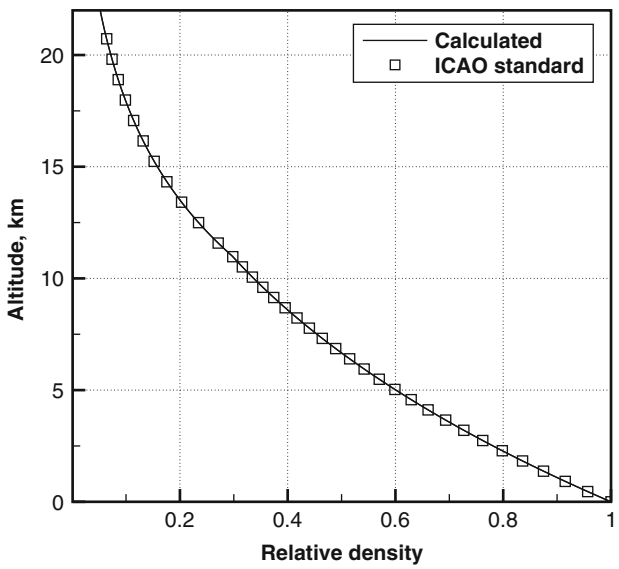


Figure 2.1-6 ICAO air density and approximated function.

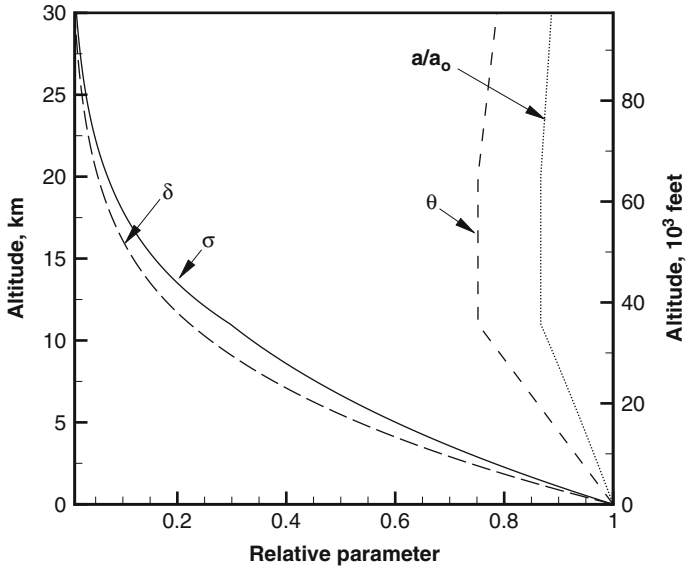


Figure 2.1-7 ISA relative parameters as a function of altitude.

Figure 2.1-7 shows the ratios of density, pressure, temperature and speed of sound from sea level to the altitude of 20,000 m. These data are calculated from Eq. 2.1.20, Eq. 2.1.33, and from the definition of speed of sound:

$$a \simeq \sqrt{\frac{p}{\rho}} = \sqrt{\gamma \mathcal{R} T}, \tag{2.1.35}$$

where γ is the ratio between specific heats ($\gamma = 1.4$).

Finally, the value of the air viscosity from the air temperature can be found from the following relationship

$$\frac{\mu}{\mu_o} = 8.14807 \cdot 10^{-2} \frac{T^{3/2}}{T + 110.4}. \tag{2.1.36}$$

From a computational point of view, we can construct a routine called

```
atmosphere (h, sigma, delta, theta, asound)
```

that returns the relative air density, pressure, temperature and speed of sound for any input altitude h. This routine will be useful in all the performance calculations presented in later chapters.

Computational Procedure

1. Calculate the local temperature from Eq. 2.1.20.
2. Calculate the relative pressure from Eq. 2.1.33.
3. Calculate the relative density from Eq. 2.1.34.
4. Calculate the speed of sound from Eq. 2.1.35.

The inverse problem (calculation of the altitude h corresponding to relative density σ) is more elaborate, because it requires to solve a non-linear equation in implicit form. The solution can be found with a bisection method. For the bisection method to work, one has to choose two points at which the function has opposite values. It is safe to choose $\sigma_1 = 0.01$ and $\sigma_2 = 1$, to make sure that the method converges to a solution.

Example

Commercial passenger jets cruise at altitudes between 9,000 and 12,000 m (30,000 to 40,000 ft). The standard temperature at those altitudes is between −56°C and −50°C. The relative pressure is between 0.29 and 0.19. Without cabin pressurization, air conditioning and heating, it would not be possible to fly passengers. Airplanes such as the Boeing B-737 have two air conditioning systems. If one system does not work, the aircraft can still fly, but it has

to maintain a cruise altitude below 7,620 m (25,000 ft, $\delta = 0.3711$). At this altitude it is possible to breathe with some difficulty. In case of failure of the air conditioning, the aircraft must descend to 4,267 m (14,000 ft, $\delta = 0.5875$), for which about 4 minutes are required. In order to maintain air supply to the passengers and the crew, emergency oxygen masks are installed for emergency.

2.1.7 Non-Standard Conditions

The ISA values for the atmospheric parameters are useful to compare aircraft performance over all the range of atmospheric altitudes. In fact, most of the calculations shown in this book are done under ISA conditions. Obviously, this is an idealized case that does not occur in practice. It is not uncommon to encounter temperature inversions, e.g. cooler air at the ground level and warmer air at low altitudes, contrary to the standard model. A comprehensive introduction to weather processes and climatic conditions around the world has been provided by Barry and Chorley.

A detailed performance analysis requires consideration of large deviations from the standard values, to deal with extreme environmental conditions: winters in the northern hemisphere, very hot weather on the ground. In addition, airport altitude, humidity and precipitations, atmospheric winds, lateral gusts, and global air circulation have strong influence on flight and safety. Rain and snow can be so heavy that take-off may have to be aborted. To simplify these matters, the US Department of Defense defines four non-standard atmospheres, referred to as hot, cold, tropical, and arctic (MIL-STD-210A). These profiles are shown in Figure 2.1-8.

Three important classes of weather-related flight problems are icing, downbursts and atmospheric turbulence. Ice accretion on the lifting surfaces during cruise can give rise to loss of longitudinal control; at take-off it may lead to loss of lift, stall angle reduction, drag penalties and longer ground run. Ice formation on the ground and in flight is discussed by Asselin. The two relevant reviews on the physics and modeling of icing have been provided by Lynch and Khodadoust, and Kind and co-workers.

The downburst is a weather phenomenon that produces a downward flow of great danger, particularly during take-off and landing. It is created by heavy winds, with speeds up to 30 m/s. These downbursts generally consist of closely spaced small cells (a few km wide) that are separated by relatively calm atmosphere. They act as jets flowing downward and spreading radially from the ground. A basic physical model of the downburst is discussed in detail by Zhu and Etkin. A vortex-ring model has been proposed by Ivan. Take-off and landing simulations with a downburst have been published by Hahn, while Zhao and Bryson optimized the flight path of an aircraft under a downburst with a non-linear feedback control program. The measurements near the airport and dynamic modeling have been discussed by Frost and co-workers.

Turbulence is a more familiar weather pattern to the frequent flyer. It includes cases of free air and convective air turbulence, atmospheric boundary layers and mountain ridge waves. No airplane is immune to the powerful gusts of

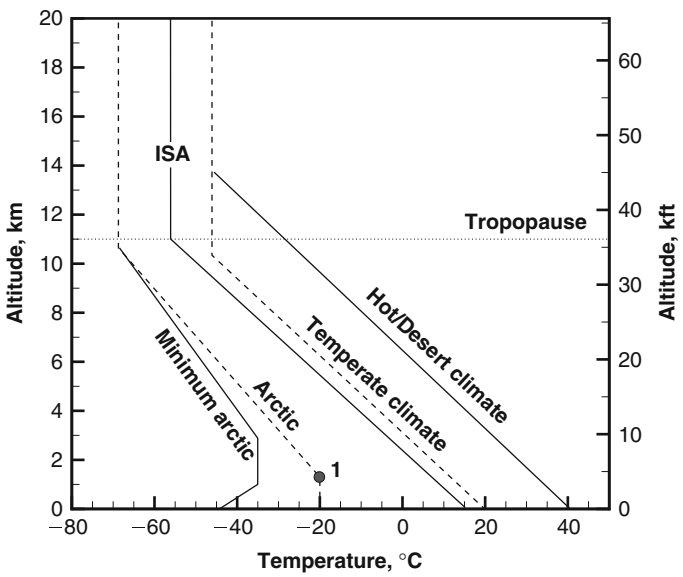


Figure 2.1-8 ICAO standard atmospheric temperature and reference atmospheres for flight mechanics calculations.

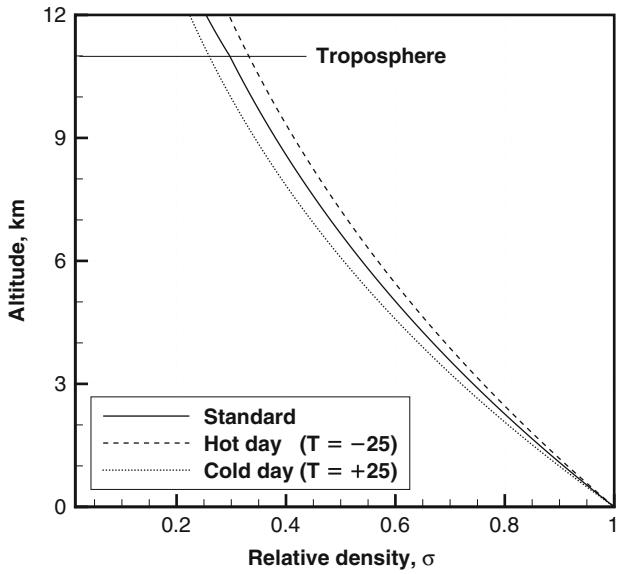


Figure 2.1-9 ICAO standard density and deviations due to ±25 degrees around the standard value.

the atmosphere, not even airplanes the size of the Boeing B-747. A thoughtful discussion of turbulence and flight has been provided by Etkin and Houbolt.

Among the global circulation effects there is the jet stream. This is an atmospheric wind with an West–East prevalence, strongest around the troposphere, that affects transatlantic flights between Europe and North America. The jet stream can add or subtract 1 flight hour on such flights.

Temperature variations can be of the order of 80 degrees (–40°C to +40°C). Since the temperature does not appear in any of the performance equations, a useful relationship with pressure and density is required.

If the temperature has a constant deviation from the standard value, say a constant $\pm\Delta\mathcal{T}$, the method of §2.1.6 can still be used, because the temperature gradient is the same (Eq. 2.1.20). The only difference is that the symbol $\mathcal{T}_o$ denotes the sea level temperature, whatever that may be. The result of such a model is shown in Figure 2.1-9 for the relative density.

For a temperature profile such as the Arctic temperature in Figure 2.1-8, a solution method is the following. All the quantities denoted with $(.)_0$ denote non-standard sea level condition, the quantities denoted with $(.)_1$are relative to point 1 in Figure 2.1-8. Starting from Eq. 2.1.27, we have

$$\int_o^h \frac{dp}{p} = -\frac{g}{\mathcal{R}} \int_o^h \frac{dh}{\mathcal{T}}, \tag{2.1.37}$$

$$[\ln p]_o^h = \ln\left(\frac{p}{p_o}\right) = \ln \delta = -\frac{g}{\mathcal{R}} \int_o^h \frac{dh}{\mathcal{T}}, \tag{2.1.38}$$

$$\ln \delta = -\frac{g}{\mathcal{R}}\left(\int_o^{h_1} \frac{dh}{\mathcal{T}} + \int_{h_1}^h \frac{dh}{\mathcal{T}}\right) = -\frac{g}{\mathcal{R}} \frac{h_1}{\mathcal{T}_o} - \frac{g}{\mathcal{R}} \int_{h_1}^h \frac{dh}{\mathcal{T}}, \tag{2.1.39}$$

$$\ln \delta = -\frac{g}{\mathcal{R}} \frac{h_1}{\mathcal{T}_o} + \frac{g}{\mathcal{R}\lambda_1} \ln\left(\frac{\mathcal{T}_1 - \lambda_1 h}{\mathcal{T}_1 - \lambda_1 h_1}\right). \tag{2.1.40}$$

It is possible to show that if $h_1 \to 0$, then Eq. 2.1.40 is equivalent to Eq. 2.1.31 (see also Problem 10). An amendment to the standard program to calculate deviations from the standard atmosphere is the following.

Computational Procedure

1. Construct a table of actual temperatures $\mathcal{T}(h)$.
2. Calculate the relative pressure by solving numerically Eq. 2.1.38.
3. Calculate the relative density from Eq. 2.1.24.

Problems

1. Write a program that solves the equations of the International Standard Atmosphere and on output provide the ratios σ, δ and θ as a function of the altitude h.
2. For a given relative air density $\sigma = 0.2, 0.4, 0.6$, find the corresponding altitude in ISA atmosphere by solving Eq. 2.1.34 for the unknown h.
3. Plot the air viscosity as a function of altitude, by solving Eq. 2.1.36. Discuss the origin and the meaning of this equation.
4. Due to improper closing, the door of a certain aircraft is lost in flight. At the time of the incident the aircraft is flying at an altitude $h = 6{,}000$ m with a speed $U = 400$ km/h. Describe the effects on the aircraft, on the passengers and the cargo. Motivate your answer. Calculate the pressure drop in the cabin by using the data of the International Standard Atmosphere at the flight altitude.
5. Calculate the air mass in the atmosphere from sea level to 5,000 m using the ISA data. Compare this mass with the mass from sea level to 20,000 m. Hint: neglect the changes of surface area with the altitude (this leads to an error less than 0.1%), use the buoyancy law $dp = \rho g dh$; consider a constant value of the gravitational acceleration g; use Eq. 2.1.34 to evaluate the relative air density. (About 50% of the air mass above the Earth should be contained within the first 5,000 m from sea level, an indication that the pressure created by the higher layers forces the air closer to the Earth's surface.)
6. From the equation of ideal gases, the relationship between the relative density, pressure and temperature is $\delta = \sigma\theta$. Find the parameter r such that $h = \sigma^r$ and

   $$\delta\sqrt{\theta} = \sigma^r$$

   which is valid in the troposphere.
7. Which aero-thermodynamic parameters are discontinuous at the tropopause? Can you explain the reasons for this discontinuity?
8. Concorde's cruise performance was found to be quite sensitive to the atmospheric temperature at its normal cruise altitude. Describe how the temperature can affect the cruise conditions and the fuel consumption, and find an engineering solution (*problem-based learning: additional research is required*).
9. Calculate the relationship between the leading-edge and the quarter-chord sweep angle for a wing with a straight leading edge. The taper ratio is λ, the tip and root chords are c_{tip} and c_{root}, respectively; the wing span is b, and the wing area is A.
10. Calculate the relative pressure, density and temperature for the Arctic profile (Figure 2.1-8), by using the method highlighted in §2.1.7. The reference point 1 has coordinates: $h_1 = 1.5$ km, $\mathcal{T}_1 = -70°$C. The temperature gradient is estimated by $\lambda_1 \simeq -0.0055°$C/m. Stop the calculation at the troposphere.

Aerospace Engineering Desk Reference; ISBN 9781856175753

2.2 Weight

Antonio Filippone

> *… future growth potential looks unlimited … one gross weight doubling, possibly two, is predicted by 1985; nuclear power can drive the optimum weight to 5 to 10 million pounds by the year 2000.*
>
> F.A. Cleveland, 1970

The aircraft's weight influences the flight performance more than any other parameter, including engine power. Weight has been of concern since the earliest days of aeronautics; just recall the fact that the first attempts to fly were based on lighter-than-air concepts. Therefore, we devote this chapter to the weight analysis and the relative aspects, such as weight definitions, useful loads, weight reporting formats, and the relative approximations. The discussion will be confined to the operational phase of the aircraft.

2.2.1 The Aircraft's Weight

By 1914, it was believed that the limiting weight of the airplane could not exceed 2,000 lb (about 800 kg). By comparison, Baumann wrote an alarming report in 1920, prompted by the construction of *giant airplanes* in Germany during the war–airplanes weighing as much as 15.5 tons, and powered by as much as 260 hp (195 kW) engines. Interest in large airplanes was sparkled by excitement in the very early days, with F. W. Lanchester and Handley Page expressing their views. Lanchester's opinions on this matter are recorded by Kingsford.

Cleveland's forecast in 1970 turned out to be wrong in the opposite direction; nuclear power has never been considered a serious option. Lockheed persevered along these lines for some years, and in 1976 Lange proposed aircraft concepts in the 900 ton (2 million lb) class, including a 275 MW nuclear power plant. By contrast, the design office at Boeing proposed the span-loader concept – a 1,270 ton aircraft without nuclear power (project 759).

In his paper "Quest for a Novel Force", Allen speculated on *antipodal megaliners,* monsters of the future capable of transporting 1,200 passengers from London to Sidney through a transatmospheric flight trajectory. The highly speculative content of Allen's paper is food for thought at a time when aircraft design is a conservative discipline. A further review, focusing on configuration alternatives and economic viability of the big airplanes, has been provided by McMasters and Kroo.

At the start of the 21st century, even the biggest airplanes do not exceed a gross weight of 600 tons (1.3 million lb). The Antonov AN-225, the largest prototype airplane ever built, can lift up to 250 tons of cargo at its design point. Its 88.40-m wing span is a wide as a football field. Its 18.1-m height reaches the top of a six-storey building. The airplane was designed to carry a spaceship as an external load.

The Airbus A-380 commercial liner has a maximum take-off weight of 562 tons and a wing span of 79.8 m. It is powered by four jet engines delivering a static thrust of over 1,000 kN. It can load up to 310,000 liters of fuel, and fly a distance of 15,000 km with over 500 passengers. The volume occupied by the fuel alone is a cube with a 6.8-m side. By way of comparison, the A-380's weight is equivalent to about 15 railway carriages, or five diesel locomotives – by all means a long passenger train.

Cleveland's analysis (recommended reading material to those interested in very large aircraft) contains a discussion of historical growth in size that leads to a *square-cube law.* The argument is that, if technology had not improved, growth would have been halted by the fact that the load stress in airplane structures increases with the linear dimension, when the load is proportional to the weight.

The concept is well illustrated by using as an example a single rectangular beam. The load is proportional to the weight W; the weight is proportional to the cube of the linear dimension l^3; the cross-section of the beam is proportional to the square of the dimension l^2. Therefore, in first approximation load/cross-section $\sim l$. At some point this increase in load reaches the structural limits of the material, and the beam collapses under the effect of its own weight.

While the Airbus A-380 has a wing span six times larger than the Wright Flyer (1903), a wing area about 16.5 times larger, the weight has increased by a factor of 1,650, which appears to defeat the law. The square-cube law would

imply that the wing loading on the A-380 would be the same as the Wright Flyer, or about 6.2 kg/m^2. By this rule, the weight of the A-380 should not exceed 5,237 kg! As it turns out, the actual weight is a factor 10^2 times the value calculated at *constant technology*.

If the wing is scaled up while holding wing loading and structural stresses as constant, its weight will grow roughly as $W^{1.4}$. However, when one looks at the details of the components, they do not scale up with the same factor. Cleveland showed that by doubling the gross weight and the payload of the aircraft the wing weight would have to increase by a factor of 2.69; the airframe would grow by a factor of 1.84; and the electrical systems would grow by a factor of 1.40.

Figure 2.2-1 presents a trend of aircraft maximum take-off weight (MTOW) and corresponding wing span b. The analysis shows that the wing span increases slower than the gross take-off weight, according to a function $W \simeq b^n$, with $n < 1$.

From a productivity point of view, the most important factor is not the absolute weight and size of the aircraft, but its useful payload. Historically, this has increased from about 10% of the Wright Brothers Flyer to over 30% of the current generation of airplanes; this ratio has also increased with the increased gross weight. The increase is driven by commercial requirements and by the need to move bulky equipment and machinery. By comparison, Concorde had a payload of less than 1% of its MTOW, though this is admittedly a completely different design concept.

For aircraft designed to operate in war zones the payload is essentially limited to ordnance. The amount of ordnance that can be carried varies greatly, depending on the type of operation. Figure 2.2-2 shows some estimated maximum loads for two categories of military aircraft. Figure 2.2-3 shows the payload ratio of several transport aircraft. The dotted line is a least-squares fit of the available data, and although the points are scattered, it indicates that the payload ratio increases with the size of the airplane. Figure 2.2-4 shows the payload ratio relative to the operational empty weight (OEW) and a least-square fit of the data (dotted line). Again the pay load performance increases with the size and weight of the airplane.

Table 2.2-1 summarizes the weight/payload data of the largest cargo airplanes currently in service. The data in the fourth column, R, indicates the maximum range at maximum payload.

To properly understand the sheer size of the vehicles in the table in terms of bulk load that can be carried, a Galaxy C-5B is reported to carry any of the following items: two Ml Abram battle tanks (61,700 kg each); six M2/M3 Bradley infantry vehicles; six Apache helicopters with folded blades; 113,740 kg of relief supplies; a 74,000 kg mobile bridge for the US Corps of Engineers, and all the arsenal of the US Army Ordnance.

A passenger Boeing B-747-400 can carry up to 416 paying passengers (or 7.5 coaches), and up to 216,840 liters of fuel. This is enough fuel to fill 4,300 mid-size cars, which would be able to cover a cumulative 3 million km journey. The Boeing B-747-400 at maximum take-off weight would be equivalent to lifting 10 train coaches or 300 mid-size cars at once. The payload/range diagram of this aircraft is shown in Figure 2.2-5, and was elaborated from Boeing's data.

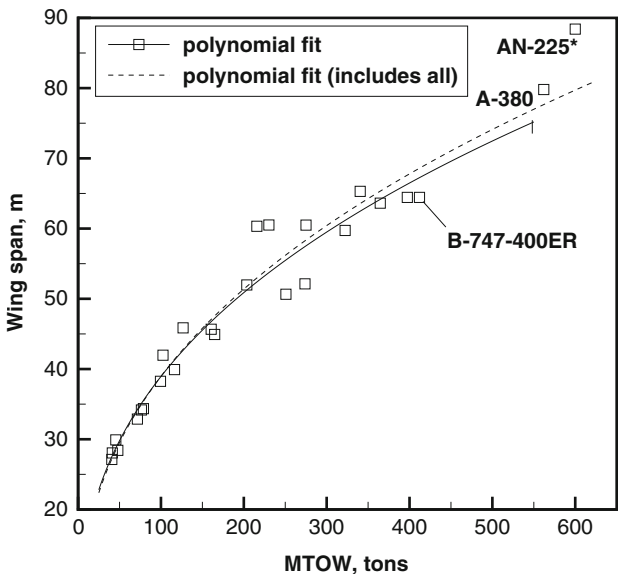


Figure 2.2-1 Wing span versus MTOW for large commercial aircraft; AN-225 at its design point.

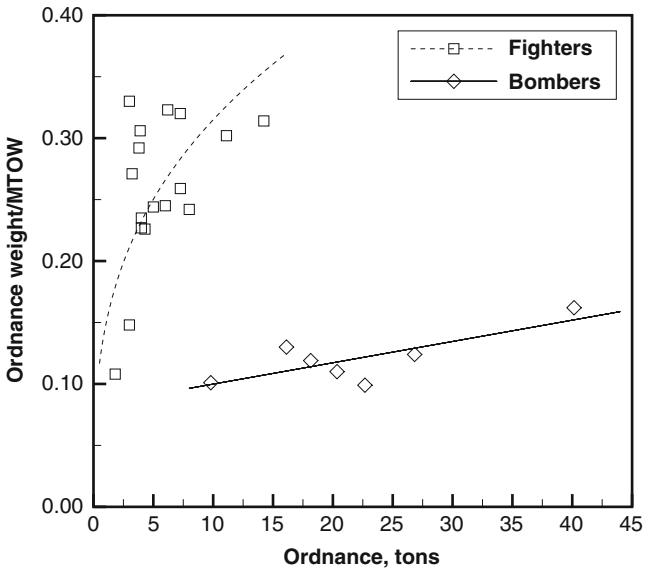


Figure 2.2-2 Maximum ordnance versus MTOW for fighter/attack and heavy bomber aircraft.

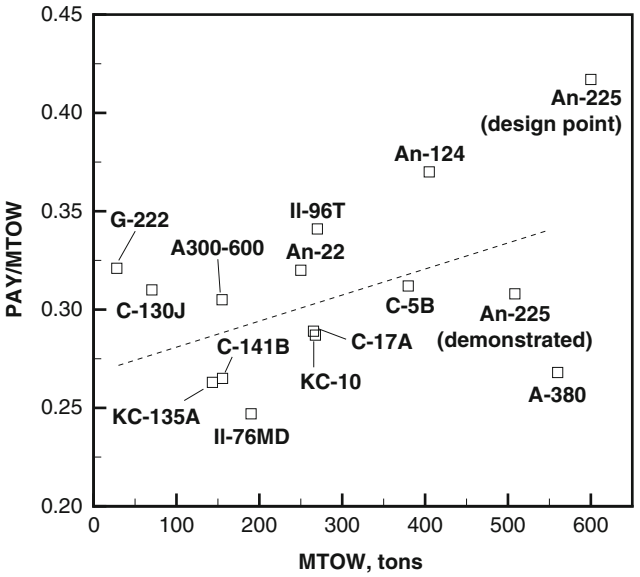


Figure 2.2-3 Cargo airplanes maximum payload PAY versus MTOW.

Freight is increasingly transported in standard containers called *pallets*. Loading and unloading of pallets is done quickly with ground-based vehicles with conveyors. The loading of pallets also rationalizes the available space in the cargo hold. A suitable cargo performance parameter is

$$E = \frac{PAY}{MTOW} R. \tag{2.2.1}$$

where R is the air range. This parameter is a measure of how much payload can be carried over a given distance. It emphasizes the fact that a certain payload can be flown over a longer or shorter distance; or that for a certain flight distance a larger or lower payload can be carried. This parameter has been estimated for a number of aircraft and is shown in Figure 2.2-6.

2.2.1.1 Wing Loading

In the previous discussion the concept of *wing loading* was used without much elaboration. The wing loading is the ratio between the aircraft gross weight and the area of the wing. This definition does not take into account the fact

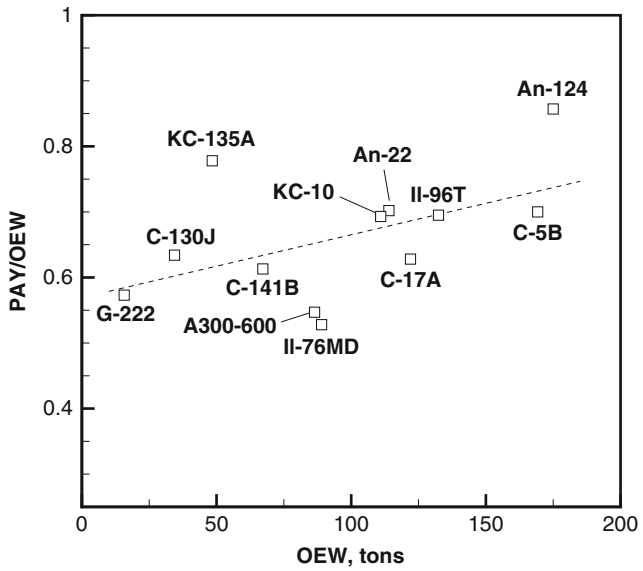


Figure 2.2-4 Payload ratio versus the OEW for some transport aircraft.

Table 2.2-1 Cargo aircraft PAY versus MTOW data (estimated); *R* is the range at maximum payload.

Aircraft	MTOW	PAY/MTOW	*R* (km)	Notes
Antonov AN-225, Ruslan	600.0	0.370	4,500	design
Lockheed C-5B, Galaxy	381.0	0.311	5,526	
Lockheed C-130J, Hercules	79.4	0.245	5,400	
Boeing C-17, Globemaster	264.5	0.288	4,700	
Boeing B-747-400F	396.8	0.284	8,240	
Boeing B-747-400ER F	412.8	0.290	9,200	
Airbus A-380-F	560.0	0.268	n.a.	
Satic A-300-600, Beluga	155.0	0.305	1,666	

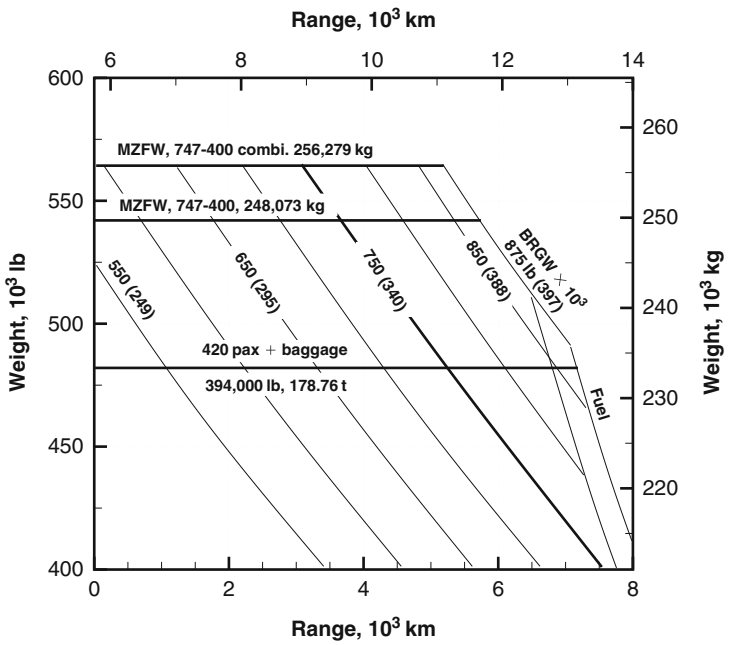


Figure 2.2-5 Payload ratio relative to the OEW of some transport aircraft.

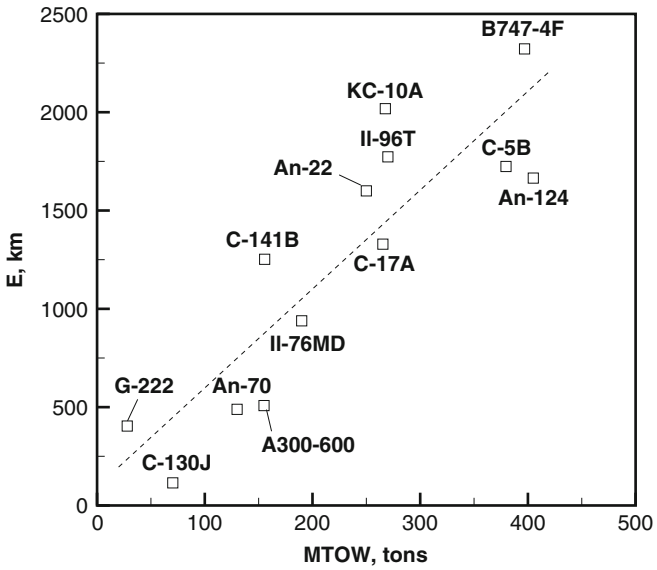


Figure 2.2-6 Cargo airplanes maximum range.

Table 2.2-2 Average wing loading *W/A* of different classes of aircraft (conversion between units is approximate).

Aircraft	kgf/m^2	Pa	lb/ft^2
Light aircraft	150–250	1,500–2,500	31–52
Turboprop transport aircraft	300–450	3,000–4,500	63–95
Business jet	300–450	3,000–4,500	63–95
Subsonic jet aircraft	350–650	3,500–6,500	73–137
Heavy-lift aircraft	450–750	4,500–7,500	94–158
Supersonic fighter	500–950	5,000–9,500	105–200

that a (small) portion of the lift may be created by the fuselage or the tail surfaces. Therefore, the wing loading implies that the aircraft's weight rests entirely on the wing. Examples of average wing loadings for different classes of aircraft are given in Table 2.2-2.

The wing loading of several aircraft has been plotted in Figure 2.2-7. The MTOW has been considered in the evaluation of *W/A*. The wing loading is not a constant; it decreases in flight, due to fuel consumption.

A comparison of interest is the one relative to bird flight. Figure 2.2-8, elaborated from data in Pennycuick and Greenewalt, shows a trend of birds' wing loading versus the birds' gross weight. Data are shown for over 100 birds, averaged between the two sexes; they make no distinction between soaring and flapping flight. The wing loading increases with the weight. The dotted line represents a linear fit of the data within the weight range considered, which can be considered as "equal technology". This is a case where the cube/square law is applicable. Flapping wings require higher wing loading at a given weight, and generally fall above the trend lines. For example, the razor bill (*Alca torda*) is indicated with a high point in the chart, and flies with rapid wing-beats. The same applies to the common loon (*Gavia immer*), a bird that has small wings, beating fast and steady; this bird is unable to soar or glide. The present trend line is compared to von Helmholtz's cube/square law, and shows only a little deviation. On the other hand, exceptionally good fliers like the bald eagle (*Haliaeetus leucocephalus*) and the golden albatross have a comparatively low wing loading, a sign of aerodynamic efficiency.

The interest in bird flight goes further back in time than aeronautics. The scale effects on birds and aircraft have been discussed by several authors.

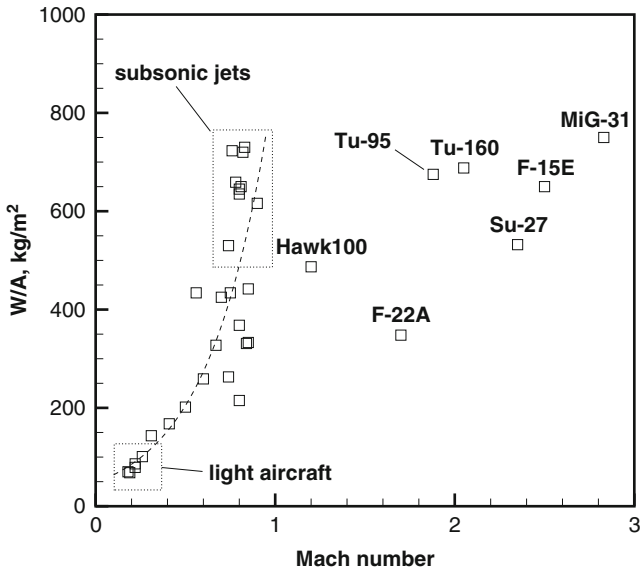


Figure 2.2-7 Wing loading of several classes of airplane (estimated).

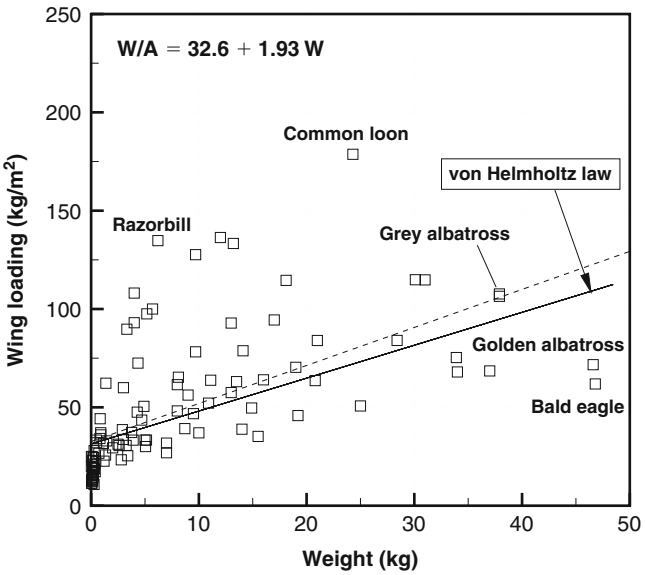


Figure 2.2-8 Wing loading of some birds, and trend lines, including von Helmholtz.

2.2.2 Definition of Weights

There are several definitions of aircraft weights that are part of the aircraft performance and operations. For some weights, the corresponding acronyms of wide use are reported. A comprehensive discussion of aircraft weights (including historical trends) is given by Staton and Torenbeek. The aircraft's weight has an effect on range, endurance, ceiling, climb rate, take-off and landing distances, maneuverability, to say nothing of direct operating costs and production costs.

The *Manufacturer's Empty Weight* (MEW) is the weight of the aircraft including non-removable items, such as all the items that are an integral part of the aircraft configuration. It is also called *dry* weight.

The OEW is the aircraft's weight from the manufacturer, plus a number of additional removable items, due to *operational* reasons. These are the engine oil, the unusable fuel, the catering and entertaining equipment, flight and navigation manuals, life vests and emergency equipment.

The MTOW is the maximum aircraft weight at lift-off, e.g. when the front landing gears detach from the ground. The MTOW includes the payload and the fuel. Sometimes there is a reference to a gross take-off weight (TOW); it is intended that this weight is an average value at take-off.

The *Maximum Payload Weight* (PAY) is the allowable weight that can be carried by the aircraft. The sum OEW + PAY is less than the MTOW, the difference being the fuel weight. The payload includes passengers and their baggage, bulk cargo, military weapons, equipment for surveillance and early warning systems. This weight is seldom given for passenger aircraft and high-performance military aircraft, though for different reasons. We need to differentiate between maximum payload based on weight (obviously), volume (space-limited payload, such as number of pallets) or capacity (due to seating limitations). The useful load is the sum between the payload and the mission fuel.

The *Maximum Landing Weight* (MLW) is the weight of the aircraft at the point of touchdown on the runway. It is limited by load constraints on the landing gear, on the descent speed (and hence the shock at touchdown), and sometimes on the strength of the pavement. Permissible loads on the pavement are regulated by the ICAO.

The difference between MTOW and MLW increases with aircraft size. For example, the Boeing B-777-200-IGW (Increased Gross Weight) has MTOW = 286,800 kg and MLW = 208,600 kg. This yields a difference of 78,200 kg, corresponding to 27% of MTOW. In extreme cases (landing gear or engine failure), fuel can be jettisoned for unscheduled landing. On the Boeing B-777 fuel is jettisoned through nozzle valves inboard of each aileron. This operation is done more routinely by military aircraft.

The *Maximum Ramp Weight* (MRW) is the aircraft weight before it starts taxiing. The difference between the maximum ramp weight and the maximum take-off weight, *MRW – MTOW,* corresponds to the amount of fuel burned between leaving the air terminal and lift-off. This difference is only relevant for very large aircraft, since it affects the operations on the runway. For example, the Boeing B-747-400 has an MRW = 398,255 kg and MTOW = 396,800 kg. The difference of 1,455 kg is the fuel that can be burned from the moment the aircraft starts taxiing and the take-off point, about 1,750 liters. This corresponds to the fuel tanks of about 35 mid-size cars!

The *Maximum Brake-Release Weight* (MBRW) is the maximum weight at the point where the aircraft starts its take-off run. The *All Up Weight* (AUW) generally refers to the weight of the aircraft in cruise conditions. Due to variations of weight, as a consequence of fuel burn, there is a change in AUW in flight.

Finally, the *Maximum Zero-Fuel Weight* (MZFW) is the aircraft weight on the ground without usable fuel, and the *Maximum Taxi Weight* (MTW) is the certified aircraft weight for taxiing on the runway. The latter is defined by the structural limits of the landing gear. The unusable fuel is the amount of fuel in the tanks that is unable to reach the engines under critical flight conditions, as specified by the aviation authorities.

There are, of course, the weights of the various aircraft components, such as the engines and the engine installation. For the engines a *dry weight* is sometimes reported, which indicates the weight of the engine without usable lubricant and fuel.

The large take-off weight involved in some airplanes also has consequences on the type of paved runway. For weights above 120 tons, the thickness of the pavement must be around 0.6 m. A thinner pavement would not be capable of sustaining a take-off and landing of such a massive aircraft. Other operational weight restrictions exist, such as the maximum weight for balanced field length, obstacle clearance, noise emission, and available engine power.

2.2.3 Weight Estimation

One may wonder how the aircraft's weight is measured or estimated, particularly for the very large aircraft. In fact, there is no balance than can hold a Boeing B-747 or an Airbus A-380 at their maximum take-off weight.

There are four classes of methods: conceptual (or component) weight estimation, statistical methods (based on previous aircraft design), quasi-analytical methods, and analytical methods. The aircraft's operating empty weight is estimated from the components method, e.g. by summing up the weight of its systems. These are: the wing system (with the control surfaces), fuselage, horizontal and vertical tails, landing gear, power plants, hydraulic and electrical systems, pneumatics, air conditioning and pressurization systems, auxiliary power units, instruments, and furnishings. The latter item is very much dependent on the operator of the aircraft. Therefore, the empty weight tends to change accordingly. These methods are important in aircraft design. What is essential in the present context is the weight management for the operation of the aircraft.

For commercial aviation, the operational weight of the aircraft includes the fuel, payload, and flight crew. For passenger operations, the weight of each passenger is calculated as 95 to 100 kg (including baggage); the volume to be allocated for baggage is 0.015 to 0.018 m^3. Methods for estimating the empty weight are given in a number

Table 2.2-3 Weight breakdown for some representative aircraft. All weights in kg.

System	B-747-100	% weight	C-5A	% weight
Wing	40,200	11.450	37,048	11.234
Tail	5,417	1.543	5,592	1.696
Airframe	31,009	8.833	52,193	15.826
Landing gear	14,596	4.157	17,046	5.169
Nacelle	4,703	1.340	3,838	1.164
Propulsion system	4,352	1.240	3,087	0.936
Flight controls	3,120	0.889	3,143	0.953
Auxiliary power unit	815	0.232	484	0.147
Instruments and navigation	674	0.192	333	0.101
Hydraulics and pneumatics	2,296	0.654	1,956	0.593
Electrical system	2,404	0.685	1,495	0.453
Avionics	1,873	0.534	1,871	0.567
Furnishings	21,748	6.195	3,539	1.073
Air conditioning	1,647	0.469	1,179	0.357
Anti-icing	188	0.054	106	0.032
Load and handling system	104	0.030	124	0.038
Operating empty weight	134,934	38.434	133,028	40.338
Dry engine weight	16,173	4.607	13,137	3.984
Empty weight	151,106	43.041	146,164	44.321
Take-off gross weight	351,076	100.0	329,785	100.0

of textbooks on aircraft design. Comparisons of weight breakdown for a number of aircraft, such as in Table 2.2-3, were published by Beltramo and co-workers.

Table 2.2-3 shows the weight breakdown from two large aircraft, earlier versions of the B-747 and the C-5. The main difference is in the weight of the furnishings, which is obvious, since the B-747 is a passenger airplane, and the C-5 is a military utility airplane, designed to carry bulky cargo. The weight of the airframe of the C-5 is strengthened for the same reasons. By comparison, the weight of the wing of the -400 version of the B-747 is 43,090 kg, an increase of 9% over the original wing, but its contribution to the total weight is slightly below 11% – an indicator of the technological advances in wing design.

2.2.4 Weight Management

The weight of almost all aircraft grows over time during their service life. Weight grows due to a number of reasons, namely new performance specifications, re-engineering of the power plant, exploitation of structural design margins, and not least the correction of design flaws, which may come after several years of service.

Weight and balance logbooks are maintained to keep a check on all the modifications done to the aircraft. The manufacturers also provide charts showing the basic weights and position of the center of gravity. Loading of commercial aircraft is done according to the instructions provided. The airline flight management performs basic calculations of passengers and baggage, by assuming a uniform loading of the aircraft. There are models that provide rapid solutions to the aircraft weight and balance as a function of passengers, baggage and fuel.

Generally, fuel consumption must obey special priorities to maintain the balance of forces on the aircraft at different flight regimes. Fuel must be used from the inboard tanks first. Optimal distribution reduces the requirements on aircraft trim, and therefore the drag associated with it, thereby maximizing the profitability of the aircraft.

Concorde was a special case also from the weight management point of view. Its 11 fuel tanks were distributed forward, centrally and aft of the aircraft. The transition from subsonic to supersonic cruise moved the center of pressure about 2 m aft. This movement required trimming of the aircraft, which in subsonic aircraft can be done by operating on the control surfaces. However, at supersonic speeds the trim drag would be an unacceptable penalty, and therefore trimming was done by moving the center of gravity aft by pumping fuel from the forward tanks to the rear tanks. Up to 33,000 kg of fuel can be pumped back and forth to trim the aircraft during subsonic to supersonic transition.

Commercial airlines operate scheduled and unscheduled passenger services, that carry checked-in baggage, cabin luggage and a number of items for passenger comfort. These are primarily food, drinks, magazines, television sets and other entertainment items. The operational weight refers to the weight of the aircraft fully equipped for these operations. This is somewhat higher than the OEW. To reduce this operational weight, some airlines have stopped offering meals altogether; others provide lunch packs to be collected upon boarding.

2.2.5 Range/Payload Diagram

The distance that can be flown by a given aircraft depends not only on the size of its fuel tanks, but also on the weight of its payload. As anticipated, in most cases the combination of maximum payload and maximum fuel load exceeds the MTOW. If one knows how to calculate the aircraft range, then it is possible to construct a chart showing how some of the aircraft weights are related to the aircraft range. Since different cruise techniques are available, in principle we could construct different weight/payload diagrams, corresponding to each of the flight programs. Furthermore, there are effects of atmospheric conditions, climb and descent techniques, reserve fuel policy that can change the weight/payload performance considerably. Unless all these conditions are specified, it will not be possible to compare the range/payload performance of two aircraft.

We are interested in range at MTOW and range at maximum PAY, plus some intermediate cases. The range difference is considerable, particularly for cargo aircraft. Consider the Airbus A-300 Beluga, which transports heavy aerospace equipment between industrial plants. Its range at maximum PAY is 1,660 km (about 900 nm); it increases to 2,780 km (1,500 nm) with a 40 ton PAY, and to 4,630 km (2,500 nm) with a "small" 26 ton payload.

Figure 2.2-9 shows the range/payload diagram for three commercial subsonic jet aircraft of the Airbus family.

Start the analysis by considering the following equivalence:

$$\frac{W_e}{W}+\frac{W_p}{W}+\frac{W_f}{W}=1, \tag{2.2.2}$$

where W_e is the OEW, W_p is the PAY and W_f is the fuel weight. From Eq. 2.2.2 the fuel fraction $\xi = W_f/W$ can be written as

$$\xi = 1-\frac{W_e}{W}-\frac{W_p}{W}. \tag{2.2.3}$$

Since the empty weight is fixed, Eq. 2.2.3 is a linear relationship between payload fraction and fuel fraction, with the gross weight W being a parameter. However, the fuel ratio as a function of the aircraft gross weight is different, and is shown in Figure 2.2-10 for the aircraft model B, at three values of the payload weight.

Sometimes a fourth term appears in Eq. 2.2.2 – the engine weight ratio, e.g. an item separated by the operational empty weight. This specification is not useful in discussing the range capability of the aircraft, though it is of interest when considering the efficiency of the propulsion system.

Figure 2.2-11 shows a complete summary of the weight/pay load range. The chart shows the range and weight limits due to payload and fuel.

The longest range is achieved at zero payload and full tanks, which corresponds to a gross weight less than MTOW. This is called *ferry range* and determines the maximum distance that an aircraft can fly non-stop and without in-flight refueling (range at maximum fuel load). The minimum range is achieved at maximum payload and it is shown how the range at maximum payload is less than half the ferry range (this value is typical of other aircraft). Also of interest is

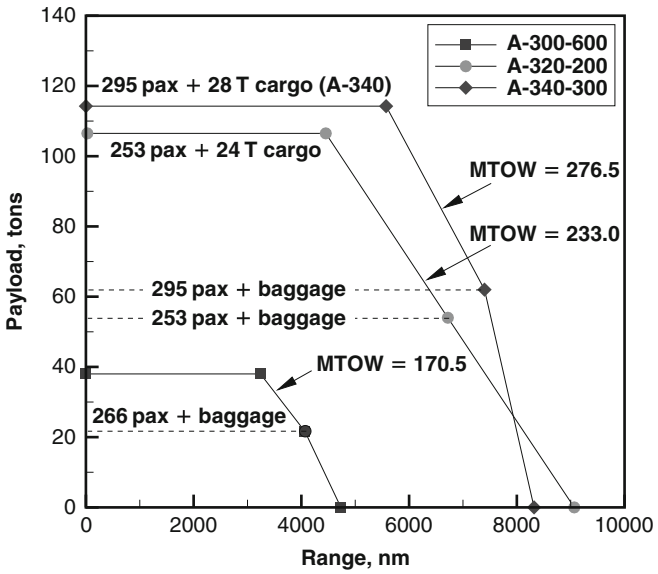


Figure 2.2-9 Maximum payload range for Airbus airplanes.

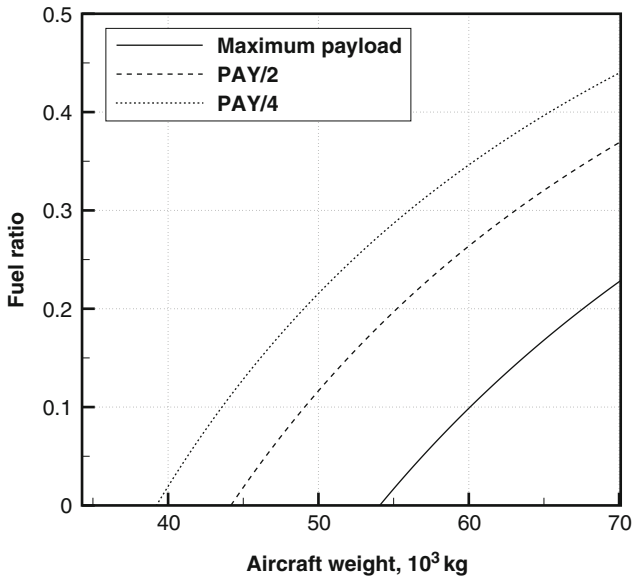


Figure 2.2-10 Fuel ratio for aircraft model B.

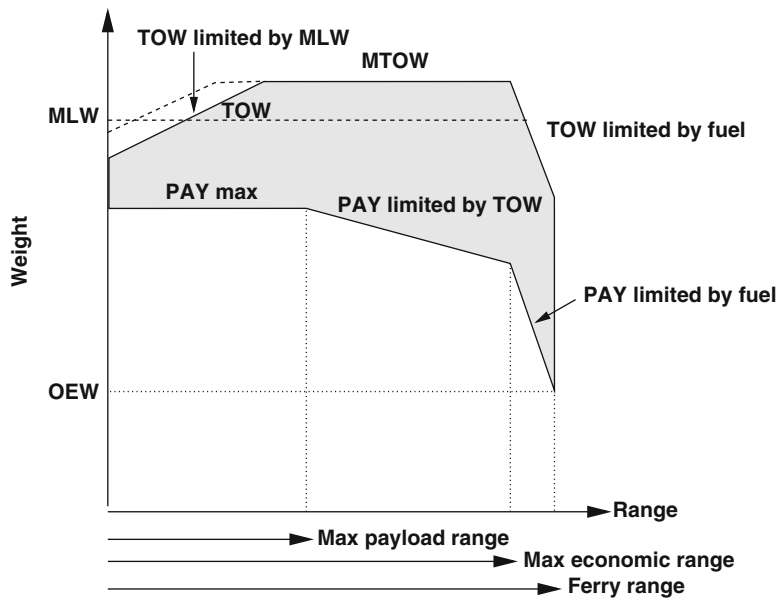


Figure 2.2-11 Weight/range chart for a commercial airplane.

the aircraft range at maximum fuel and maximum take-off weight, which requires the payload to be considerably less than the maximum.

First, consider an aircraft at MTOW configuration. If the mission fuel increases while keeping the aircraft weight constant with some payload, the range increases linearly with it. When the fuel tanks are full, the aircraft is prepared for maximum range at that weight. To fly a longer distance, the aircraft must have a gross weight less than MTOW. The minimum weight will be obtained by flying without useful payload.

There is a short range that can be flown at maximum payload, a medium range in which the payload is limited by MTOW, and a long range in which the payload is limited by the fuel capacity. In the short range flight plan the take-off weight increases linearly to the MTOW, it remains fixed at MTOW for a medium range flight, and decreases due to limited fuel capacity in the long range cruise. Although the diagram starts from a range equal to zero, there is a minimum range where the take-off weight is limited by maximum landing weight.

2.2.6 Direct Operating Costs

Direct Operating Costs (DOCs) are the costs incurred by the owner of an aircraft to operate scheduled or unscheduled flights. It includes the cost to fly, insure and maintain the aircraft airworthy. The aircraft costs money even if it stays on the ground. Characterization of these costs is difficult. An example of analysis is given by Beltramo *et al.*[78], who developed cost and weight estimating relationships and weight estimating relationships for commercial and military transport aircraft. Even the parametrization of the fuel costs is aleatory, due to the costs of aviation fuel on the international markets, and to the cost of fuel at different airports around the world.

Kershner[79] has demonstrated that while DOCs have consistently decreased over time, the impact of the fuel cost has remained high, and in some historical circumstances has increased to over 50%.

The appearance of *budget airlines* on the major world markets has contributed to a substantial change in the structure of the DOCs. It is sometimes exciting to learn that we can buy a ticket to an international destination within Europe or the continental USA at a cost comparable to the cost of this book. Cost items such as ticketing, customer service, seat allocation, baggage handling, ground services, in-flight catering, airport taxes, leasing contracts, etc. have been dissected, reduced or removed altogether from the DOCs. However, the cost of the fuel remains essentially the same. This cost is an essential part of the aircraft performance.

Problems

1. The load on a wing is proportional to the aircraft's gross weight. Describe how and why the stress on the structure increases with the linear dimension of the wing.
2. Discuss the limits to aircraft weight growth, including structural, aerodynamic, propulsion, landing gear, and systems limits; handling qualities, ground services, air traffic control, and costs.
3. You are required to study the feasibility of a military aircraft that has a radius of action of 1,500 km and can carry 5,000 kg of ordnance. The maximum take-off weight cannot exceed 20,000 kg. Assume a reserve fuel ratio equal to 5% of the mission fuel, and a range constant $c = 16{,}000$ (based on average performance of similar military aircraft).
4. For aircraft model B plot the block fuel ratio ξ as a function of the payload fraction W_f/W, where W_f is the fuel weight, W is the aircraft gross weight, for a weight up to MTOW. Discuss the effects on increasing the payload fraction on the mission fuel, and hence on the range.
5. Discuss reasons why the sum of maximum payload weight and maximum fuel weight generally exceeds the maximum take-off weight of an aircraft.
6. Investigate the items of direct operating costs of a modern commercial subsonic jet. Estimate by further research the percentage of each item on the overall costs, and the effects of increasing the fuel cost (*problem-based learning: additional research is required*).

Aerospace Engineering Desk Reference; ISBN 9781856175753

2.3 Aerodynamics

Antonio Filippone

> *The aircraft [replacement of the SR-71A] has been seen moving at high supersonic speed, with the resultant sonic bangs, over Southern California. It is believed to be powered by a revolutionary new engine which leaves a distinctive "sausage-string" shaped contrail at high altitude, coupled with an unmistakable sound.*
>
> The Encyclopedia of World Aircraft, 1997

Aerodynamics is one of the fundamental aspects of aircraft flight. Most books on flight mechanics and performance have prominent chapters on the aerodynamics of the wing and the wing section. In view of the amount of topics to deal with in performance analysis, it will not be possible to review airfoil and wing characteristics. This chapter deals with the aerodynamics of the aircraft as a whole, particularly drag and lift characteristics as a function of the main parameters.

2.3.1 Aerodynamic Forces

The aerodynamic forces and moments are not directly contributed by the propulsion system. For a symmetric aircraft in level flight, the aerodynamic forces are applied somewhere on the symmetry plane. For first-order calculations, there is a general consensus that these forces are applied at the center of gravity, as discussed in Chapter 2.1. In reality, the aerodynamic forces are applied at the center of pressure, which is dependent on the flight Mach number. As the aircraft accelerates past the speed of sound, the center of pressure tends to move aft. The opposite happens during deceleration to subsonic speed. The shift of the center of pressure and the consequent stability problems can be solved by pumping fuel aft and fore, as discussed by Leyman for the Concorde (Problem 1). An important calculation method for the center of pressure of wing/body and wing/body/tail combinations was developed by Pitts and co-workers for bodies at subsonic, transonic, and supersonic speeds.

The aerodynamic force component in the direction of the velocity vector is called *drag*. The force component normal to the drag and pointing upward is the lift. The magnitude of these forces is, respectively,

$$D = \frac{1}{2} C_D \rho A U^2, \tag{2.3.1}$$

$$L = \frac{1}{2} C_L \rho A U^2, \tag{2.3.2}$$

where C_D and C_L are dimensionless force coefficients, A is a reference area, and U is the air speed. The reference area is the *wing area*, e.g. the area of the neutral wing projected on the ground and including the portion inside the fuselage. The wing area is not a constant, because of the use of high-lift devices in take-off and landing operations. However, it is calculated with the control surfaces in the neutral position. The presence of additional aerodynamic surfaces, such as horizontal tail and canards, affects the position of the lift force, the magnitude and the share of lift generated by the aircraft's subsystems.

Two important non-dimensional quantities are the Reynolds number and the Mach number. The Reynolds number is the ratio between the inertial and viscous forces, $Re = \rho U l/\mu$, with l a reference length in the flight direction. The Mach number is the ratio between the air speed and the speed of sound, $M = U/a$.

The force coefficients depend on the Mach number, on the Reynolds number, on the angle of attack, on the geometrical shape of the aircraft, and on whether the flow is steady or transient. For a given aircraft configuration

$$C_L = C_L(\alpha, Re, M, t), \qquad C_D = C_D(\alpha, Re, M, t). \tag{2.3.3}$$

In cases of engineering importance, the influence of the Reynolds number is considerably smaller than the remaining parameters, and is generally neglected. More specifically, this influence can be characterized as a slight decrease in C_D with the increasing speed, due to boundary layer effects. Unless the aircraft is maneuvering at high angles of attack, the C_L and C_D are taken from steady state operation; hence, the time is excluded from the functional parameters. Therefore, we will reduce Eqs 2.3.3 to

$$C_L = C_L(M, \alpha), \qquad C_D = C_D(M, \alpha). \tag{2.3.4}$$

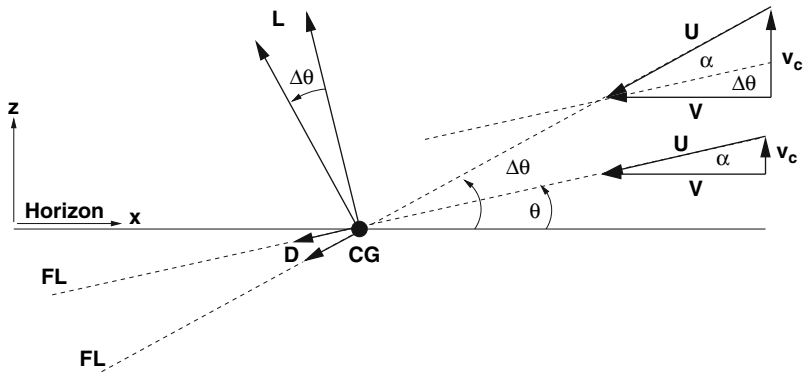


Figure 2.3-1 Aerodynamic forces, aircraft attitude and angle of attack in the vertical plane (angles exaggerated for clarity).

When the coefficients cannot be expressed in analytical terms, they are tabulated as a function of Mach number and angle of attack.

It is important to understand that the direction of D and L changes with the direction of flight. If the aircraft is flying on a flight path contained in a plane normal to the ground, its position can be characterized by the position of its CG and its attitude γ (three degrees of freedom). The angle between the velocity vector and the longitudinal axis of the aircraft is the nominal angle of attack. The same value of the angle of attack corresponds to an infinite number of attitudes. As shown in Figure 2.3-1, if the attitude increases, the velocity vector can be adjusted by augmenting the climb velocity by

$$\Delta v_c = V \tan \Delta\theta. \tag{2.3.5}$$

Under these conditions the angle of attack remains constant. The graph also shows how the direction of the aerodynamic forces is rotated by $\Delta\theta$.

An important aerodynamic parameter is the zero-lift angle of attack, α_o. This is the angle between the velocity vector $\mathrm{V} = (V, v_c)$ and the aircraft's longitudinal axis FL at which the lift force vanishes. The zero-lift angle makes an angle, generally small, with the fuselage line.

Assume an aircraft flying in the vertical plane, as in Figure 2.3-2. The aircraft must rotate around the CG so as to have the fuselage line FL coincident with $V = (V, v_o)$ for the lift to vanish. At attitudes above this, the aircraft will have a positive angle of attack. If a perturbation is introduced to α, the direction of the lift and drag force does not change.

2.3.2 Lift Equation

From aerodynamic theory, the lift coefficient can be written as

$$C_L = C_{L_o} + C_{L_\alpha}(\alpha - \alpha_o), \tag{2.3.6}$$

where C_{L_o} is the zero-angle of attack lift coefficient and $C_{L_\alpha} = dC_L/d\alpha$ is the lift curve slope. At low Mach numbers and low angles of attack C_{L_α} is constant. As the incidence is increased, a noticeable difference with the linear behavior

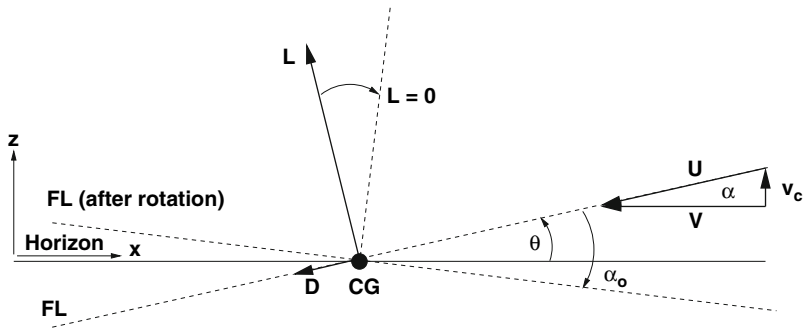


Figure 2.3-2 Zero-lift conditions on the aircraft (angles greatly exaggerated).

starts to appear; C_{L_α} may decrease or increase. A case of incipient wing stall is reflected in a decrease of the C_{L_α}. Flow separation starts at the trailing edge and at the outboard regions of the wing; it gradually expands to cover most of the wing surface. When C_{L_α} increases, it is generally a case of vortex lift, discussed further in §2.3.3. Most airplanes operate at relatively small angles of attack; therefore, the linear assumption is adequate for most performance calculations. However, modern fighter jet aircraft are capable of operating at very large angles of attack; this involves a great deal of unsteady separated flow around the wing, in which case Eq. 2.3.6 is a useful representation of the lift only if the non-linear function $C_{L_\alpha}(\alpha)$ is known.

The lift-curve slope reflects the general lifting capabilities of a wing, because it is related to its aspect-ratio $\mathcal{AR}$, hence to its main geometrical characteristics. A useful equation between C_L and wing aspect-ratio for low-speed low angle of attack flight is given by

$$C_{L_\alpha} = \frac{2\pi}{1 + 2/\mathcal{AR}}(\alpha - \alpha_o), \tag{2.3.7}$$

which is valid only for elliptic spanwise loading. Equation 2.3.7 allows the construction of the three-dimensional wing lift from the airfoil and the geometry of the wing. This equation does not contain the angle of attack; therefore, it shows as a straight line in the α–C_L graph.

2.3.3 Vortex Lift

One of the sources of non-linearity in the lift equation is the *vortex lift*. At high angles of attack the wing flow is dominated by two large counter-rotating vortices developing from leading-edge separation. These vortices are always associated to low pressures on the upper side of the wing, and grow linearly or super-linearly downstream. They are often associated with secondary flow separation and augment the lift-curve slope in a variety of ways. Whatever the mechanisms of vortex lift generation, a performance analysis relies on lift curves that can be expressed in a compact form, such as Eq. 2.3.6. An example of vortex lift on the C_L is shown in Figure 2.3-3 for a Δ-wing with a 74 degree sweep. The wing has an aspect-ratio $\mathcal{AR} = 0.5735$. The data show that the leading-edge vortex increases the lift at a super-linear rate. The resulting C_L increasingly diverges from the "conventional" curve, $C_L = 2\pi\alpha$. Also, the stall angle is dramatically increased to about 40 degrees and is more gradual, sometimes yielding a nearly constant lift over a wider range of incidences. This is not uncommon for this type of wing. Therefore, it is an indication that the aircraft can maneuver at angles of attack beyond stall. Aircraft capable of maneuvering in post-stall conditions are called *super maneuverable*.

Angles of attack that can be sustained by modern military aircraft during high-power climb, turning, and aerobatic maneuvers can exceed 50 degrees. This is done to obtain a positional advantage during combat, although the advantage comes at the expense of some energy loss and some stability problems. Calculations under these

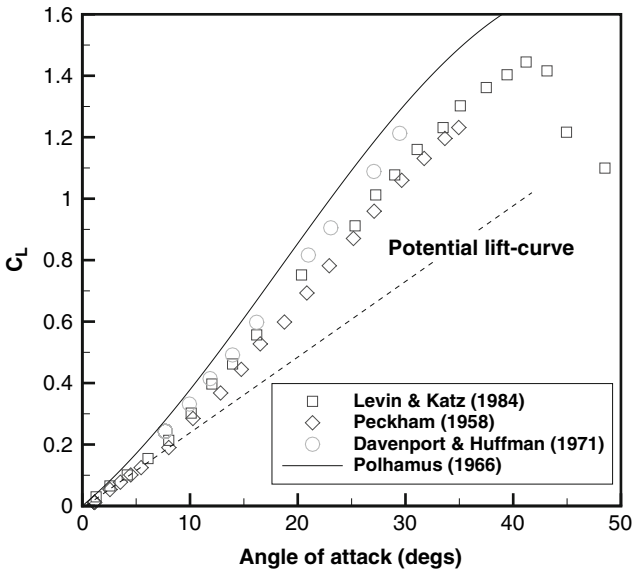


Figure 2.3-3 Vortex lift of a slender Δ-wing, sweep angle $\Lambda = 74$ degrees at the apex.

conditions can be done by considering average values of the lift and drag force, for example from tabulated data like those in Figure 2.3-3.

A relatively simple method, due to Polhamus, allows the calculation of the C_L of pointed Δ-wings at incidences below stall, and recovers some of the non-linearities of the vortex lift. The resulting equation is

$$C_L = k_p \sin\alpha \cos\alpha^2 + k_v \cos\alpha \sin\alpha^2, \tag{2.3.8}$$

where k_p and k_v are coefficients for the potential flow lift and vortex lift components, respectively. These coefficients depend on the aspect-ratio of the wing and their calculation has to be done by other means, for example a vortex lattice method. These coefficients have been calculated by Polhamus, and are shown in Figure 2.3-4. The value of the coefficients is important, because the calculation of the vortex lift from Eq. 2.3.8 can be done quickly in semi-analytical form. Therefore, a high-performance aircraft having such a wing would be relatively easy to simulate in a high-α maneuver.

The physical phenomena responsible for vortex lift are discussed in some detail by Peake and Tobak, and Delery. These authors, along with Ashley and co-workers and Gursul give an overview of the problem, mostly at low speeds. A good discussion of vortex lift at transonic and supersonic speeds is given by Wood and co-workers, which summarizes the research at NASA. Data for Δ-wings at supersonic speeds and angles of attack up to 47 degrees have been provided by Hill. The aerodynamics of the aircraft at high angle of attack is reviewed by Erickson, who discussed post-stall maneuver, yaw control, longitudinal stability and other problems.

Aerodynamic surfaces that promote leading-edge separation and control the evolution of the separation vortex are another reason for vortex lift. These additional surfaces are *canards*, *strakes*, leading-edge extensions, and double Δ-wings. Canard surfaces reach a value of up to 20% of the main wing (Grumman X-29), although they are more likely to be slightly above 10% (the Eurofighter 2000 has a canard surface less than 5% of the main wing). Wind tunnel and flight data on strakes and leading-edge extensions (lift and drag polar, longitudinal and lateral stability at angles of attack) are available in several publications and co-workers.

Figure 2.3-5 shows wind tunnel results of experiments with strakes empirically designed. The C_L curve is smoother; the performance results in lower C_D for a given C_L. The effect of the strakes includes a change in stability characteristics, with a positive pitching moment slope, instead of a negative one. This makes the aircraft unable to fly without proper control.

Example

Consider the pointed Δ-wing discussed in Figure 2.3-3. The aspect-ratio of the wing is $\mathcal{AR} = 2 \tan \Lambda/2$, with Λ the apex angle. Therefore, $\mathcal{AR} = 2 \times \tan 37 = 1.507$. The Polhamus coefficients for this case are estimated as $k_p \simeq 3.20$ and $k_v \simeq 1.82$. The corresponding lift curve is shown as a solid line in Figure 2.3-3.

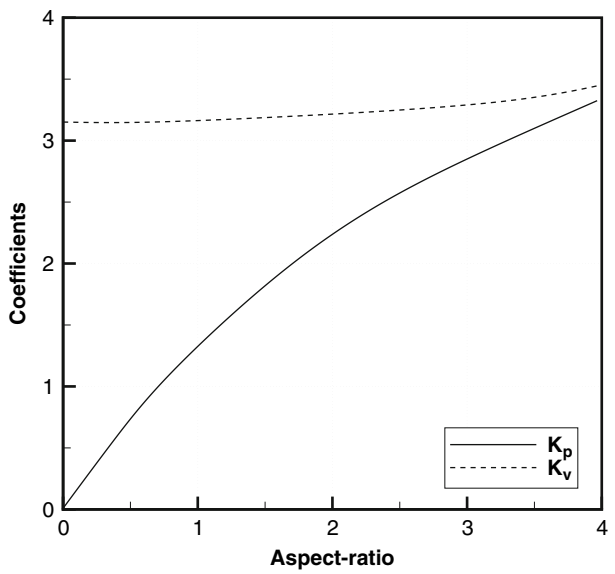


Figure 2.3-4 Polhamus coefficients for a pointed Δ-wing.

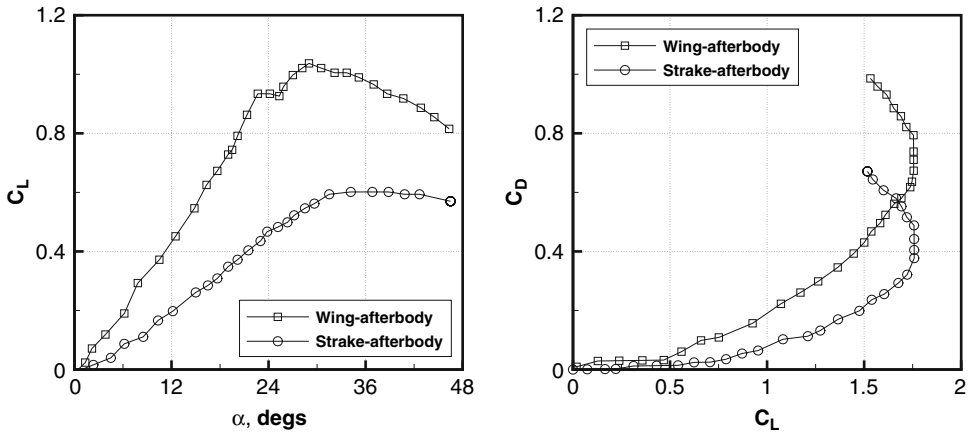


Figure 2.3-5 Wind tunnel data for strakes and wing/body combinations at high α.

2.3.4 High-Lift Systems

The term *high-lift system* denotes all the technical means for increasing lift during the terminal phases of the flight. Low take-off speeds and high climb-out rates are a general requirement for good airfield performance.

Most of the systems are unpowered, and consist of leading-edge slats and trailing-edge flaps in several segments. A single-slotted flap is used when the wing system is not capable of supplying a $C_{L_{max}}$ adequate for approach and landing. If this is still not enough for low-speed control, the landing flap has to be replaced by a double-slotted flap. Calculation of the maximum C_L obtainable from a high-lift system is a complicated matter. Therefore, most of the data available rely on practical experiments and on flight testing. An aircraft operating near its $C_{L_{max}}$ is at risk of wing stall. For take-off and landing operations there are practical rules that avoid operating at too high C_L. For safe operation we have to consider

$$C_{L_{approach}}/C_{L_{to}} \simeq 1.10\text{—}1.12, \tag{2.3.9}$$

and

$$C_{L_{to}} \simeq 0.7\text{—}0.75 C_{L_{max}}, \quad C_{L_{approach}} \simeq 0.6\text{—}0.65 C_{L_{max}}. \tag{2.3.10}$$

Therefore,

$$C_{L_{land}} \simeq 1.3 C_{L_{to}} \tag{2.3.11}$$

A summary of high-lift and control systems on current airplanes is shown in Figure 2.3-6, where we have plotted the maximum C_L against the mechanical complexity of the systems (arbitrary unit). To the left of the shaded bar is a summary of unpowered systems, which are operated by proper actuators; to the right are some examples of powered systems, which are operated with energy from the main power plant. The maximum C_L achieved by the unpowered systems seldom exceeds 3.0, though it is often below this limit. Some data are the following: the Fokker F-28 has $C_{L_{max}} \simeq 3.35$; the Boeing YC-14 has $C_L \simeq 3.57$.

In the top right of Figure 2.3-6 there are the high-lift performance data of some experimental aircraft, for example the NASA/Boeing QSRA (Quiet Short Haul Research Aircraft), flown in the 1980s. This aircraft required only a 300 meter runway. The jet engine exhaust was directed over the wing, which shielded the ground from jet noise during take-off. Figure 2.3-7 shows a typical C_L polar.

2.3.5 Drag Equation

Drag is a vast subject in aircraft aerodynamics. A typical chart of drag components on the aircraft is shown in Figure 2.3-8.

The drag equation specifies the functional dependence of the C_D on the main state parameters of the aircraft (Eq. 2.3.4). A general expression to the second order of the angle of attack is

$$C_D = C_{D_0} + \eta C_{L_\alpha} \alpha^2, \tag{2.3.12}$$

where C_{D_0} is the zero-lift drag coefficient (profile drag), η is an induced drag coefficient, and α is the angle of attack in radians. In Eq. 2.3.12 all the coefficients are a function of α, the Mach number. They are generally tabulated in the

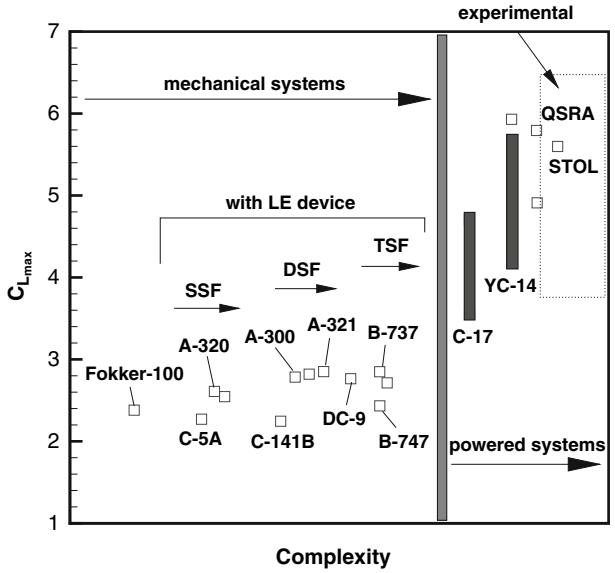


Figure 2.3-6 High-lift systems on current airplanes; QSRA and STOL are experimental airplanes. SSF = Single-Slot Flap, DSF = Double-Slot Flap, TSF = Triple-Slot Flap.

form $C_{D_0}(M), \eta(M), C_{L_\alpha}(M)$. It is assumed that the aircraft operates within the linear range of angles of attack, even at supersonic speeds. At subsonic speeds a common drag equation is

$$C_D = C_{D_0} + kC_L^2. \tag{2.3.13}$$

In this equation the angle of attack does not appear explicitly. This is the preferred drag equation for performance calculations of low-speed aircraft and commercial jets. In fact, due to the quadratic term C_L^2, Eq. 2.3.13 yields several useful closed form solutions. Also, the C_L can be easily associated to the aircraft's gross weight. Figure 2.3-9 shows some flight test data for the Douglas DC-10 passenger aircraft. The linear fit demonstrates that the parabolic drag equation 2.3.13 is valid over a wide range of lift coefficients, as they can be used in ordinary flight conditions.

For an estimate of the induced drag factor k an equation often used from low-speed aerodynamics is

$$k = \frac{1}{e\pi\mathcal{AR}}, \tag{2.3.14}$$

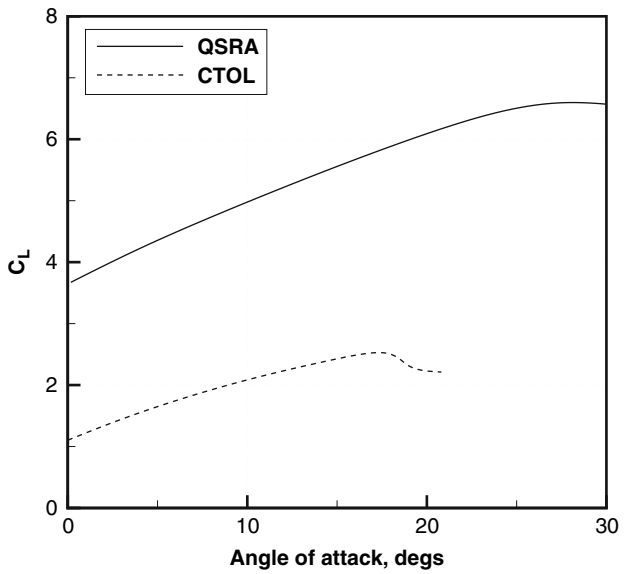


Figure 2.3-7 Lift coefficient of NASA QSRA compared with that of a conventional aircraft.

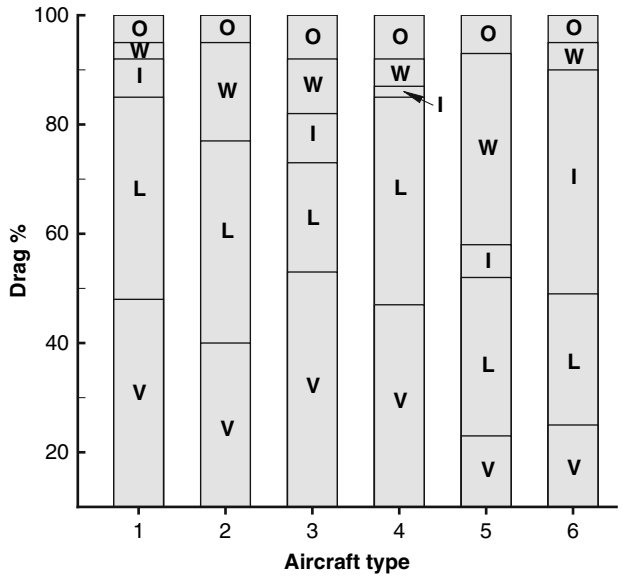


Figure 2.3-8 Contributing factors to overall drag. Nomenclature: (1) subsonic transport; (2) SST; (3) business jet; (4) fighter aircraft (subsonic); (5) fighter aircraft (supersonic); (6) helicopter; V = viscous/parasite; L = lift-induced; I = interference drag; W = wave drag; O = other causes.

where e is a factor variable between 0.74 and 0.88 (Oswald factor), which depends on the spanwise load distribution. For an ideally elliptically loaded wing $e = 1$. Equation 2.3.14 is useful for checking the validity of the coefficient k in the drag equation.

The C_L and C_D can be plotted against each other in a single graph, to yield the *drag polar*; for a given aircraft speed or Mach number. An example is shown in Figure 2.3-10, which is relative to the experimental aircraft Lockheed YF-16. The relationship between C_L^2 and C_D is not linear; therefore, the drag (Eq. 2.3.13) is not suitable for this aircraft.

2.3.5.1 Zero-Lift Drag

The zero-lift drag (profile drag) of the aircraft is the resistance due to viscous effects and other causes not directly related to the production of lift, such as the resistance of various subsystems (landing gears, probes, antennas, gaps, etc.). It depends on a number of factors, the Reynolds number and the surface roughness being the most important

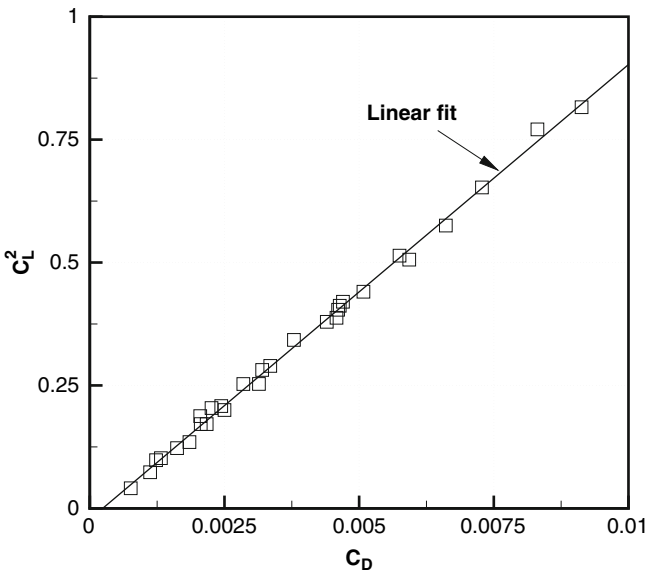


Figure 2.3-9 Drag equation for the Douglas DC-10.

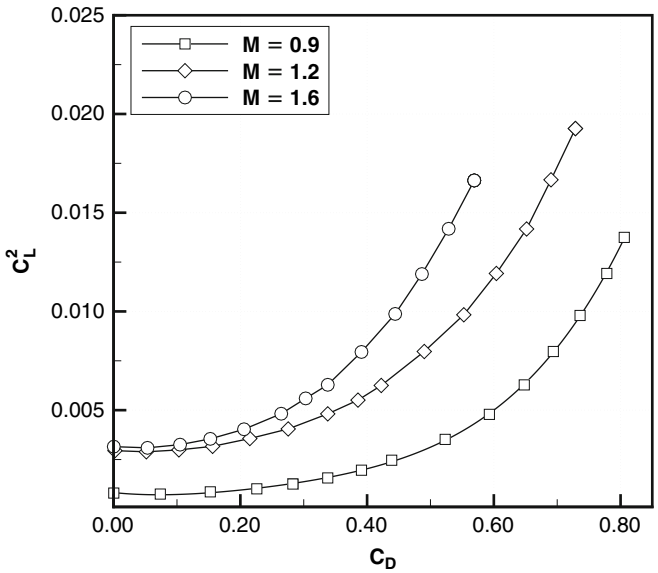


Figure 2.3-10 Drag polar of YF-16 at Mach numbers up to $M = 1.6$ (estimated data).

ones. The percentage of this drag on the total drag count depends on the type of aircraft and its operational configuration, and varies from 25–30% for a supersonic jet fighter to 70–80% in the case of a VTOL aircraft. For a commercial subsonic jet aircraft the contribution of the fuselage to the zero-lift drag is of the order of 30%. The skin friction drag of modern large aircraft is close to the theoretical value from turbulent flat plate theory at the corresponding Reynolds number.

The effect of Reynolds number can be evaluated from a number of semi-empirical relationships. At the flight Reynolds numbers these expressions are equivalent to each other.

Figure 2.3-11 shows the values of the skin friction drag coefficient due to a fully turbulent flow past a flat plate, as a function of the Reynolds number. The estimated C_{D_0} values of modern aircraft are shown in the box for comparison. The Reynolds numbers in these cases are calculated using the average wing chord as a reference length, and the air viscosity at the conditions of the cruise altitude.

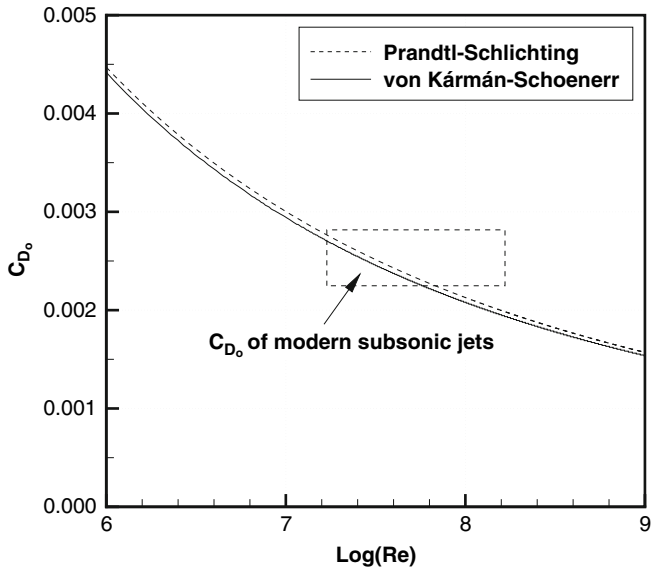


Figure 2.3-11 Estimated value of zero-lift coefficient of modern subsonic transport aircraft and comparison with theoretical values from flat plate theory.

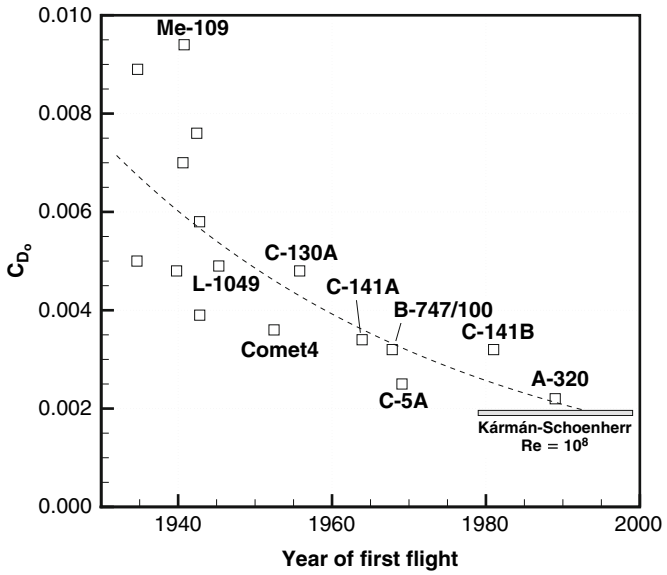


Figure 2.3-12 Profile drag coefficient of selected aircraft.

The streamlined design of modern aircraft shows that the current values are not too far from the drag obtainable by a corresponding flat plate at the same flight Reynolds number, which leaves little room for improvement, unless a new technology is developed (boundary layer control, flow control, etc.). The current values of the zero-lift drag coefficient for large commercial jets are of the order of 220 to 250 drag counts (1 drag count = 0.0001), and sensibly higher for smaller airplanes. By comparison, the Douglas DC-3 of the 1930s had 249 drag counts of skin friction drag – a very respectable value, which demonstrates that progress in this area has been relatively slow. A historical trend of the profile drag coefficient for selected aircraft is shown in Figure 2.3-12. The dotted line shows a trend toward the minimum theoretical value of a turbulent flat plate.

2.3.6 Glide Ratio

This parameter is the ratio between lift and drag; sometimes it is called *glide factor* or *aerodynamic efficiency*. The best glide ratios of current aircraft do not exceed 20, although they are contained within the 14 to 18 range. Some high-performance gliders with very large aspect-ratio wings have glide ratios in excess of 25. By comparison, the best low-drag wing section will have a glide ratio of the order of 150. The L/D values are known to decrease with the increasing Mach number.

The glide ratios of large aircraft at cruise conditions are not that different from those of the most efficient birds. It is estimated that the albatross *(Diomeda exulans)* achieves an $L/D \simeq 20$. The best glide ratio of the Boeing B-52 is about $L/D \simeq 20$, while Concorde struggled to achieve $L/D \simeq 9$; the Wright Flyer had a respectable $L/D \simeq 8.3$. The airfoil Liebeck L-1003 has an optimal $L/D \simeq 220$ at a Reynolds number $Re = 10^6$, e.g. a factor 10 compared to the Boeing B-52.

The glide ratio is a parameter variable with the speed and the weight of the aircraft, and changes during a long-range cruise. Table 2.3-1 summarizes glide ratio data for some known aircraft in cruise conditions.

The glide ratio is one of the essential performance parameters; therefore, it is of interest to calculate optimal values with respect to the aircraft's mass (weight), speed and flight altitude. By using the drag Eq. 2.3.13, the glide ratio becomes

$$\frac{L}{D} = \frac{C_L}{C_D} = \frac{C_L}{C_{D_o} + kC_L^2}, \tag{2.3.15}$$

and using the definition of C_L,

$$\frac{C_L}{C_D} = \frac{c_1/U^2}{C_{D_o} + c_2/U^4}, \tag{2.3.16}$$

Table 2.3-1 Average aerodynamic data for selected aircraft (estimated).

Aircraft	L/D	M(L/D)	M	Notes
Boeing B-52G	20.5	16.4	0.80	
Lockheed C-5A	18.6	14.5	0.78	
Boeing B-707/200	18.2	14.4	0.79	
Boeing B-747/100	17.6	14.8	0.84	
Douglas DC-9-30	17.2	13.2	0.77	
Douglas DC-3	14.7	4.1	0.28	propeller
Concorde	9.0	18.0	2.05	
McDonnell-Douglas F-15C	10.0	9.0	0.90	transonic
McDonnell-Douglas F-15C	4.0	6.4	1.60	supersonic
XB-70A	7.55	5.74	0.76	subsonic
XB-70A	5.14	6.22	1.21	
XB-70A	8.72	24.35	2.79	supersonic

with $c_1 = (2/\rho)(W/A)$, and $c_2 = kc_1^2$. To find the speed corresponding to maximum L/D, we need to find its derivative with respect to the speed and set it to zero:

$$\frac{\partial}{\partial U}\left(\frac{c_L}{C_D}\right) = \frac{2U(C_{D_o} + c_2U^4) - 4U^2c_2U^3}{(C_{D_o} + c_2U^4)^2} = 0, \tag{2.3.17}$$

$$-2c_1U^{-3}(C_{D_o} + c_2/U^4) - 4(c_1/U^2)c_2U^{-5} = 0. \tag{2.3.18}$$

By further simplification, we find

$$U^4 = \frac{k}{C_{D_o}}\left(\frac{2W}{\rho A}\right)^2. \tag{2.3.19}$$

Another way of expressing Eq. 2.3.15 is in terms of the mass (or weight) at given flight speed. Again, using the definition of C_L, we find:

$$\frac{C_L}{C_D} = \frac{c_1m}{C_{D_o} + c_2m^2}, \tag{2.3.20}$$

where this time

$$c_1 = \frac{2g}{\rho AU^2}, \qquad c_2 = kc_1^2. \tag{2.3.21}$$

As in the previous case, the glide ratio is optimal at only one mass – for a given flight speed. This can be found from differentiating Eq. 2.3.20 with respect to the mass and setting the derivative to zero:

$$\frac{\partial}{\partial m}\left(\frac{C_L}{C_D}\right) = \frac{c_1(C_{D_o} + c_2m^2) - 2c_1c_2m^2}{(C_{D_o} + c_2m^2)^2} = 0, \tag{2.3.22}$$

which leads to

$$m = \sqrt{\frac{C_{D_o}}{c_2}} = \frac{1}{c_1}\sqrt{\frac{C_{D_o}}{k}} = \frac{\rho AU^2}{2g}\sqrt{\frac{C_{D_o}}{k}}. \tag{2.3.23}$$

It is straightforward to read the results of Eq. 2.3.23: the optimal mass (or weight) of the aircraft decreases with the increasing speed and flight altitude – all other parameters being constant. The term under square root is an aerodynamic factor that will appear again in other optimal expressions.

A parametric study of Eq. 2.3.23 is shown in Figure 2.3-13. The maximum value of L/D does not change with the cruise altitude, but the mass corresponding to its maximum value does. In fact, the mass for maximum L/D decreases as the aircraft climbs. The mass required for the global optima may well exceed the design mass of the aircraft.

Another parametric study, which will turn out to be useful for the cruise analysis, is shown in Figure 2.3-14. In the figure we show the estimated L/D of the reference subsonic jet as a function of the cruise altitude, for a fixed mass. Again, the global optimum is achieved at conditions beyond the limits of this aircraft's performance.

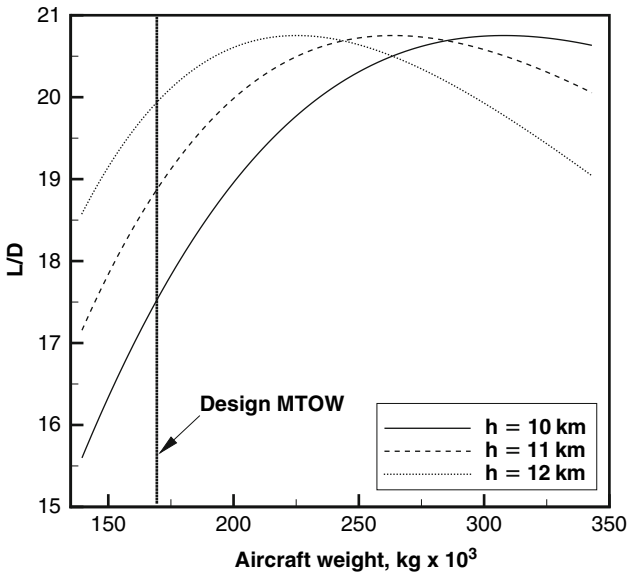


Figure 2.3-13 Glide ratio versus aircraft mass at different flight altitudes (as indicated), at a cruise Mach number $M = 0.80$, aircraft model A.

2.3.7 Glide Ratio at Transonic and Supersonic Speed

The glide ratio at transonic and supersonic speed cannot be expressed in a simple form. In general, the data are tabulated as a function of the Mach number. We use the lift and drag equation for high-speed flight, Eq. 2.3.6 and Eq. 2.3.12, and calculate the glide ratio as

$$\frac{C_L}{C_D} = \frac{C_{L_\alpha}\alpha}{C_{D_0} + \eta C_{L_\alpha}\alpha^2} = \frac{1}{C_{D_0}/C_{L_\alpha}\alpha + \eta\alpha} = f(M, \alpha). \tag{2.3.24}$$

Exact evaluation of Eq. 2.3.24 can only be done numerically. The algorithm is the following.

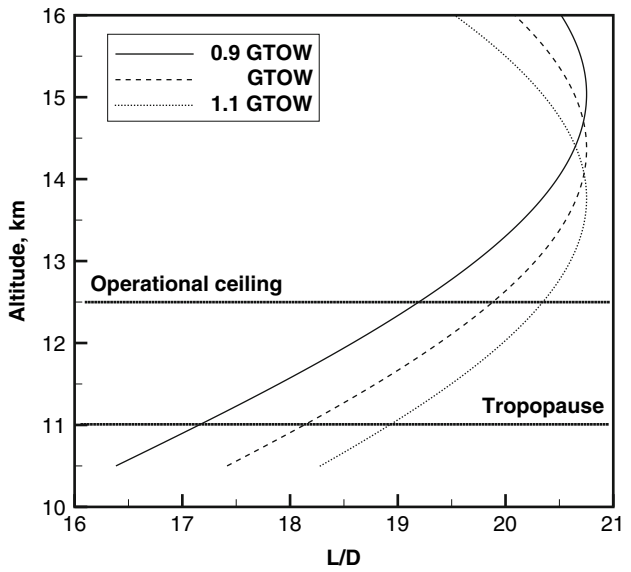


Figure 2.3-14 Glide ratio versus altitude at fixed weight, cruise Mach number $M = 0.80$. The graph shows the effects of gross weight from a nominal value GTOW = 155,000 kg.

Computational Procedure

- Set the Mach number. Hence, Eq. 2.3.24 is only a function of α.
- The condition of min/max is found from the derivative of Eq. 2.3.24 with respect to α:

$$\frac{\partial}{\partial \alpha}\left(\frac{C_L}{C_D}\right) = 0. \tag{2.3.25}$$

After some algebra, the optimal condition becomes

$$\alpha^2 = \frac{C_{D_0}}{\eta C_{L_\alpha}}, \qquad \alpha = \sqrt{\frac{C_{D_0}}{\eta C_{L_\alpha}}}. \tag{2.34.26}$$

- The value of the root α of Eq. 2.3.26 is inserted in Eq. 2.3.24; the corresponding value of C_L/C_D is stored in memory.
- If the value of C_L/C_D has increased, then we set the maximum to the current value.
- Increase the Mach number and repeat the procedure, as above, until the maximum Mach number is reached, or until the C_L/C_D starts decreasing.

The result of this algorithm is presented in Figure 2.3-15. The maximum value of $C_L/C_D \simeq 16$ at low Mach number ($M < 0.2$), and a cruise angle of attack $\alpha \simeq 3.6$ degrees. Note the transonic dip at low supersonic speeds. The L/D is relatively low, and it reaches a value of about 6 at full supersonic conditions (Concorde managed an L/D around 9). There is a small recovery at high supersonic speeds. This behavior of the glide ratio as a function of the Mach number is characteristic of many high-speed vehicles, as demonstrated by Küchemann.

2.3.8 Practical Estimation of the Drag Coefficient

The aerodynamic performance of most aircraft is not widely advertised. Nevertheless, there are methods that allow for a practical estimation of the drag characteristics. The values of the lift are not as interesting as the drag. In fact, the lift is mostly depending on the weight. The drag depends on both the weight and the type of aerodynamic surface. We consider as a reference a subsonic jet aircraft, although the following considerations are valid for other fixed-wing aircraft whose drag equation is parabolic.

The glide ratios are known from statistical data. In the present context we will work with $C_L/C_D \simeq 18$, which is neither the best nor the worst in its class (see Table 2.3-1). If the weight of the aircraft is W, then the lift coefficient is

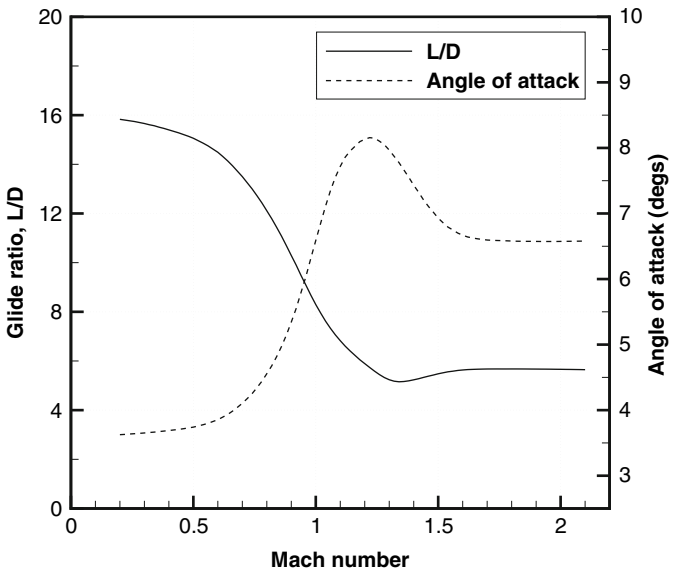


Figure 2.3-15 Glide ratio versus Mach number for the supersonic aircraft in baseline configuration.

found from Eq. 2.3.2. For example, if the AUW of the aircraft is 160,000 kg, the C_L at cruise conditions ($h = 11{,}000$ m, $M \simeq 0.80$) is estimated at $\simeq 0.558$.

The aircraft drag will be $C_D = (L/D)C_L \simeq 0.031$, a relatively high figure, although not far off the true value. The next problem is to find the contribution of the zero-lift drag and the lift-induced drag. One possibility is to use the equation

$$C_{D_0} = q\overline{C}_f A_{wet},$$

which requires the knowledge of the total wetted area, A_{wet}, and the average skin friction coefficient, $\overline{C}_f$. These data are not easily found. There are practical methods to estimate the wetted area of an aircraft the average skin friction is an integral quantity. The other idea is to use again statistical data; for a modern airliner the profile drag will not be too far from 230 drag counts, e.g. $C_{D_0} \simeq 0.0230$. The factor k is then found from

$$C_{D_0} + kC_L^2 \simeq 0.031 \quad \Rightarrow \quad k \simeq 0.0257.$$

These data can be further refined. If the glide ratio were 18.5, then $k \simeq 0.0230$, a 10% reduction over the first estimate. If, instead, we improve the C_{D_0} down to 220 drag counts (a lower limit at the current technology level), then with the initial estimate of L/D we find $k \simeq 0.0289$.

Another method is to use the engine thrust data. For example, an aircraft at cruise conditions will have a measurable fuel flow. From the fuel flow we could extract the effective engine thrust, by using the charts of the specific fuel consumption provided by the engine manufacturer, and hence the aircraft drag. From Eq. 2.3.1 the calculation of C_D is straightforward.

2.3.9 Compressibility Effects

As the free stream Mach number is increased beyond $M = 0.4 - 0.5$, some compressibility effects start to appear on the lifting surfaces. These effects are compounded by local accelerations and angles of attack. Systems that are particularly affected are the propellers and rotors. A full compressible flow theory is not necessary until transonic Mach numbers are achieved. Semi-empirical corrections are applied to the aerodynamic characteristics to take into account some of these effects. A number of corrections for lift, drag, pressure and pitching moment coefficients exist to deal with flows at speed range. Prandtl–Glauert, Kármán–Tsien, Chaplygin, Busemann, and others provided relatively simple expressions for the correction of the two-dimensional airfoil characteristics. These correction formulas can be found in most books on applied aerodynamics. The Prandtl–Glauert formulas represent a first-order correction, and are given by

$$C_D = \frac{C_D}{\sqrt{1 - M^2}}, \qquad C_{L_\alpha} = \frac{C_{L_\alpha}}{\sqrt{1 - M^2}}. \tag{2.3.27}$$

At full supersonic conditions this correction can be replaced by Ackeret's linearized flow theory. In this case the aerodynamic coefficients of the wing, corrected for three-dimensional effects, are

$$C_L = \frac{4}{\sqrt{M^2 - 1}}\left[1 - \frac{1}{2\mathcal{AR}\sqrt{M^2 - 1}}\right], \qquad C_D = C_L\alpha. \tag{2.3.28}$$

These expressions are valid only when there is no interaction between the Mach cones arising from shocks at the wing tips. The transonic regime, being highly non-linear, is not covered by corrective factors. There exist a number of theories to calculate the lift-curve slope of thin wings at supersonic speeds, for example wings with subsonic leading edges and supersonic leading edges.

2.3.10 Transonic Drag Rise

At Mach numbers beyond the validity of the compressibility correction a number of non-linear transonic phenomena start to occur. These phenomena are reflected in Eq. 2.3.12 by the functional dependence of the coefficients from the Mach number. For reference, the beginning of transonic drag rise is set by the *divergence Mach number*, which is conventionally defined by the point where the derivative

$$\frac{\partial C_D}{\partial M} > 0.1. \tag{2.3.29}$$

The Mach number at which this occurs is a complex function of the aircraft configuration and operational conditions (angle of attack, lift coefficient, altitude). For subsonic commercial jets it is of relatively limited implication, because

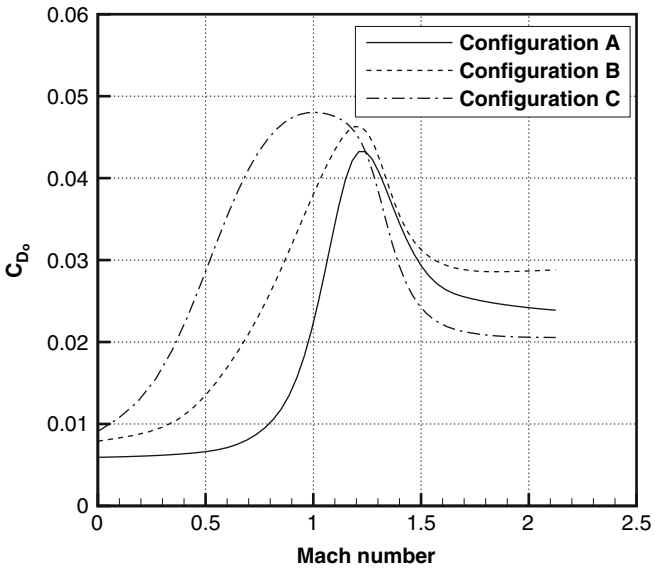


Figure 2.3-16 Zero-lift drag coefficient for three different configurations. Configuration A is the baseline.

both the operating angle of attack and the cruise speed vary over a narrow range. For a high-performance fighter jet the issue is more complex, because it depends strongly on its configuration. An example is reported in Figure 2.3-16, for the reference supersonic aircraft model C. The graph is indicative of some operational configurations. The drag reaches a maximum somewhere between $M = 0.8$ and 1.2, then it decreases and settles to a level higher than the value before the transonic drag rise. Therefore, having an engine powerful enough, the aircraft can go past the *sonic barrier* and fly at supersonic speeds.

The penalty to be paid in supersonic flight is sometimes quantified by the transonic drag jump; this is the drag increase between two reference speeds, which are $M = 1.2$ (supersonic) and $M = 0.8$ (subsonic). The corresponding *transonic drag penalty* is plotted in Figure 2.3-17, for a number of fighter and experimental aircraft. On the horizontal axis there is the fineness ratio l/d'. A clean configuration has a lower drag penalty. Also, more modern aircraft are less sensitive to the transonic flight conditions. For reference, also the drag discontinuity of a Sears–Haack body is plotted.

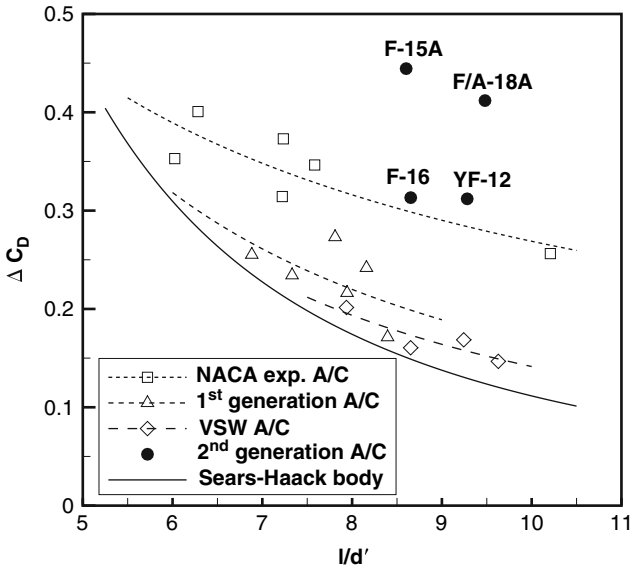


Figure 2.3-17 Transonic drag rise for some aircraft.

The Sears–Haack body is a body of revolution pointed at both ends, having the minimum supersonic drag or given length and volume. It represents an ideal situation, to which an aerodynamic performance should tend. A full discussion of the Sears–Haack body is given in some good textbooks of aerodynamics, for example that of Ashley and Landhal.

2.3.11 Lift and Transonic Buffet

The transonic effects on the lift can be considerable. The loss of lift is due to aeroelastic response under unsteady aerodynamic loads. A loss of lift due to wing buffet is called *transonic dip*, and sometimes *shock stall*. Figure 2.3-18 shows the transonic dip at the inception of wing buffeting. One of the first military jet aircraft, such as the North American F-84, suffered badly from loss of lift at transonic speeds. A 10-year older design, such as the North American F-100 Super Sabre, quickly recovered the lift, although it showed some decrease in the transonic lift.

Another effect of the Mach number is on the maximum lift coefficient of the aircraft, $C_{L_{max}}$. For a subsonic transport aircraft $C_{L_{max}}$ is nearly constant up to the drag divergence point, then it is gradually reduced. The angle of attack corresponding to $C_{L_{max}}$ also decreases. For a supersonic fighter jet aircraft the function $C_{L_{max}}(M)$ is more complicated and shows a transonic dip with a partial supersonic recovery. Its value at supersonic speeds can be considerably lower than that at subsonic speeds. These effects are discussed in detail by Abbott and von Doenhoff for NACA airfoil sections, and in the NACA publications reported in that book.

The behavior of $C_{L_{max}}(M)$ is important in high-performance maneuver. The values of $C_{L_{max}}$ are often approximate, and a satisfactory evaluation of this parameter is difficult.

2.3.12 Aero-Thermodynamic Heating

Heating due to the aerodynamics of high-speed flight is due to the conversion of kinetic energy of the flow into thermal energy. The amount and intensity of this exchange is dependent on the free stream Mach number and on the local geometry. It is highest at the stagnation points of the aircraft. This temperature increase is due to the difficulty of transferring momentum to the outside flow domain, therefore the momentum is dissipated in heat. From supersonic aerodynamic theory at moderate Mach numbers ($M < 3$), the temperature rise at the stagnation point can be found from the energy equation

$$C_p\mathcal{T} + \frac{1}{2}U^2 = C_pT_o, \tag{2.3.30}$$

which is valid for an iso-entropic transformation; T denotes the absolute temperature, C_p is the specific heat at constant pressure, and "o" denote stagnation conditions. The temperature rise at the stagnation points is found from

$$\Delta\mathcal{T} = \frac{U^2}{2C_p} = \frac{1}{2}\frac{a^2M^2}{C_p} = \frac{1}{2}\frac{\gamma\mathcal{R}\mathcal{T}M^2}{C_p}, \tag{2.3.31}$$

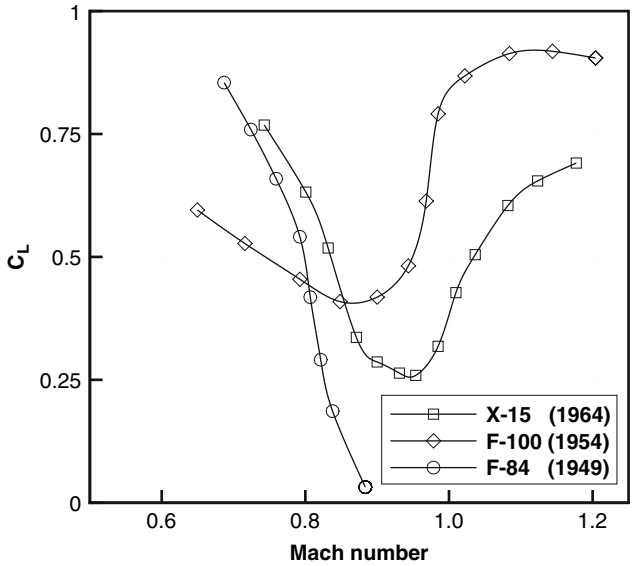


Figure 2.3-18 Transonic dip for three reference jet aircraft.

Table 2.3-2 Aero-thermodynamic heating at stagnation points (estimated data).

Aircraft	M	Max temp., K
Concorde	2.0	380
Convair B-58	2.0	420
XB-70	3.0	550
X-15	6.0	900
Waverider	6–8	1,000
Space shuttle	10–12	1,500
ICBM	20–25	6,000

where in this case γ is the ratio between specific heats and is taken equal to 1.4 (approximation valid at temperatures below 900 K). This heating will have consequences on the flight envelope of the aircraft. A summary of estimated stagnation point temperature for some high-speed flight vehicles is reported in Table 2.3-2.

Figure 2.3-19 shows the map of aero-thermodynamic heating at the stagnation point for supersonic speeds in atmospheric flight, as deduced from Eq. 2.3.30. The problem of hypervelocity vehicles (waveriders, re-entry vehicles and missiles) is complicated by the fact that the flow does not obey the law of ideal gases, and the aero-thermodynamic heating has to be calculated using real gas relationships. A discussion can be found in specialized studies on hypersonics and missiles and planetary entry flight mechanics.

Example

The actual skin temperature will be the value shown minus the difference between altitude and sea level temperature. For example, the estimate for Concorde would be about 173 degrees above the local temperature. Therefore, the sea level temperature would be about $173 - (56 + 15) = 102°C$. The actual temperature found in the technical publications on Concorde report most of the nose above 100°C, the leading edge of the wing at a temperature of about 100°C and most of the fuselage at temperatures of 90 to 97°C. Therefore, the approximation is acceptable.

2.3.13 Aerodynamic Penetration and Radius

We define two parameters (flight mechanics and design quantities) that have the dimensions of a length. Consider a rigid-body aircraft of mass m at a flight speed U without engine thrust, and subject to a drag force D. The equation of motion in the flight direction is

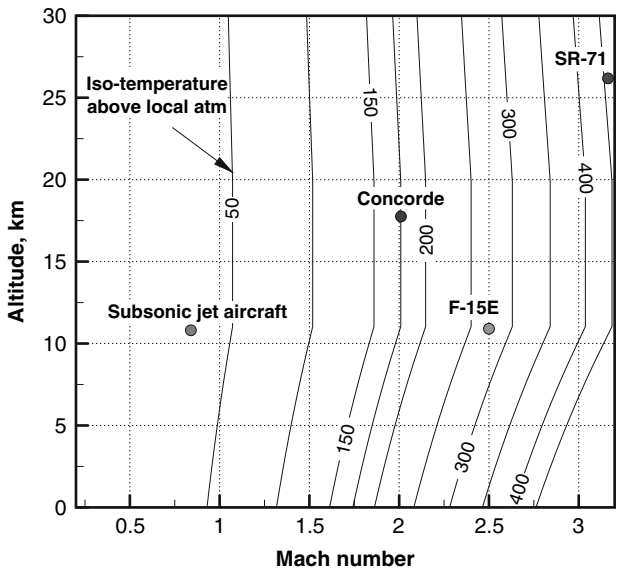


Figure 2.3-19 Stagnation point temperatures above local atmospheric temperature.

$$a = -\frac{D}{m}, \tag{2.3.32}$$

or

$$a = \frac{\partial U}{\partial t} = U\frac{\partial U}{\partial x} = \frac{\rho A C_D U^2}{2m}. \tag{2.3.33}$$

Further simplification of this equation leads to

$$\frac{dU}{U} = \frac{\rho A C_D U^2}{2m} dx. \tag{2.3.34}$$

The factor

$$s_D = \frac{2m}{\rho A C_D U^2} \tag{2.3.35}$$

is called *aerodynamic penetration* and represents the distance that the aircraft of mass m can travel at constant density-altitude if subject to a drag force that is proportional to the speed squared. Therefore, if C_D is constant, the solution of the equation above is

$$\frac{U}{U_{ref}} = e^{x/s_D}, \tag{2.3.36}$$

where U_{ref} is a reference speed. Now assume that the aircraft is accelerated in a direction normal to its flight path by lift alone. The equation of motion along the flight path is

$$U\frac{\partial \gamma}{\partial t} = \frac{U^2}{\chi} = \frac{\rho A C_L}{2m} U^2, \tag{2.3.37}$$

where χ is the radius of curvature of the flight path, and γ is the local attitude. This equation can also be written as

$$\frac{1}{s_L} = \frac{1}{\chi} = \frac{\rho A C_L}{2m}. \tag{2.3.38}$$

The *aerodynamic radius* is the radius of curvature of the flight path due to lift. This parameter will be called $S_L = \chi$.

These concepts were first introduced by Larrabee to describe several problems in atmospheric flight and atmospheric re-entry. They both have the dimension of a length. They can be used for the solution of Lanchester's phugoid. Additional applications of these concepts include the wind drift and the lateral maneuver. Wind drift is important in ballistics, parachute drops and unguided atmospheric re-entry. Lateral maneuver is important in air-to-air missiles and combat aircraft.

2.3.14 Aircraft Vortex Wakes

When the fixed-wing aircraft generates lift from its lifting surfaces, it creates a complex vortex system (tip vortices, flap vortices) that travels downwards (down-wash). The strength and shape of these vortices depend essentially on the main wing parameters. The downward mass flow depends on the weight of the aircraft. In fact, it is possible to prove that the downwash w creates a downward mass flow rate $\dot{m} \propto W/b$. From low-speed aerodynamic theory, under the approximation of large aspect-ratio and elliptic wing loading downwash, the mass flow rate is

$$\dot{m} = \rho A_{ref} \overline{w} = \frac{2}{\pi}\frac{W}{b}. \tag{2.3.39}$$

We have assumed that the reference area A_{ref} behind the aircraft is $A_{ref} = \Delta x b = \Delta t\, U b = U b$ in the unit of time*. This flow rate is not dependent on the flight altitude. Therefore, for a Boeing B-747 at a cruise speed $M = 0.80$ with an AUW = 350 tons, we have $\dot{m} \sim 5{,}400$ kg. At a density altitude $\sigma = 0.29$ this corresponds to about 15,200 m^3 of downward flow every second, at an average downwash speed of $\overline{w} = 6$ m/s. If a small airplane of the same class as the Piper 28 flies just below the B-747 at an average speed of 120 kt (222 km/h), the downwash reduces the inflow by about 5.5 degrees – a potentially catastrophic effect.

* Note that W is a weight and $\dot{m}$ is a mass flow. The discrepancy in units is only due to the choice of the reference area. To avoid confusion, express all the quantities in SI units.

Figure 2.3-20 Vortex trails in the sky: trails intersecting at various degrees of dissipation.

The vortex system creates a hazard for vehicles following at a short distance, or below. This problem is of interest in congested air spaces and around airports with large volumes of traffic. Research in this field has been buoyant, in the attempt to reduce flight hazard and separation times between take-off and landing operations. Some research now focuses on two broad areas: (1) the technical means and design solutions to diffuse the vortex system, and (2) on the optimal flight paths that would reduce the runway occupation time. For example, the shortest method to fly away from a trailing vortex is to follow a curved path (by a banked turn). Rossow suggested that with individual flight corridors, created by small changes of operation, the time interval between take-offs can be reduced to the order of seconds.

The air traffic control authorities prescribe separation distances based on the weight of the leading and trailing aircraft. For example, an airplane with a gross weight $W = 100{,}000$ kg following a larger aircraft with mass of the order of 300,000 kg must be flying at least 5 nautical miles behind.

The increase in weight of some aircraft brings them up in the separation distance. Obviously, the problem becomes more urgent for the very large aircraft. The Airbus A-380, with a gross weight that is 50% higher than the Boeing B-747, is one of the most recent cases.

There are different separation distances: the *vertical separation* is obtained by assigning different cruising altitudes; the *lateral separation* is obtained by maintaining a distance between the flight paths of two airplanes; the *longitudinal separation* is the distance between the leading and trailing aircraft, if these were to follow the same flight path.

Similar considerations will have to be made for the helicopters, although helicopter wakes are considerably more complicated than the wakes from fixed-wing aircraft. The strongest helicopter wakes occur when the helicopter is operating at low speed (less than 80 km/h).

Another problem, mostly environmental, is due to aircraft contrails. These are condensation trails left behind by the jet engines. As the contrails dissipate, they tend to assume the shape of some clouds, Figure 2.3-20. The contrails often turn into cirrus clouds and are a sign of changing weather. The environmental conditions required for contrail formation and persistence are discussed by Jensen and co-workers, and Schumann. Detwiler and Jackson treated the contrail formation under standard atmosphere, with focus on the engine cycle. The IPCC estimated that contrails covered about 0.1% of the Earth surface in 1992, and that this figure is set to rise to 0.5% by 2050.

2.3.15 Aerodynamics and Performance

For details on aerodynamics we invite the readers to consult the vast literature on the subject. We recommend, by increasing order of complexity, the books of Bertin, Kuethe and Chow, Ashley and Landhal, and most of the volumes of the Princeton Series in *High Speed Aerodynamics and Jet Propulsion*[134]. The book of R.T. Jones is a classic on wing theory. Katz and Plotkin published a focused book on panel methods and lifting surface methods. These methods are rapid in generating aerodynamic characteristics and stability derivatives of complex wing systems and of the full airplane.

The theoretical foundations in relation to aircraft shape, aerodynamics and performance have been analyzed by Küchemann and Weber. A more applied approach to the aerodynamics of the airplane has been extensively discussed in the literature.

Some of the aircraft flights discussed in this book include non-steady flight. Aircraft maneuvering is simulated with steady state aerodynamics. This approximation is adequate for moderate accelerations on a straight flight path. There is scant research in this area. At the very low end of the speed range (pertinent to birds' flight), experimental results indicate that sudden accelerations increase the lift to values considerably higher than steady state. For accelerations past the speed of sound, the main difference in aerodynamic response is in the transonic area.

The accelerations involved in some maneuvers, such as free roll, turning, and V/STOL operations, occur at high angles of attack and high lift. In these cases accurate unsteady aerodynamics is required. Rom's book is one of the few comprehensive publications on aircraft aerodynamics at high angles of attack.

Aerodynamic performance data for airfoils or wing sections are abundant in the specialized literature. More advanced wing sections include the double wedge (Lockheed F-117 A), segmented (North American XB-70) and biconvex wings (Lockheed F-22A, Grumman F-104). All the high-speed aircraft have a supercritical wing section. The problem of estimation of the aircraft aerodynamics from flight testing is discussed in detail by Klein.

The current technology level requires wing sections and planforms to be designed ad hoc, using appropriate numerical and optimization methods. These methods include constraints on geometry and aerodynamics; they may include multipoint design, and off-design analysis.

Aircraft manufacturers are not likely to venture to the public and show their best aerodynamic data. Nevertheless, some aerodynamic performance data of engineering interest are sometimes found in specialized publications. For example, data on the Douglas DC-10, Lockheed L-1011, and Lockheed YF-16 are found in AGARD CP-242. Data are available for the Boeing B-737-100, B-747-100, Boeing B-707 (AGARD LS-67), Lockheed C-141B and C-5A in AGARD R-723; data for the Fokker 50 and 100 can be found in AGARD CP-515. Data for the North American F-100, Boeing B-707, Convair B-58, the Bell X-l, and NASA X-15 were published by Seckel. Data for the trainer aircraft Northrop T-38 have been provided by Brandt and co-workers Heffley and Jewell provide data for the Northrop F-4C and the North American XB-70.

Problems

1. A supersonic aircraft is to accelerate from a cruise speed of $M = 0.8$ through the speed of sound. During the acceleration, there can be a change in aerodynamic response from the lifting surfaces, including a loss of lift and a drag rise. Past the speed of sound, at fully developed supersonic flow, another problem appears: the center of pressure moves aft. Investigate how the transonic dip and the drag rise problems are solved. Then investigate the nature of the shift in the center of pressure, and consider solutions to the stability problem that minimize the trim drag (*problem-based learning: additional research is required*).
2. The aerodynamic drag is one the most important performance parameters of an aircraft. Discuss how this force is affected by (1) the speed and the Mach number; (2) lift and gross weight; (3) angle of attack; (4) cruise altitude; and (5) configuration at landing and take-off.
3. Calculate the aero-thermodynamic heating at stagnation points of an aircraft having a Δ-wing with a sweep $\Lambda = 55$ degrees. The aircraft is flying at a supersonic speed $M = 1.8$ at an altitude $h = 9{,}000$ m. Consider as stagnation points the nose of the aircraft and the leading-edge line of the Δ-wing. (The heating is to be considered as an equilibrium temperature above the atmospheric temperature at the flight altitude.)
4. A generic aircraft has a zero-lift angle of attack α_o inclined over the horizon by $+2$ degrees. The aircraft is on a steady level flight with an angle of attack $\alpha = +3$ degrees. Sketch the direction of the lift and drag forces.
   - If the aircraft rotates around the CG by $\alpha = +1$, how do the lift, drag, normal and axial forces change?
   - Now assume that the aircraft starts a climb, and assumes a constant angle of climb $\gamma = +4$ degrees. Sketch the direction of the aerodynamic forces and indicate the relevant flight angles (angle of attack and climb angle).
5. Estimate the maximum temperature at stagnation points due to aerothermo dynamic heating for aircraft C by using Eq. 2.3.31. Assume hat the aircraft is flying level at altitude $h = 12{,}000$ m and find the corresponding values of the standard atmosphere. Discuss the approximation of the results obtained, and how the heating will be affected by changes in atmospheric conditions.
6. Calculate the mass of airplane model A to cruise at an altitude $h = 11{,}000$ m with maximum glide ratio L/D. Also, calculate the change in optimal mass if the aircraft drifts down to an altitude $h = 10{,}000$ m.
7. Prove that the maximum glide ratio L/D for a given jet-driven airplane does not depend on the cruise altitude (refer to the discussion and results of Fig. 2.3-13).
8. Discuss methods for increasing the lift coefficient of a high-performance fighter jet. Make sketches to explain your ideas.

9. Discuss the reasons why the maximum lift coefficient, $C_{L_{max}}$, of an aircraft is difficult to estimate, and which methods can be used for its evaluation. Also, how do the pilots find out that they are operating the aircraft at (nearly) maximum C_L?
10. Calculate the speed U corresponding to a minimum of the ratio $C_D/C_L^{3/2}$. Consider a parabolic drag equation, with $C_{D_0} = 0.023, \quad k = 0.034$.
11. Assume that the subsonic jet transport in cruise conditions has a zero-lift drag equal to 50% of the total drag. The lift-induced drag is estimated at 35% of the total drag. The remaining drag is due to interference, compressibility effects, excrescences and gaps. Calculate the C_{D_0} and the coefficient k in the parabolic drag equation. Consider an aircraft mass equal to 145,000 kg at cruise conditions. The cruise thrust is $T = 45$ kN.
12. You are required to provide a performance analysis of a certain aircraft. You start from wind tunnel aerodynamic data for the drag. You know that the aircraft model was tested in the wind tunnel as a 1:5 scale, and that the Reynolds number used, $Re = \rho U l/\mu$, matched the actual flight conditions (this does not happen very often). The wind tunnel used was of the variable-density type, with average density at the test section equal to 2 bar. The laboratory conditions were standard atmosphere. The total drag force in the wind tunnel had a value D. Calculate the ratio between the drag force of the actual aircraft and the wind tunnel value.

Aerospace Engineering Desk Reference; ISBN 9781856175753

2.4 Engines

Antonio Filippone

> *A new impetus was given to aviation by the relatively enormous power for weight of the atomic engine; it was at last possible to add Redmaynes's ingenious helicopter ascent and descent engine to the vertical propeller that had hitherto been the sole driving force of the aeroplane without over-weighting the machine …*
>
> H.G. Wells, in *The World Set Free*, 1914

In this chapter we will look at the engine/airframe as a single system. When an engine is mounted on the airframe, it becomes an integral part of the aircraft. The performance parameters considered are the thrust, power, specific fuel consumption, fuel flow, and the free parameters needed to evaluate them (flight speed, altitude and throttle position). Additional parameters are involved in propeller propulsion. These are the propeller's geometric configuration and its rotational speed. For variable throttle settings, the engine response will be considered instantaneous. Therefore, there will be no time lag between throttle and power or thrust.

We will not provide details on how a gas turbine or a reciprocating engine functions. Data and performance on aeroengines are published regularly by Jane's Information Systems. Chart performance, if not already available in the open literature, is almost impossible to obtain from the manufacturers – unless you buy an engine from them. Some basic concepts will be reviewed in this chapter, with the scope of providing means of calculating some propeller and helicopter performance.

It would not be exhaustive to talk about aircraft engines without emphasizing the extraordinary progress that has been made since the beginning of powered flight, and the technological advances that continue to be made. The words of H.G. Wells quoted above were written in 1914 and imagined the world of aviation in the 1950s. Today, the atomic engine is not a reality, and not even an option. Nevertheless, the progress in aircraft propulsion has been extraordinary – thanks to the jet engine.

2.4.1 Gas Turbine Engines

The term *gas turbine* is associated with a jet engine consisting of a compressor, a combustion chamber, a turbine and an exhaust nozzle, although the name refers to both jet thrust engines and shaft power engines. The main types of gas turbine engines are the turbojet, the turbofan and the turboprop. The gas turbine is the core of the engine. However, there are other parts whose function is essential (inlet, fuel lines, fuel nozzles, sensors, collectors, thrust reverser).

The turbojet belongs to the first generation of gas turbine engines*. It consists of a single gas flow. The operation of the engine requires a number of aero-thermodynamic stages: (1) compression of the inlet flow via a number of axial compressor stages; (2) the transfer of the compressed air into the combustion chamber, where it is mixed with fuel; (3) combustion in radially spaced combustion chambers; (4) discharge into a multistage turbine, rotating on the same shaft with the compressor; and (5) ejection of all the exhaust gases as a high-speed hot jet through the nozzle (jet speeds at $M = 2$ and above). The scheme of the engine is shown in Figure 2.4-1a. The air is captured by an inlet, whose other function is to provide pre-compression by an aero-thermodynamic mechanism called *ram compression.* This is an adiabatic compression in the engine inlet due to flow deceleration.

The main function of the turbine downstream of the combustion chamber is to operate the compressor. A considerable amount of power is generally required by the compressor. The rest of the thermal and kinetic energy associated with the mass flow is transformed into a high-speed, high-noise jet released from the nozzle.

The very first turbojets had centrifugal compressors, but as the engines became better understood, they were replaced by more efficient axial compressors. As the thrust requirements increased, the compressor's architecture became more complicated, with low- and high-pressure units, each with several rotor stages. The exhaust gas leaving the combustion chamber has a high temperature (about 1,000°C). When the compressor and the turbine are connected to the same shaft, their rotational speed is the same. This coupling is referred to as *spool.*

* Officially, the first flight of a jet-powered aircraft took place on August 24, 1939. It was the Heinkel He-178, powered by an engine designed by Hans von Ohain.

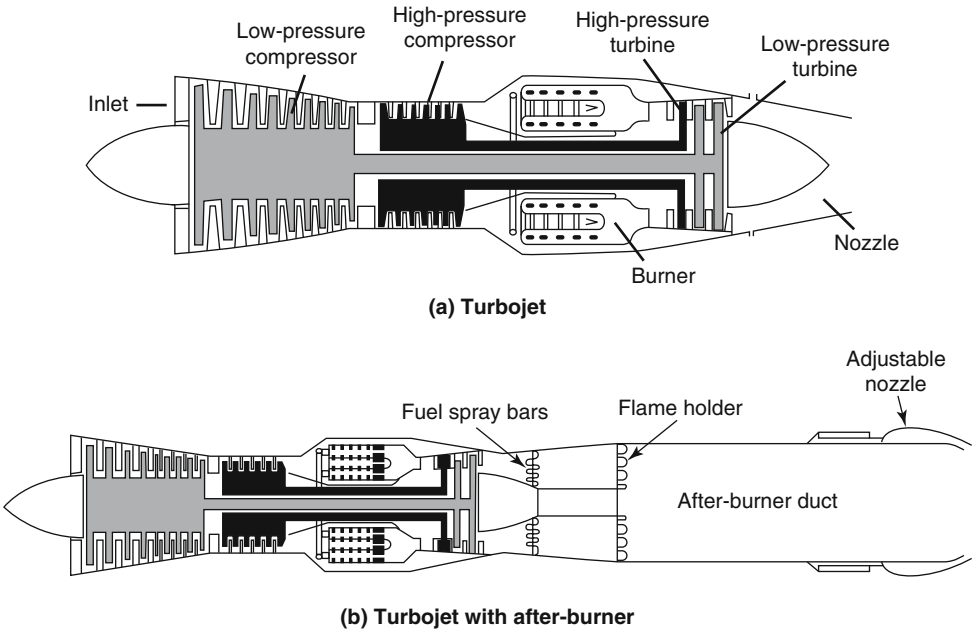


Figure 2.4-1 (a) Turbojet; (b) turbojet with after-burner.

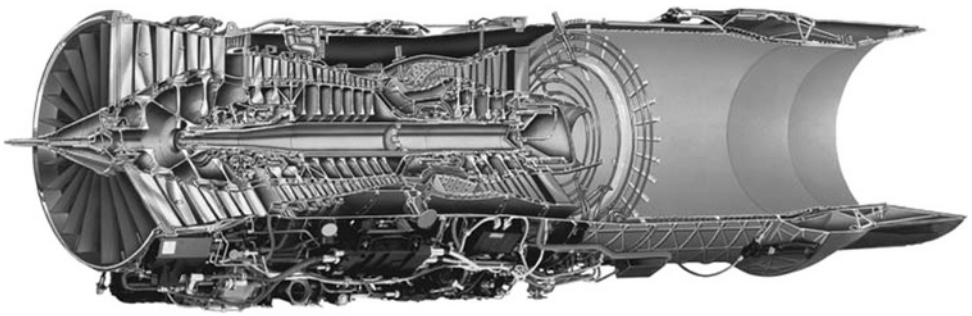


Figure 2.4-2 Low by-pass jet engine P&W F-100, with after-burner (McDonnell-Douglas F-15, Lockheed F-16). By-pass ratio 0.36; pressure ratio 32.0; max thrust about 130 kN (sea level, ISA, with after-burner), max diameter: 1.18 m; gross weight: 1695 kg. (Source: Pratt & Whitney.)

This basic design was applied to engines that powered the early jet airplanes (the Douglas DC-8, the Boeing B-707, the British Comet, and the French Caravelle); it included the JT8D, by most accounts one of the noisiest engines ever to fly. All other engines, leading to the modern high by-pass ratio engines, are derivatives of this basic idea.

The gas turbine can have an additional combustion stage (reheat, or after-burning). Fuel is injected after the primary combustion for the purpose of increasing the engine thrust. After-burning uses the excess air (about 75%) that does not support the primary combustion.

Gas turbines with this capability operate without reheat most of the time, because the increase in thrust is derived at the expense of considerable fuel consumption. Their application is limited to some military jet aircraft, due to the large fuel consumption by the reheat stage. One exception in the commercial arena is the Rolls-Royce Olympus 593, which powered Concorde.

Figure 2.4-2 shows a cutaway of the military engine P&W F-100, a low by-pass engine with after-burning thrust. Note the length of the engine and the length of the reheat section.

A turbofan is another derivative of the turbojet. In this engine the excess air that does not support the combustion is channeled through an external annulus, and by-passes the combustor. The *by-pass ratio* (BPR) is the ratio between the by-pass flow rate and the core flow.

This ratio has been increasing over the years, from about 1.1 to values above 5 in modern engines (although the General Electric GE-90 has a BPR = 8.4, and the P&W GP-7000 series has a BPR = 8.7). This results in engines of considerable size, as shown, for example, in Figures 2.4-3 and 2.4-4. This is one of the engines that powers the Airbus

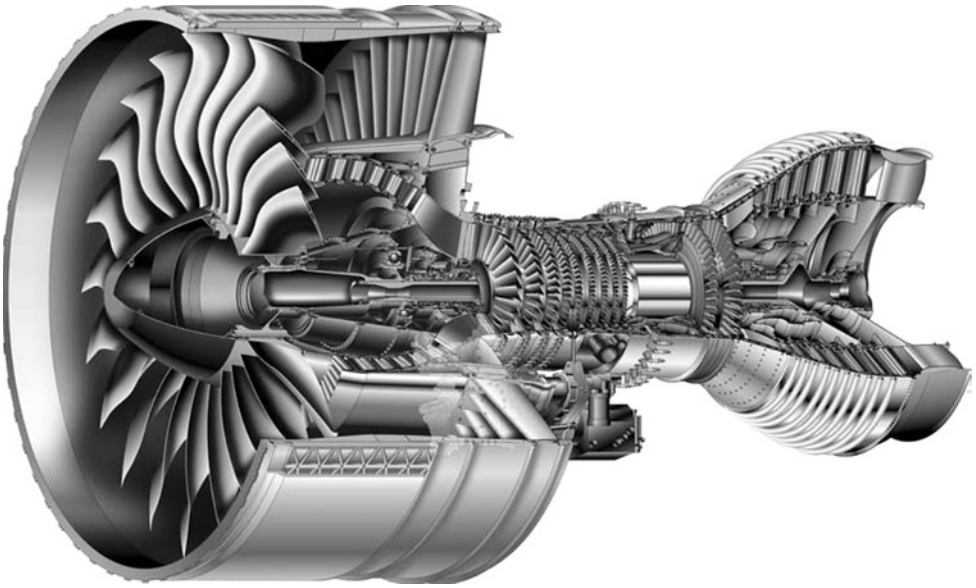


Figure 2.4-3 High by-pass turbofan engine P&W GP-7270, that powers the Airbus A-380. By-pass ratio 8.7 (cruise); pressure ratio 45.6 (climb); take-off thrust 311.4 kN (sea level, ISA), flat rated to 30°C; fan tip diameter: 2.95 m; max diameter: 3.16 m. Noise level: 25.6 EPN dB margin to ICAO Stage 3. Combustion emissions: NOx = 0.0509 kg/kN, CO = 0.0227 kg/kN. Other data: single annular combustor. (Source: Pratt & Whitney.)

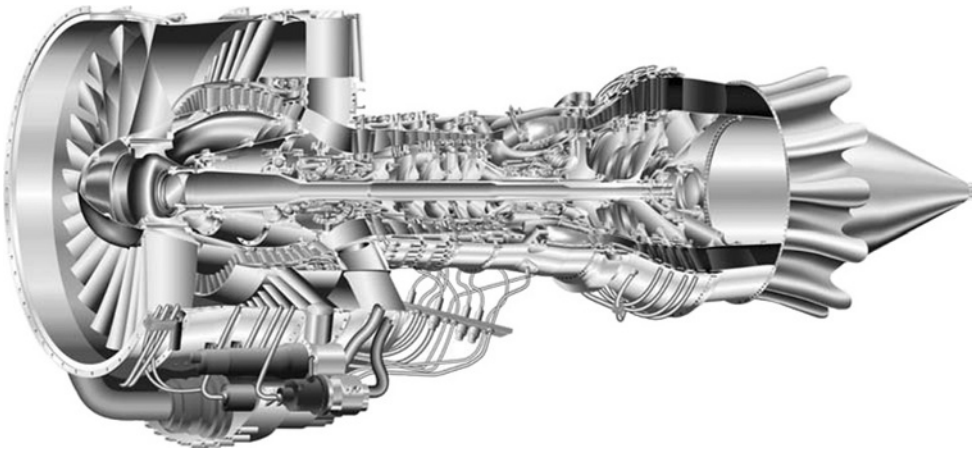


Figure 2.4-4 Medium by-pass turbofan engine P&W PW-6000 (Airbus A-318). By-pass ratio 4.9 (cruise); pressure ratio 28.7 (max); max take-off thrust 10.88 kN (sea level, ISA), flat rated to 32.5°C; fan tip diameter: 1.435 m. (Source: Pratt & Whitney.)

A-380. The other difference, compared to the basic engine, is a large-diameter fan placed in front of a multistage axial compressor. The function of the fan is to increase the capture area of the inlet, and to channel the by-pass flow through the annulus of the engine. The fan is powered by the engine itself, either on the same shaft as the compressor, or on a separate shaft (dual compressor engine). The advantages of this engine are that the exit flow has lower speed, lower average temperature, and produces far less noise.

A *turboprop* is an aero-engine consisting of a gas turbine unit coupled with a propeller, Figure 2.4-5. The thrust can be derived both from the jet engine and the propeller, although in practice most of the useful thrust is imparted by the propeller. Due to the different speed between the gas turbine and the propeller, these engines have a reduction gear. A gas turbine rotates at speeds of the order of 10,000 rpm, while the propeller's speed is less than one-third of this, as limited by the tip Mach number.

Jet engines for helicopter applications are a variant of the turboprop, and consist of one or two gas turbines, a reduction gear of considerable size and a rotor shaft. The reduction in rotational speed is higher than the turboprop, because helicopter rotor speeds rarely exceed 350 rpm.

The subject of aircraft engines is a specialized one and goes under the field of aerospace propulsion or gas turbines.

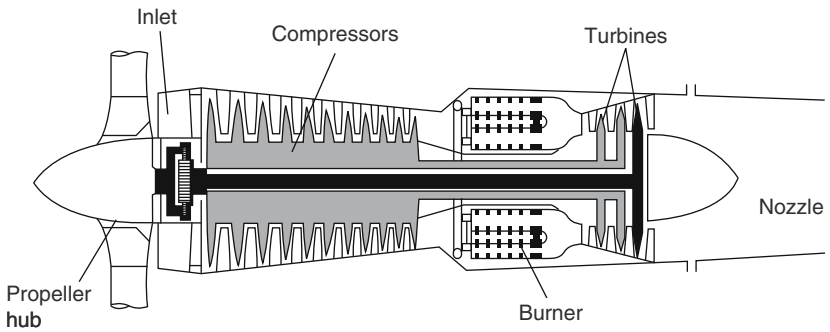


Figure 2.4-5 Basic scheme of turboprop engine. In general, the engine has a gear group between the main shaft and the propeller shaft.

2.4.2 Internal Combustion Engines

These engines, also called *reciprocating* or *piston* engines, dominated aviation in its first half century. Today their application is confined to small aircraft and small helicopters. They account for over 60% of the airplanes flying at the present time. At the low end of the propulsion envelope, they are more flexible, lighter and cheaper than gas turbine engines. These engines are always coupled to a propeller; therefore, their limitation is compounded by the limits on propeller propulsion.

Over time, the reciprocating engines grew to deliver shaft power in excess of 3,000 kW. The first engine to be employed for aircraft propulsion was the Wright Brothers six-cylinder, 12 kW, 3.5 liter engine, capable of running continuously at 850 rpm for up to 15 minutes, after which it would melt because of poor cooling capabilities. The engine weighed 82 kg, and therefore it delivered power at a rate of about 0.15 kW/kg. By comparison, this power loading increased to 0.6 kW/kg by the early 1950s, and engines such as the Pratt & Whitney R-4360 Wasp delivered a maximum power of about 3,200 kW, with a total displacement of 71.4 liters from 28 cylinders (four banks of seven). This engine powered a variety of aircraft, from the Boeing B-29 to the Lockheed Constitution.

Reciprocating engines have the advantage of being consistent in terms of efficiency, which does not change much with air temperature or rpm. The most significant factor affecting the efficiency is the throttle setting. A given amount of power can be produced by an infinite number of manifold pressure and rpm combinations, although the engine will be more efficient at higher manifold pressure and lower rpm combinations.

Flight speeds and altitudes can be increased by using turbo-chargers. Turbo-charging can be done by direct coupling of a turbine to the crankshaft (turbo compound), or by supercharging of the intake air. However, the turbo-charger also tends to increase the engine's temperature, because the inlet air is heated by compression. For this reason, above some critical altitude the engine will overheat, unless some preventive action is taken.

For an aspirated reciprocating engine, the brake horsepower decreases with increasing altitude. A turbo-charger maintains the power with altitude until a critical altitude is reached, after which point the power decreases. One of the main differences between automobile and aviation engines is that aviation engines operate at relatively high power most of the time, and at full power at every take-off. Figure 2.4-6 shows the altitude performance of the Lycoming IO-540, a six-cylinder reciprocating engine for light aircraft. For this engine, a cruise setting of 65% power at 2,400 rpm is recommended.

Generally, curves of this type include lines of constant engine manifold pressure (which are nearly vertical) and correction details for difference in temperature for the inlet air.

Large-diameter propellers are affected by Mach number effects, noise emission, and vibration; hence, large step-down gearshafts have to be used, which add to the mechanical complexity and weight of the engine.

2.4.3 Engine Flight Envelopes

Like the aircraft they power, aero-engines have a flight envelope. This is an enclosed area in the altitude-Mach number space where safe and efficient operation is guaranteed. Another way to characterize the flight envelope of the engine is to use the concept of specific impulse.

First, let us consider the comparative flight envelope of various propulsion systems, as shown schematically in Figure 2.4-7. The lines in the graph show the limits of operation of each engine configuration.

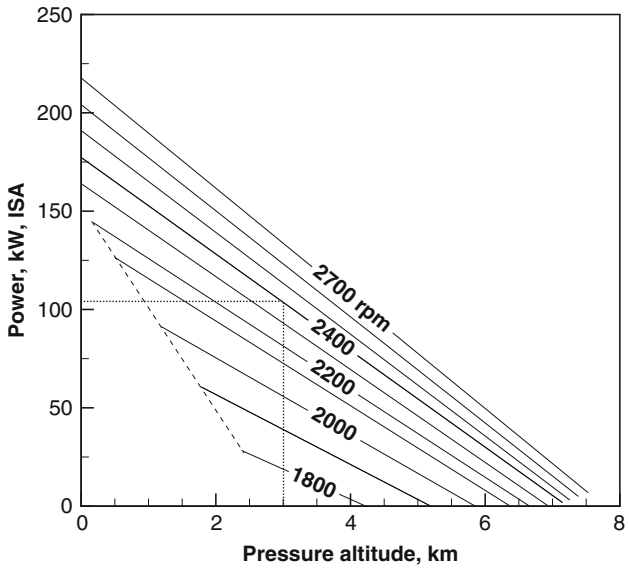


Figure 2.4-6 Lycoming IO-540 performance curves at altitude.

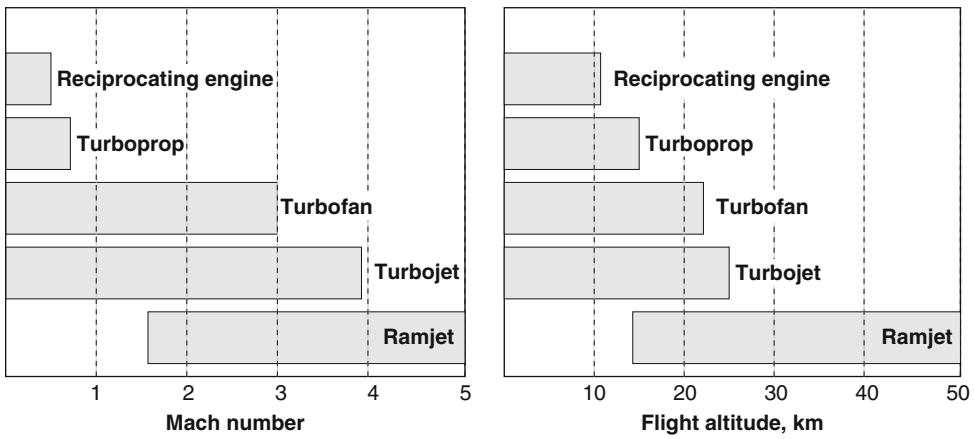


Figure 2.4-7 Schematic operation envelopes of different aircraft engines.

The lower speeds and altitudes are exclusively the domain of the reciprocating engines. Essentially, this type of engine has a low altitude limit, due to the air breathing requirements. Turbo-charged engines provide relatively constant power output up to a critical altitude, $P/P_o = 1$ for $h < h_{crit}$. The power then decreases according to the following semi-empirical relationships

$$\frac{P}{P_o} = \left(\frac{\sigma}{\sigma_{crit}}\right)^{0.765}, \quad h_{crit} < h < 11,000 \text{ m}. \tag{2.4.1}$$

Reciprocating engines provide very limited power above the troposphere.

There is not a single power plant configuration capable of taking the aircraft from take-off to hypersonic speeds. For this reason current research focuses on hybrid engine configurations that include two different power plants. For example, when in the 1960s NASA worked on hypersonic flight with the experimental aircraft X-15, it used the Boeing B-52 as a carrier (under the wing). The X-15's rocket engine was ignited at the cruise speed and altitude of the Boeing B-52.

SpaceShipOne, the first suborbital aircraft to take off from a paved runway (2004) and to land as an airplane, has a hybrid engine system. This consists of a jet engine and a hybrid rocket motor. The carrier aircraft climbs by jet propulsion to about 16,500 m (54,133 ft) with a Mach number $M = 0.85$, then it shifts to a gliding flight and releases the space ship. The ship fires the rocket motor and accelerates to $M = 3.5$ in a steep climb to about 100 km of altitude.

2.4.4 Power and Thrust Definitions

Engine and aircraft manufacturers, as well as the aviation regulations, refer to a variety of power and thrust ratings. The term *uninstalled power/thrust* refers to the engine before the integration with the airframe. The installed power and thrust refer to the air-frame/power plant combination, and are more pertinent to performance calculations. When doing performance calculations, we call thrust the *net thrust,* i.e. the effective thrust delivered by the engine, which accounts for the intake drag and air bleed.

The *shaft power* usually refers to a propeller engine, and denotes the mechanical power delivered by the engine, not accounting for the efficiency of power conversion and the airframe/engine integration. The installation losses include inlet losses, the exhaust gas losses, and the extracting bleed air.

The *equivalent shaft power* is a parameter applied to turboprops, and refers to the sum of the shaft power delivered to the propeller and the power due to the jet of the exhaust gases from the gas turbine, through the nozzle.

The power output of a gas-turbine engine depends on the thermal stresses it is capable of withstanding, e.g., the maximum temperature at any part of the engine. For this reason the manufacturers quote a *maximum continuous power*. This is the engine power level that can be maintained indefinitely, without compromising the thermal and structural integrity of the engine. The *maximum emergency power* is the engine power that can be delivered by the engine for a short time. This includes events such as One-Engine-Inoperative (OEI) and take-off. The maximum continuous power is lower than the emergency power. *Take-off power* and maximum emergency power are sometimes the same. Otherwise the take-off power rating refers to a time of 5 minutes, or 10 minutes in case of OEI. The *intermediate power* rating is a power that can be maintained over a brief to medium time (less than 30 minutes), and is used for emergency situations.

Since most aircraft require additional power for auxiliary units (cabin pressurization, air conditioning), some air can be extracted from the compressor (*bleed air*) to provide the auxiliary services. This, in turn, reduces the engine's useful power, because the mass flow into the engine and its combustor is reduced. Some engine charts show the effects of bleed air on the power output. These effects appear as discontinuities in the power curve.

For helicopter applications the manufacturers refer to a *maximum transmission power*. This is a power limited by the maximum torque on the transmission system (shaft, gear and hub). Increasing the engine power normally requires re-engineering of the transmission system.

The engine thrust, like the power, is dependent on aircraft speed, flight altitude and rpm. The *static thrust* is the thrust at zero aircraft speed. For a military aircraft flying without the optional after-burning, the maximum thrust is sometimes called *military thrust*.

Sometimes engines must be run at part-throttle, to avoid exceeding a maximum rated thrust. In such cases, they are called *flat-rated* engines. For example, the CFM56-3 engine is flat-rated at ISA +15° (or +30°). This means that the engine is guaranteed to give the rated thrust at full throttle when the ambient temperature is below ISA +15°. Above this temperature, the engine will give less thrust because the air has lower density.

For aircraft powered by jet engines, a performance parameter is the *specific thrust T/W* (also called *thrust-ratio);* for the propeller-driven aircraft the corresponding parameter is the *power loading, P/W* (or its inverse). These parameters give an indication of the amount of power or thrust available for a given aircraft gross weight. The values of these parameters vary with the flight altitude and Mach number; one therefore has to be clear as to the conditions at which they are taken: sea level or specified altitude; static ($U = 0$) or in flight conditions. The data quoted by the manufacturers (when reported at all) generally refer to static thrust at sea level. Useful data, including historical trends, have been published by McCormick.

Another performance parameter, more related to the engine itself than the aircraft, is the *engine's specific weight,* i.e., the weight of the engine per unit of power (or thrust). The *engine's specific power* is the power delivered per unit of weight of the engine. However, as engines become bigger, also the installation and integration requirements on the airframe change. This means that the nacelles have to be strengthened to sustain the increased engine weight, and the engine's auxiliary systems have to scale up as well. The weight of engine mounts, pylons, thrust reversers and nozzles are roughly proportional to the engine's thrust. On the other hand, the increase in weight due to accessories (fuel controls, sensors, etc.) grows less rapidly.

The use of the concept of power loading and thrust-ratio for propeller and jet aircraft makes the comparison between the two power plants not completely obvious. The specific thrust is a non-dimensional number (thrust force divided by weight); the power loading is a dimensional number, given in kW/kg or kW/N. A jet engine fixed on the ground does not do any useful work, although it burns fuel and delivers an amount of thrust. However, a useful power can be deduced from a flight condition at speed U, $P = TU$. This expression can be compared to the shaft engine power of the turboprop engine. By this analysis we find that the installed engine power of the Airbus A-380 is a factor 8,500 of the Wright Brothers Flyer.

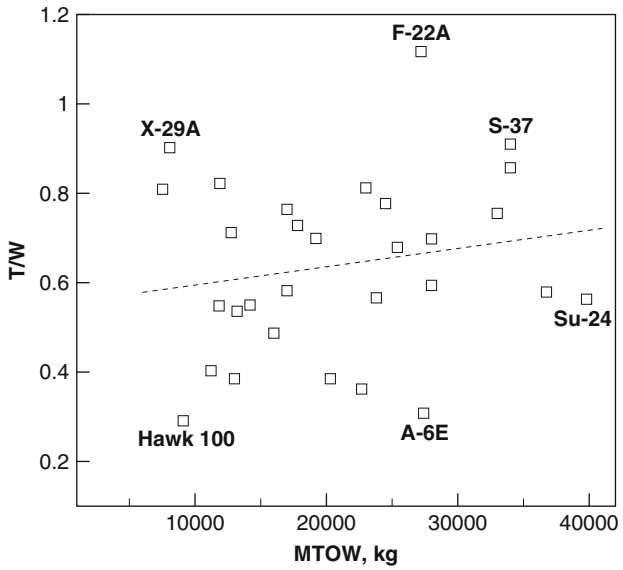


Figure 2.4-8 Fighter jet aircraft thrust-ratios versus MTOW. Dotted line is a least-squares fit of the available data.

Extrapolating from a large number of data, we found that the average thrust-ratio of subsonic commercial jets is about $T/W \sim 0.2$ to 0.4; for a fighter jet aircraft these data double, and become even larger if after-burning is used. From the data in Figure 2.4-8 there is no clear relationship between the thrust-ratio and the MTOW. The complication of the chart also arises from a combination of configurations (external loads, combat mission, use of after-burning).

Example

Let us now take a brief look at the power/weight ratio of some engines. Consider the Wright Flyer's reciprocating engine. Although it was not the best engine of its time, it represented about 20% of the aircraft weight. By comparison, each of the engines on the Airbus A-3 80 delivers a maximum sea level static thrust of 311.4 kN, or 70,000 pounds*. The weight of the GP-7000 engine is estimated to be 4,500 kg, which leads to a total engine weight at MTOW equal to less than 3.5%.

2.4.5 Generalized Engine Performance

The jet engine operation depends on the rotational speed (rpm), the air speed U and the flight altitude h. At ISA conditions the altitude defines the local values for pressure, density and temperature of the inlet flow. The mass flow rate $\dot{m}_f$ and the thrust depend on the inlet diameter d. Therefore, the propulsion parameters are the thrust T, rotational speed, the air speed U (or Mach number M), the temperature T, the pressure p, the fuel flow $\dot{m}_f$, and the diameter d. Since this is a relatively large number of parameters, the engine performance is described by a reduced number of non-dimensional or engineering quantities.

The essential parameters in engine performance are the forward speed, the rotational speed, the thrust T, mass flow rate $\dot{m}_f$, and specific fuel consumption f_j. Therefore, a normalization of these parameters is performed by using the remaining three quantities $(\mathcal{T}, p, d)$.

First, the speed U can be replaced by the Mach number (dimensionless) by means of $M = U/a$. The speed of sound is derived from the flight altitude. Second, the rotational speed of the engine is normalized with a/d, to yield the parameter rpm d/a. Third, the thrust T is replaced by the parameter T/pd^2 (pressure $\times$ length^2 = force). Finally, the mass flow rate $\dot{m}_f$ is replaced by the dimensionless parameter $\dot{m}_f/(pd^2/a) = \dot{m}_f\, a/pd^2$. A further step is required for some parameters. By using the relative pressure δ from Eq. 2.4.33

$$\frac{T}{pd^2} = \frac{T}{\delta}\frac{1}{p_o d^2} \simeq \frac{T}{\delta}. \tag{2.4.2}$$

where p_o is the standard atmospheric pressure at sea level. With the definition of speed of sound, $a = \sqrt{\gamma \mathcal{R} \mathcal{T}}$, we have

*This performance is set to increase to above 80,000 pounds in later versions.

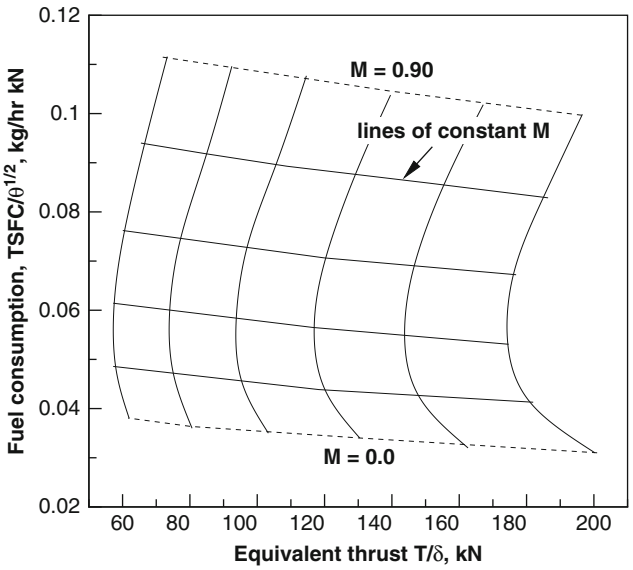


Figure 2.4-9 Corrected thrust versus corrected TSFC for the turbofan engine CF6-50 (estimated). Units for corrected TSFC are kgf of fuel flow per hour per kN of engine thrust.

$$\frac{\text{rpm}\, d}{a} = \frac{\text{rpm}\, d}{\sqrt{\gamma \mathcal{R} T}} = \frac{\text{rpm}}{\sqrt{T}} \frac{d}{\sqrt{\gamma R}}, \tag{2.4.3}$$

$$\frac{\dot{m}_f a}{p d^2} = \frac{\dot{m}_f \sqrt{\gamma \mathcal{R} T}}{p d^2} = \frac{\dot{m}_f \sqrt{T}}{\delta} \frac{\sqrt{\gamma R}}{p_o d^2}. \tag{2.4.4}$$

Instead of using dimensionless parameters, it is a practice in engine performance to use the corrected thrust, rotational speed, and mass flow rate

$$\frac{\mathrm{T}}{\delta}, \quad \frac{\text{rpm}}{\sqrt{T}}, \quad \frac{\dot{m}_f \sqrt{T}}{\delta}, \tag{2.4.5}$$

along with the Mach number M. The remaining factors in Eqs 2.4.2 – 2.4.4 are all constant. The parameters in Eq. 2.4.5 have dimensions. The corrected thrust, the rotational speed and the mass flow rate maintain their physical dimensions. Thrust and mass flow rate are the parameters most widely used in aircraft performance simulation. The rotational speed is generally not given, and calculations are performed at fixed throttle settings, but variable Mach numbers. However, the general dependence of the corrected parameters is

$$\frac{T}{\delta} = f_1\left(M, \frac{\text{rpm}}{\sqrt{T}}\right), \quad \frac{\dot{m}_f \sqrt{T}}{\delta} = f_2\left(M, \frac{\text{rpm}}{\sqrt{T}}\right). \tag{2.4.6}$$

Figure 2.4-9 shows the CF6 corrected engine performance at constant (full) throttle, at Mach number up to 0.9, and altitudes up to 14,000 meters (about 45,931 ft). Figure 2.4-10 shows the estimated thrust specific fuel consumption (TSFC) of the F-100 engine (Figure 2.4-2) versus Mach number at selected altitudes. The TSFC was normalized with the static value at sea level. The results of the analysis show that the TSFC is slightly increasing at cruise altitudes.

2.4.6 Fuel Flow

The fuel flow is the rate at which fuel is burned by the engines. It is given in units of mass (or weight or volume) per unit time. For a jet engine the fuel flow is proportional to the thrust. The constant of proportionality is the TSFC

$$\dot{m}_f = \frac{dm_f}{dt} = TSFC \cdot T. \tag{2.4.7}$$

Here a word of caution is needed, because generally engine manufacturers, textbooks, flight manuals and technical publications quote the most extravagant units for the TSFC, including kg/kg/h, kg/kN/h, kg/N/h, gr/N/s, mg/N/s, lb/lb/h (where kg and gr are units of force!). Sometimes the units are missing, which leaves the figures open to interpretation. In fact, it is not unusual to find wrong data. If one uses international units, the TSFC will be expressed in N/N/s, e.g. newton of fuel weight per newton of engine thrust per unit of time (seconds). While this

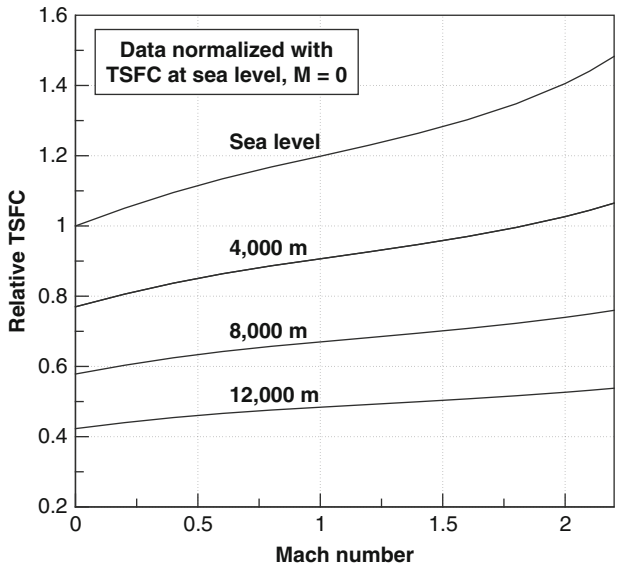


Figure 2.4-10 Relative TSFC versus Mach number for the F-100 engine model (military thrust).

seems rigorous, the unit N/N/s is hardly a useful figure. Similar confusion exists on the fuel flow data (lb/h, 1/h and gallons/h, where the gallon is sometimes the US gallon and other times the Imperial gallon).

Part of the confusion stems from the fact that from a practical point of view one wants to know how much fuel is burned in one hour; and part of the confusion is due to outdated engineering units (pounds, shaft horsepower); and finally there is the confusion between mass, weight and force. One justification for the use of the units lb/h/lb (or N/h/N) is that the SFC is of the order of unity for all jet engines.

It is always a good idea to check the data provided with the gross thrust of the engine. Divide the fuel flow by the engine thrust and make sure that all the units are coherent. Some engine fuel consumption data are provided in Table 2.4-1. Further fuel consumption data and engine efficiency, with historical trends, are given by Babikian and co-workers.

The fuel flow is dependent on the aircraft speed. In fact, from Eq. 2.4.7, at level flight conditions we have

$$D = T = \frac{\dot{m}_f}{TSFC}. \tag{2.4.8}$$

If the TSFC is constant, then the speed of minimum drag is also the speed of minimum thrust and speed of minimum fuel flow. The specific fuel consumption is inversely proportional to the thermal efficiency of the engine, and for propeller propulsion it is also inversely proportional to the propulsion efficiency, η.

The specific fuel consumption of gas turbine engines has been decreasing since their first inception. The chart shown in Figure 2.4-11 was elaborated from data obtained from Rolls-Royce. The datum at the far right of the chart is the 2020 target. The 7–8% improvement on the TSFC is a considerable gain, when one considers the amount of fuel that is burned by an aircraft during one year of operation or during its lifetime in service.

Technical information on aviation fuels has been provided by Goodger and Vere, in addition to some reports on standards. Table 2.4-2 shows some basic data of aviation fuel useful for performance calculations. The stoichiometric ratios are between 0.0672 and 0.0678.

Table 2.4-1 Basic jet engine data.

Engine	SFC, S/L kgm/s/N	BPR	$\dot{m}_a$ kgm/s	T_{to} kN
GE-90	$1.051 \cdot 10^{-5}$	8.4	1,350	70.0
CF6-50C2	$1.051 \cdot 10^{-5}$	5.7	590	50.3
CF6-80C2	$9.320 \cdot 10^{-5}$	5.1	800	50.4
RR RB-211-524G/H	$1.595 \cdot 10^{-5}$	4.3	730	52.1
RR Trent 882	$1.566 \cdot 10^{-5}$	4.3	730	72.2
P&W JT-9D-7R4	$1.566 \cdot 10^{-5}$	5.0	690	176.3
P&W JT-8D-17R	$2.337 \cdot 10^{-5}$	6.6	470	30.8
CFM-56-5C2	$1.606 \cdot 10^{-5}$	1.0	150	18.9
V-2500	$1.629 \cdot 10^{-5}$	5.4	355	21.6

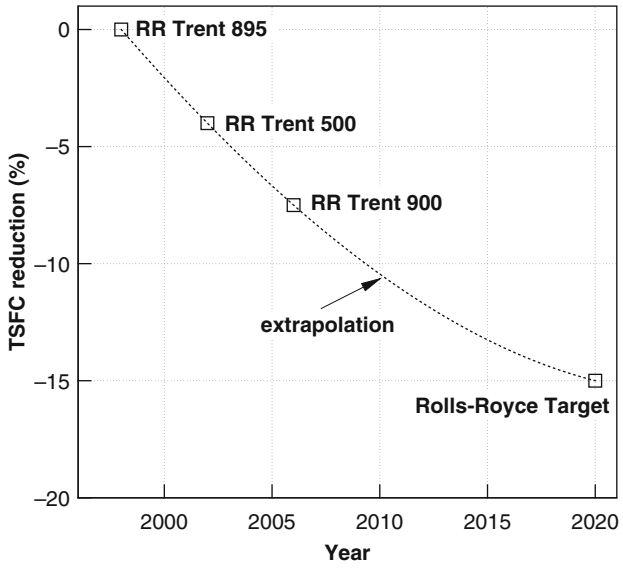


Figure 2.4-11 Specific fuel consumption reduction from technological advances on the Trent series of aero engines (data from Rolls-Royce).

Table 2.4-2 Characteristics of aviation fuels at 15°C. Data are averages.

Fuel	Wide-cut	Kerosene	AV gas
Specific weight	0.762 kg/l	0.810 kg/l	0.715 kg/l
Specific combustion heat	43.54 MJ/kg	43.28 MJ/kg	43.71 MJ/kg

The thermal efficiency is a measure of the effectiveness of the conversion of the thermal energy of the fuel by combustion to useful power. This efficiency is less than unity and reflects the inability to use the heat of the exhaust gases, and presence of other losses associated to the engine cycle, such as the losses in compression, combustion, and expansion.

The *specific impulse* is defined as the ratio between the thrust and the specific fuel consumption,

$$I_{sp} = \frac{T}{SFC}. \tag{2.4.9}$$

It represents the time that a given engine could operate by burning the amount of fuel with a weight equal to the thrust. Sometimes the specific impulse is defined as the amount of time a unit of mass of propellant will last for a unit thrust delivered. This index is useful when comparing different types of propulsion systems (propellers, jets and rockets).

The physical dimensions of the specific impulse are that of speed (force/mass flow rate $= \mathrm{N/(kg\,s^{-1})} = \mathrm{(kg\,ms^{-2})/(kg\,s^{-1})} = \mathrm{m/s}$), although it is sometimes given in seconds. Power plants that deliver large amounts of thrust have a large specific impulse. A map of specific impulse for various types of thrust-producing vehicles is shown in Figure 2.4-12. The specific impulse varies from about 70,000 seconds for a conventional helicopter to 3,500 seconds (one hour) for a pure jet engine with after-burning.

Table 2.4-3 summarizes some fuel flow data for well-known aircraft, taken at their average take-off gross weight. The engines are classified according to the Federal Aviation Administration (FAA), in terms of number of engines and type of airframe.

Example

Below is an example of a calculation of fuel flow and specific fuel consumption at cruise conditions for the airliner. The flight altitude was set at $h = 10{,}500$ m. We have calculated the net thrust from the drag force, and the drag force from an estimated C_D at cruise conditions. The TSFC is then found from solving Eq. 2.4.8.

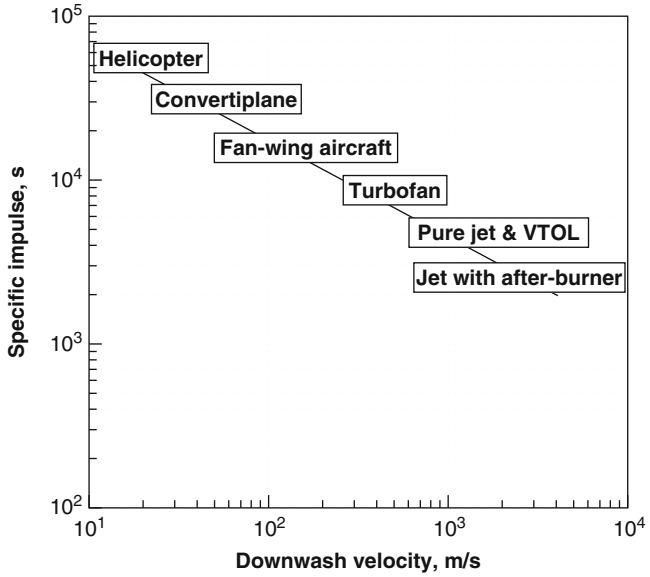


Figure 2.4-12 Specific impulse I_{sp} of various propulsive systems.

Table 2.4-3 Selected aircraft data and applications statistical value.

Aircraft	Type	Engines	*TOW*, kg	$\dot{m}_f$, kg/h
Boeing B-747/100	wide body	4	340,200	13,770
Boeing B-747/400	wide body	4	394,630	15,208
DC-8-63	narrow body	4	158,900	8,641
MD-11	wide body	3	277,930	9,319
Boeing B-727-200	narrow body	3	95,030	6,980
Airbus A-300/600	wide body	2	161,030	6,351
Boeing B-777	wide body	2	248,570	8,012
Boeing B-737/500	narrow body	2	60,240	2,827
MD-80	narrow body	2	67,810	3,531
Canadair CL-601	regional jet	2	19,550	1,317
Fokker F-100	regional jet	2	44,450	2,748
SAAB 340	regional prop	2	12,700	488
Embraer 120	regional prop	2	11,500	608

```
Aircraft mass at start
                                   mi  =  150000.0 kg
                                   fuel mass  =  45000.0 kg
                                   fuel ratio  =  0.30
Estimated cruise drag
                                                CD  =  0.02988
                                              Drag  =  85.156 kN
Estimated TSFC                              TSFC  =  0.161E-03 N/s/N
                                                     0.164E-04 kg_f/s/N
                                                         0.0589 kg_f/h/N
                                                         0.5781 lbf/h/lbf
Estimated Fuel Flow                           ff  =  13.674 N/s
                                                     5018.002 kg_f/h
                                                    11062.617 lb_f/h
Estimated endurance, full tanks     E                      9.76 hours
Estimated endurance w/ actual fuel                         8.97 hours
Glide ratio at start,                               L/D  =  17.280
Figureof Merit at start M                          (L/D)  =  13.824
```

The program includes a few lines to perform basic unit conversions. The estimate for the endurance (flight time with the available fuel) and glide ratio are also useful to validate our assumptions.

2.4.6.1 Aspects of Fuel Consumption

In 1998, worldwide consumption of aviation fuel was about 680 million liters per day, an increase of 13% from 1990. This consumption is now expected to be in excess of 700 million liters/day. Every year air travel releases 600 million tons of CO_2 into the atmosphere.

It was previously mentioned that fuel may be jettisoned in emergency situations. The fuel evaporates quickly and does not reach the ground if the altitude of release is 1,500 feet (457 m) or above. Although this practice is discouraged, some data indicate that civil aviation jettisons about 0.01% of the fuel consumed. At the current levels of consumption, this corresponds to 17,000 liters/day (5,000 tons/year) – hardly a negligible amount. Clevell also reported that fuel is jettisoned about 1,000 times a year by the US military, for a total of 7,000 tons. Regulations for civil aviation require that fuel is jettisoned at sea, from altitudes above 10,000 ft.

2.4.7 Propulsive Efficiency

For an engine flying level at speed U, the one-dimensional equation for the mass flow through the engine is found from the energy equation (or first law of thermodynamics) applied to a control volume of the flow into the engine. We consider the control volume limited by the stream tube through the inlet section, as shown in Figure 2.4-13; we also assume that the flow is in the axial direction.

If $\dot{Q}$ is the heat transferred per unit of time to the air by combustion of the fuel (thermal energy transfer rate), and $\dot{W}$ is the work done on the air, then the energy balance reads:

$$\dot{Q} - \dot{W} = \dot{m}_a\left[(1+\xi)\left(h_e + \frac{1}{2}(U_e - U)^2\right) - h_a - \xi\left(h_f + \frac{U^2}{2}\right)\right], \tag{2.4.10}$$

where the subscript e indicates the exit conditions, a indicates the air and f is relative to the fuel; $\xi = \dot{m}_f/\dot{m}_a$ is the fuel air mixture; h is the enthalpy per unit mass of the flow; and U is the air speed. The work done on the system is

$$\dot{W} = -[TU - (p_e - p_a)A_e], \tag{2.4.11}$$

where T is the thrust, p_e is the pressure at the exit of the control volume, p_a is the air pressure at the inlet of the control volume and A_e is the cross-sectional area of the exit. The first term in Eq. 2.4.11 denotes the work done by the thrust; the second term is the work done against the static pressure. By combining Eq. 2.4.10 and Eq. 2.4.11, after simplification the thrust equation becomes

$$T = \dot{m}_a[(1+\xi)U_e - U] + (p_e - p_a)A_e. \tag{2.4.12}$$

Equation 2.4.12 shows that the engine thrust is proportional to the mass flow rate into the engine.

The propulsive efficiency of the jet engine is defined as the ratio of the thrust-power (power generated by the thrust force T at a speed U) and the rate of production of propellant kinetic energy. This ratio can be written as

$$\eta_p \simeq \frac{TU}{\dot{m}_a[(1+\xi)U_e^2/2 - U^2/2]}. \tag{2.4.13}$$

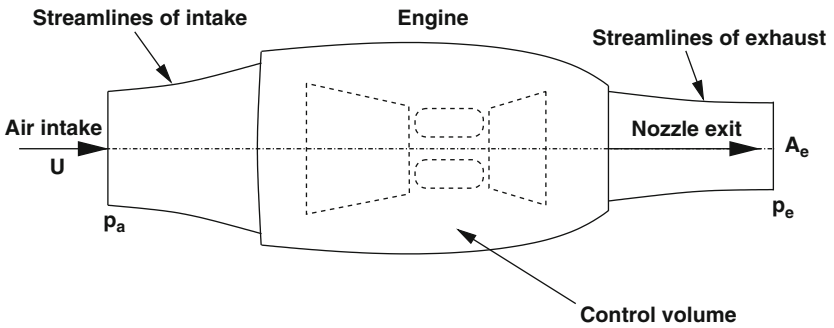


Figure 2.4-13 Control volume on the jet engine.

The fuel-to-air mixture ratio is much less than 1 (in fact, $\xi \simeq 0.067$). Often it is appropriate to neglect the pressure term, because $p_e \simeq p_a$, e.g. the exit pressure is nearly equal to the local atmospheric pressure (fully expanded jets). With these two approximations, from Eq. 2.4.12 and Eq. 2.4.13, the propulsive efficiency becomes

$$\eta_p = \frac{TU}{TU + \dot{m}_a(U_e^2 - U^2)/2}. \tag{2.4.14}$$

Equation 2.4.14 shows that the propulsive efficiency is less than 1 and increases as the jet velocity tends to the engine velocity. The difference $(U_e^2 - U^2)/2$ is the residual kinetic energy per unit mass of the jet. It is worth noting that low jet speeds are not only necessary for high propulsive efficiency, but also for engine noise reduction. One way to reduce the residual kinetic energy of the jet is to increase the mass flow, and this is in fact achieved by the high by-pass ratio turbofans.

2.4.8 Thrust Characteristics

In order to find an expression for the thrust as a function of the throttle setting, it is important to know how the mass flow rate into the engine is dependent on the pressure ratio and the rotational speed of the engine. The mass flow rate is

$$\dot{m}_a = \rho UA = (\rho UA)_c, \tag{2.4.15}$$

where the subscript "c" denotes the flow conditions at the exit of the compressor. The cross-sectional area A_c is fixed. If we use the definition of compression ratio and apply the equation of the ideal gas to the compressor flow, we have

$$\begin{aligned} \dot{m}_a &= \rho_c U_c A_c = \frac{\rho_c}{\rho_\infty} U_c A_c \rho_\infty = \left(\frac{p_c}{p_\infty}\right)\left(\frac{\mathcal{T}_\infty}{\mathcal{T}_c}\right) U_c A_c \rho_\infty \\ &= p_r \frac{U_c}{\mathcal{T}_c} A_c \rho_\infty \mathcal{T}_\infty, \end{aligned} \tag{2.4.16}$$

where p_r is the compression ratio. This can also be written as

$$\frac{p_c - p_\infty}{p_\infty} = p_r - 1 = \rho_\infty \Omega^2. \tag{2.4.17}$$

Therefore, we conclude that the engine thrust is proportional to the mass flow rate, to the pressure ratio, and the square of the rotational speed. A useful approximation is

$$\frac{T}{T_o} = \left(\frac{\Omega}{\Omega_o}\right)^2, \tag{2.4.18}$$

with T_o denoting the maximum power and Ω_o the maximum rotational speed. The axial compressor is designed to perform at its best over a narrow range of speeds, usually close to the maximum *rpm*.

As far as the speed is concerned, the mass flow into the engine is obtained through a convergent nozzle. From compressible aerodynamic theory we know that the flow can be accelerated through a convergent nozzle up to the sonic speed, where the mass flow rate is maximum (for a given cross-section and atmospheric conditions). An increase in the speed of the engine will not increase the speed of the inlet mass flow, and the engine becomes chocked. The chocked mass flow rate can be calculated starting from the St Venant equation

$$\frac{\dot{m}_a}{A} = \rho_o \sqrt{\frac{2\gamma}{\gamma - 1} p_o \rho_o \left(\frac{p}{p_o}\right)^{2/\gamma} \left[1 - \left(\frac{p}{p_o}\right)^{\gamma - 1/\gamma}\right]}, \tag{2.4.19}$$

where the subscript "o" denotes stagnation conditions. Sonic conditions at the throat of the nozzle lead to the chocked mass flow rate

$$\frac{\dot{m}}{A} = 1.8865 \sqrt{p_o \rho_o^3}. \tag{2.4.20}$$

In fact, to have chocked conditions, the flow must reach a sonic speed at the throat of a convergent-divergent nozzle. In this case, $M = 1$ and $p/p_o = 1.893$, as we find from aerodynamic theory.

In first-order performance calculations the jet engine thrust is assumed to be independent of the speed, and variable with the altitude according to the equation

$$\frac{T}{T_o} = \sigma^r \Pi, \tag{2.4.21}$$

with the power $r = 0.7$–0.8, as found from the statistical analysis of several engines. The movement of the throttle commands the opening and closing of the fuel valve. The valve controls the fuel flow into the engines, and hence the thrust and power delivered. Since excessive fuel flow can create temperature surges in some sections of the engine (combustion chamber, turbine, and nozzle), modern engines have means for overriding excess fuel flows at full-throttle condition.

It is important to mention that most performance calculations are done by assuming either a constant throttle, or an impulsive response of the engine to changes in the throttle settings. More elaborate analysis would require to take into account the time lag between control inputs and engine rotor speeds. Such an analysis may be required in helicopter maneuver, and is discussed in some detail by Newman.

2.4.9 Propeller Characteristics

Aircraft propulsion by propeller is still the most widespread method of converting engine power into useful thrust, and hence aircraft speed. The mechanism of generating thrust and torque from a propeller is the subject of propeller aerodynamics; therefore, we cannot go into details in this context, except for some basic axial momentum theory. The essential theory has been provided by Glauert and Theodorsen.

Data and propeller charts, for the purpose of basic performance calculations, can be found in some old NACA reports, such as those by Biermann and Hartman and Theodorsen and co-workers. In these reports the reader can find performance data of propellers with two to six blades. More advanced concepts, broadly related to other problems in aircraft performance, are available in AGARD CP-366.

To understand propeller performance, we need to identify the relevant parameters of operation. These parameters are the advancing speed U, the rotational speed, and the tip Mach number M_{tip}. In addition, there are several geometrical quantities that depend on the propeller: the number of blades, the diameter d, the pitch θ, the type of blade section, the chord distribution along the radius, the tip geometry, and the hub geometry.

The pitch is a measure of the orientation of the propeller on a plane normal to the axis of rotation, as shown in Figure 2.4-14. The reference line for the calculation of the pitch is the chord. Figure 2.4-14 shows how the local *inflow angle* is calculated. If the blade section is at a radial position y, the total inflow velocity is $\sqrt{U^2 + (\Omega y)^2}$. The direction of the resulting vector velocity is inclined by an angle α on the chord line: this is the local inflow angle, or angle of attack of the blade section.

The thrust generated is the resulting aerodynamic force in the direction of the forward flight. The lift and drag of the blade section are oriented in different directions to the aircraft's global forces. The aircraft's drag is parallel to the thrust element and has opposite direction.

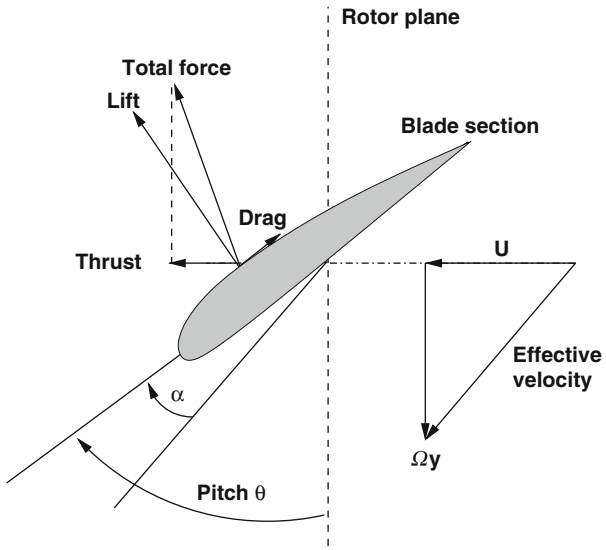


Figure 2.4-14 Nomenclature of forces on blade section.

The *solidity* σ is the ratio between the blade area (projected on the rotor disk) and the rotor disk*

$$\sigma = 2\frac{N\bar{c}}{\pi d}, \tag{2.4.22}$$

where N is the number of blades and $\bar{c}$ is the mean chord. The solidity is an important design parameter; in fact, the rotor coefficients are often normalized by σ, to express the notion of effective disk loading.

To make the data more useful from an engineering point of view, we need a measure of performance. This is the efficiency, or the ratio between propulsive power and power at the shaft

$$\eta = \frac{TU}{P}. \tag{2.4.23}$$

The propulsive efficiency expresses the ability to convert power from the engine (or power at the shaft) into useful power to fly at a speed U. The functional dependence of the efficiency from the other parameters is expressed as

$$\eta = f(U, \text{rpm}, d, \theta, \cdots). \tag{2.4.24}$$

If the aircraft is stationary on the ground, the conversion efficiency is zero, and all the shaft power generated from burning fuel is lost. The energy E is dissipated at the blades and transferred to the slipstream, which has an axial and rotational velocity. In general terms, the shaft power is

$$P = TU + E. \tag{2.4.25}$$

By using the dimensional analysis, the forward speed, the rotational speed and the diameter are replaced by another dimensionless group, the *advance ratio*

$$J = \frac{U}{\text{rpm}\, d}. \tag{2.4.26}$$

Equation 2.4.26 is just one possibility. Another definition for J is

$$J_1 = \frac{U}{\Omega R}. \tag{2.4.27}$$

The two values of the advance ratio are proportional to each other. The scaling factor is

$$J_1 = \frac{60}{\pi} J. \tag{2.4.28}$$

The advance ratio is a measure of the advancement of the propeller in one revolution, measured in number of diameters. The advance ratio is also a scaling parameter, indicating that all the propellers with the same J, and *geometrically similar*, have the same performance index. In other words, a propeller with a diameter of 2 m, operating at 2,000 rpm, and a forward speed 70 m/s ($J = 0.33$), has the same performance as a *scaled-down* propeller of 1.5 m diameter rotating at a speed of 3,000 rpm, and advancing at the same speed.

Finally, a parameter of interest is the tip speed ratio, e.g. the ratio between the propeller's speed and the tip speed:

$$\lambda = \frac{U}{\Omega R}. \tag{2.4.29}$$

For conventional aeronautic propellers, the tip speed $U_{tip} = \Omega R$ is restricted to subsonic values, in order to keep wave drag losses, noise and vibrations under acceptable limits. If we introduce the Mach number, then

$$M_{tip} = \frac{U_{tip}}{a} = \frac{1}{a}\sqrt{U^2 + \Omega^2 R^2}. \tag{2.4.30}$$

In principle, to keep the tip Mach number to subsonic values, one can operate on the aircraft speed, on the rpm and on the propeller's radius R.

The blade's Activity Factor (AF) is a non-dimensional parameter that measures the power absorbed by a propeller blade,

$$AF = 10^5 \int_{root}^{tip} \frac{c}{R}\left(\frac{r}{R}\right)^3 d\left(\frac{r}{R}\right). \tag{2.4.31}$$

* Unfortunately, this is the same symbol as the relative density of the air. We maintain the convention currently used in propeller and helicopter analysis and use σ for the solidity. Therefore, to avoid confusion, either ρ, ρ/ρ_o or σ_1 will be used for the air density.

The factor 10^5 is used to keep the value of the AF within unit values. The Total Activity Factor (TAF) is defined as the product between the number of blades N, and the single blade's activity factor,

$$TAF = N\,AF. \tag{2.4.32}$$

Among the parameters that are most effective in setting the value of the efficiency, the propeller pitch deserves special mention. One often finds propeller charts in the form of Figure 2.4-15, where the efficiency is plotted versus the advance ratio for a fixed value of the pitch. The maximum value of η depends on the advance ratio (a quantity easily measurable). As the pitch decreases, the advance ratio corresponding to maximum propulsion efficiency decreases. At all pitch settings there may be two values of the advance ratio that yield the same value of the propulsion efficiency, one below and one above the optimal advance ratio. This does not mean that the flow condition past the blades is the same. On the contrary, the point of operation past the optimal propulsion efficiency is characterized by unsteady flow separation on parts of the blade sections. This is a condition of blade stall, in a similar fashion to the elementary wing section.

The curves in Figure 2.4-15 indicate that the efficiency increases up to a maximum, then it decreases sharply; in some cases it almost plunges to zero. This is often an indication that the propeller has stalled.

Propeller charts, where the parameter is the pitch angle of the blades, are to be used in the calculation of the range and the endurance of propeller-driven airplanes. For given flight conditions, it is required to determine the rotational speed that gives the maximum ratio of the propulsive efficiency to the specific fuel consumption.

A number of other dimensionless coefficients are defined, and propeller charts can be found with these parameters, which are the power, thrust and torque coefficients:

$$C_T = \frac{T}{\rho A(\Omega R)^2}, \quad C_P = \frac{P}{\rho A(\Omega R)^3}, \quad C_Q = \frac{Q}{\rho A(\Omega R)^2 R}. \tag{2.4.33}$$

Large values of these coefficients are an indication of large thrust, power and torque absorption by the propeller. A useful relationship between the efficiency and the advance ratio involves some of the parameters in Eq. 2.4.33,

$$\eta = J\frac{C_T}{C_P}. \tag{2.4.34}$$

One type of propeller chart is shown in Figure 2.4-16. This is the advance ratio J versus the power coefficient C_P, at different values of the average C_L. One curve is the optimum propeller performance. This can be thought of as the envelope of optimal performance in the range of useful advance ratios. For a given advance ratio (e.g. operational conditions), the corresponding power coefficient is read on the vertical axis, if the blade's C_L is known.

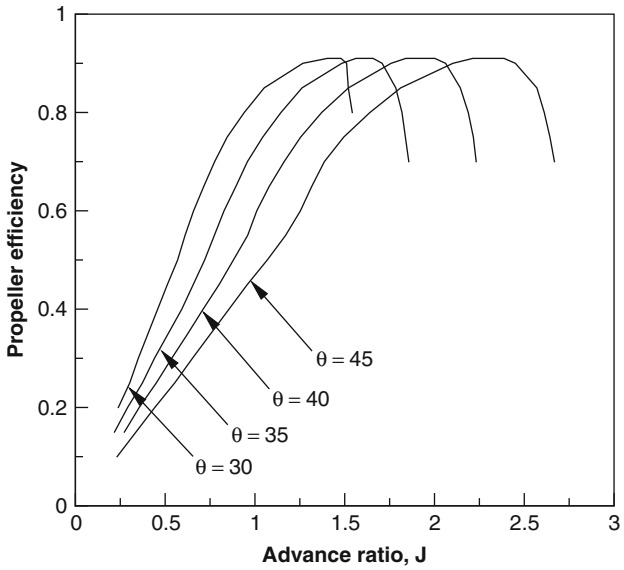


Figure 2.4-15 Propeller chart: four-bladed propeller at different pitch angles (as indicated); J versus η (data from Hamilton Sundstrand).

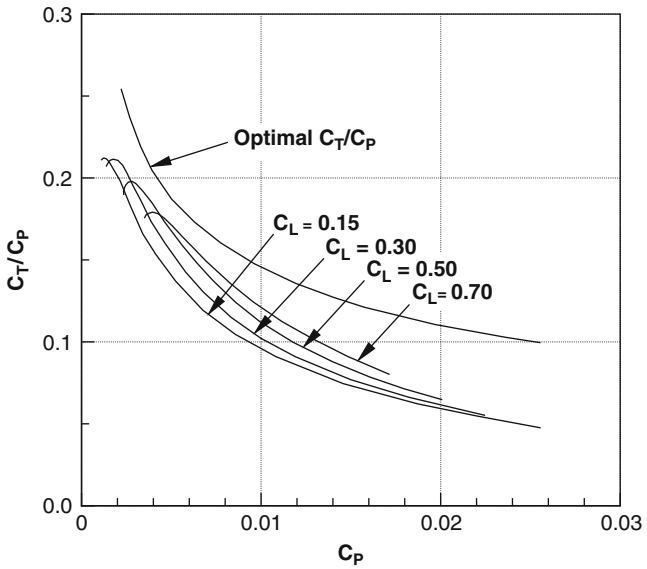


Figure 2.4-16 Propeller chart: four-bladed propeller at different design C_L (as indicated); C_P versus C_T/C_P (data from Hamilton Sundstrand).

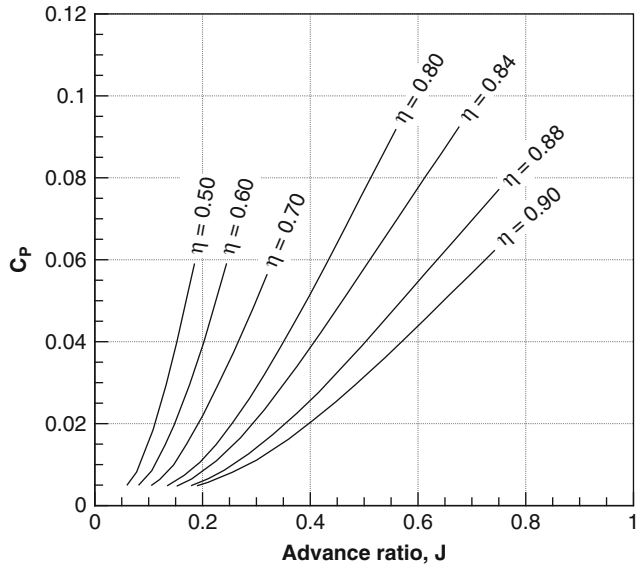


Figure 2.4-17 Propeller chart: four-bladed propeller's best efficiency in the plane $J-C_P$ (data from Hamilton Sundstrand).

Another type of chart is shown in Figure 2.4-17. The reference axes are the same. These lines of constant propeller efficiency are plotted. From Eq. 2.4.34 it is found that for a constant η, J and C_T/C_P are inversely proportional. Since C_T is positive, curves of constant efficiency have a positive slope on the plane $J-C_P$. If also C_T is constant (a reasonable approximation over a reasonable range of advance ratios), the η = constant curve is a straight line. Operation at maximum η requires operation over a limited range of advance ratios.

Although of great importance for analyzing the propeller performance, the coefficients given by Eqs 2.4.33 are not useful for determining the engine power required to fly at a given air speed, altitude and gross weight. In fact, propeller operation is analyzed either at a constant rotational speed, or at constant pitch. Operation with a mix of rotational velocities and pitch angles is also possible (see Figure 2.4-18).

For this purpose, it is necessary to have propeller data that do not depend on the rotational speed, as in Figure 2.4-15. If the pitch setting is unique, then the efficiency is a single curve, that can only be changed with the advance ratio, i.e. by a combination of rotational and forward speed.

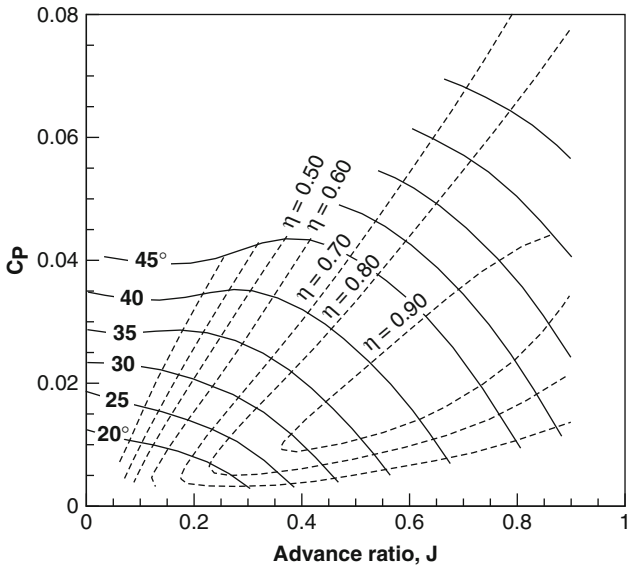


Figure 2.4-18 Propeller chart: four-bladed propeller performance. Solid lines are performance indicators at constant pitch angle (as indicated, calculated at 3/4 chord); dotted lines are lines of constant efficiency in the J versus C_P plane (data from Hamilton Sundstrand).

2.4.9.1 The Axial Momentum Theory

The most elementary theoretical mechanism for converting the rotating power of a propeller into useful thrust is based on the one-dimensional momentum theory for the flow that passes through the propeller disk (Rankine-Froude momentum theory). This theory is conceptually and practically important. First, it provides a first-order estimate of the power and thrust of an open air screw; second, the method is general and can be applied to the thrust generation from a helicopter rotor (Chapter 2.5). There is a unified theory for propellers and helicopter rotors that is useful for a wide range of performance calculations. This method is so powerful that one often forgets its shortcomings, namely the lack of reference to the blades' geometry and the rotational speed.

According to the basic momentum theory, the propeller is reduced to a rotating disk that imparts axial momentum to the air passing through it. For the purpose of this discussion the air is incompressible. For reference, consider sections 0 (far upstream), 1 (just upstream the disk), 2 (just downstream the disk) and 3 (far downstream), as shown in Figure 2.4-19. The subscripts will refer to quantities at these sections.

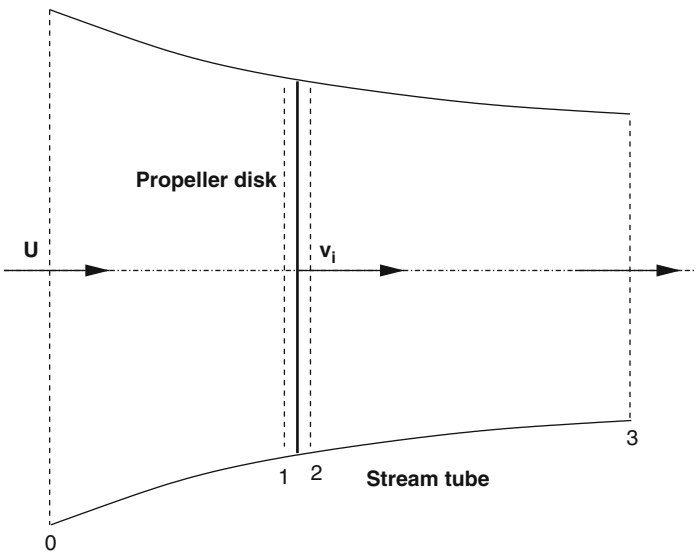


Figure 2.4-19 One-dimensional axial flow model of airplane propeller.

The theory assumes that there is no flux along the limiting streamlines and that the velocity is continuous through the propeller disk. The free stream velocity is equal to the propeller's speed $U = U_o$. The continuity equation written for the incompressible mass flow rate within the stream tube is

$$Au = const. \tag{2.4.35}$$

The propeller thrust is equal to the rate of change of axial momentum at the disk,

$$T = A(p_2 - p_1). \tag{2.4.36}$$

Now we can apply the Bernoulli equation between sections 0–1 upstream and 2–3 downstream of the propeller, Figure 2.4-19,

$$p_o + \frac{1}{2}\rho U^2 = p_1 + \frac{1}{2}\rho U_1^2, \quad p_o + \frac{1}{2}\rho u_3^2 = p_2 + \frac{1}{2}\rho U_2^2. \tag{2.4.37}$$

The difference between the two equations yields the pressure jump through the propeller

$$p_2 - p_1 = \frac{1}{2}\rho(u_3^2 - U^2), \tag{2.4.38}$$

because from the continuity equation the velocity is continuous at the rotor disk, e.g. $u_1 = u_2$. From Eq. 2.4.36 the thrust generated by the propeller becomes

$$T = \frac{1}{2}\rho A(u_3^2 - U^2) = \rho A_3 u_3(u_3 - U). \tag{2.4.39}$$

The total power is

$$P = T(u_1 + U) = \frac{1}{2}\dot{m}(u_3 + U)^2 - \frac{1}{2}\dot{m}U^2 = \frac{1}{2}\dot{m}u_3(u_3 + 2U), \tag{2.4.40}$$

that is obtained by substituting Eq. 2.4.39. The energy imparted to the slipstream is

$$E = \frac{1}{2}A_3\rho u_3(u_3^2 - U^2). \tag{2.4.41}$$

This energy is minimal when the slipstream velocity is equal to the propeller velocity. The total power is found from summing up Eq. 2.4.41 and Eq. 2.4.40. The slipstream velocity far downstream, u_3, can be related to the velocity at the disk u_1 and the free stream velocity U. In fact, the power absorbed by the propeller is

$$T = \dot{m}(u_3 + U) - \dot{m}U = \dot{m}u_3. \tag{2.4.42}$$

By combination of Eq. 2.4.40 and Eq. 2.4.42, we have

$$u_1 = \frac{1}{2}u_3. \tag{2.4.43}$$

In conclusion, the air speed at the rotor disk is the average between the propeller's speed and the speed far downstream. Also, the far downstream velocity is twice the induced velocity disk. For a static propeller this relationship provides the value of the *induced velocity* at the rotor disk:

$$u_1 = v_i = \frac{1}{2}u_3. \tag{2.4.44}$$

The corresponding induced power is $P_i = Tv_i$. The thrust becomes

$$T = 2\rho A(U + v_i)v_i, \tag{2.4.45}$$

where $\rho A(U + v_i)$ is the mass flow rate through the disk, and $2v_i$ is the total increase in velocity. The corresponding power is the product between the thrust and the velocity through the disk

$$P = T(U + v_i) = 2\rho A(U + v_i)^2 v_i. \tag{2.4.46}$$

This power expression contains two terms: (1) a useful power TU and (2) an induced power Tv_i which is a loss due to the kinetic energy imparted to the flow. The propulsion efficiency is

$$\eta = \frac{TU}{P} = \frac{TU}{T(U + v_i)} = \frac{U}{U + v_i}. \tag{2.4.47}$$

Therefore, the propulsion efficiency decreases as the induced velocity increases. The induced velocity in terms of the thrust is found from Eq. 2.4.45, which is quadratic in v_i. The only physical solution of Eq. 2.4.45 is a positive v_i,

$$v_i = \frac{1}{2}\left[-U + \sqrt{U^2 + \frac{2T}{\rho A}} \right]. \tag{2.4.48}$$

Under static conditions, $U = 0$, we have

$$v_i = \sqrt{\frac{T}{2\rho A}}, \quad P_i = \sqrt{\frac{T^3}{2\rho A}}. \tag{2.4.49}$$

These expressions will be useful for rotorcraft calculations (Chapter 2.5).

An important aspect of this elementary theory is the behavior of the pressure. For points upstream or downstream of the propeller, there is a conservation of total pressure (sum of the static pressure and dynamic pressure, $q = \rho U^2/2$). The total pressure is given by the Bernoulli equation, Eq. 2.4.37. The engine power is used to increase the kinetic energy of the air passing through the disk. Therefore, we conclude that the theory involves a sudden increase of pressure through the propeller disk, while the velocity of the air in the stream tube is continuous. The flow undergoes an acceleration; therefore, the slipstream must contract, according to the continuity in Eq. 2.4.35.

Obviously, various advancements and refinements can be done from the above assumptions. With respect to the non-uniform axial inflow, one needs to integrate the change of axial momentum over the propeller disk, and the result is

$$T = \int_A \rho u_3 (u_3 - U) dA, \tag{2.4.50}$$

A similar expression is obtained for the loss of energy:

$$E = \frac{1}{2}\int_A \rho u_3 (u_3^2 - U^2) dA. \tag{2.4.51}$$

An essential hypothesis is that of axi-symmetric flow through the disk, which is quite reasonable for a propeller in axial flight. This hypothesis leads to a slightly modified theory.

Consider an annulus of width dy, corresponding to an area $dA = 2\pi y dy$. The element of thrust generated by the mass flow is

$$\dot{m} = \rho(U + v_i) dA, \tag{2.4.52}$$

through this annulus is

$$dT = 2\rho(U + v_i) v_i dA = 4\pi\rho(U + v_i) v_i y dy. \tag{2.4.53}$$

We define the following induced velocity ratio

$$\lambda = \frac{U + v_i}{\Omega R} = \frac{U + v_i}{\Omega y}\frac{\Omega y}{\Omega R} = \left(\frac{U_n}{U_t}\right) r = \tan\phi\, r, \tag{2.4.54}$$

where U_n and U_t are the velocity components normal and tangential to the rotor plane, ϕ is the inflow angle. In general, $U_n \ll U_t$; therefore, we can make the approximation

$$\tan\phi \simeq \phi = \frac{U_n}{U_t} = \frac{\lambda}{r}. \tag{2.4.55}$$

If we introduce the inflow velocity ratio, Eq. 2.4.54, we find

$$dT = 4\pi\rho\left(\frac{U + v_i}{\Omega R}\right)\left(\frac{v_i}{\Omega R}\right)(\Omega R)^2 y dy = 4\pi\rho\lambda\lambda_i(\Omega R)^2 y dy, \tag{2.4.56}$$

with $\lambda_i = v_i/\Omega R$. Note that λ_i is the induced velocity ratio in the absence of axial flight velocity. It can also be written as

$$\lambda_i = \frac{v_i}{\Omega R} = \frac{v_i}{\Omega R} + \frac{U}{\Omega R} - \frac{U}{\Omega R} = \lambda - \frac{U}{\Omega R} = \lambda - \lambda_c, \tag{2.4.57}$$

with $\lambda_c = U/\Omega R$. The resulting thrust is found from integration of Eq. 2.4.56,

$$T = 4\pi\rho(\Omega R)^2 \int_o^R \lambda\lambda_i y dy. \tag{2.4.58}$$

The element of power is $dP = dTv_i$; therefore,

$$P = 4\pi\rho(\Omega R)^3 \int_o^R \lambda\lambda_i^2 y dy. \tag{2.4.59}$$

These integrations can only be done if the radial distribution of induced velocity is known; therefore, the problem is not closed. A solution can be found by combining the results of the blade momentum theory (next section). Various further advancements have been achieved over the years, and it is now possible to model the case of a propeller/rotor disk inclined by any angle on the free stream, with almost any load distribution on the rotor.

2.4.9.2 The Blade Element Method

The axial momentum theory provides integral quantities, such as thrust and power. These characteristics do not seem to depend on the propeller's geometry, which is clearly a shortcoming. In the blade element method, these details are yanked back into the theory. In this framework, the blade sections are supposed to operate like two-dimensional sections, with the local inflow conditions derived by appropriate means in the rotating environment, see Figure 2.4-14. Although the definition of this inflow is not obvious, and the interference between elements on the same blade and between blades is not taken into account, the method is otherwise extremely powerful. It allows the calculation of basic performance from the geometrical details and the two-dimensional blade section aerodynamics. The method described below applies to the helicopter rotor blades as well, though with some small changes.

The relationship between inflow angle, pitch angle and angle of attack is

$$\alpha = \theta - \phi. \tag{2.4.60}$$

The pitch is a geometrical setting, whilst the angle of attack of the blade section and the inflow velocity are operational free parameters. The lift and drag forces on this section are

$$dL = \frac{1}{2}\rho c C_L U^2 dy, \tag{2.4.61}$$

$$dD = \frac{1}{2}\rho c C_D U^2 dy. \tag{2.4.62}$$

These forces, resolved along the direction normal and parallel to the rotor disk give the contributions to the thrust, torque and power for the single blade

$$dT = dL\cos\phi - dD\sin\phi, \tag{2.4.63}$$

$$dQ = (dL\sin\phi + dD\cos\phi)y, \tag{2.4.64}$$

$$dP = (dL\sin\phi + dD\cos\phi)\Omega y. \tag{2.4.65}$$

These elements can be written in non-dimensional form, using the definition of coefficients given by Eq. 2.4.33. The total thrust, torque and power for N blades will require integration of the above expressions from the inboard cut-off point to the tip.

The integrals are, in fact, not solved directly, but numerically. If we divide the blade section into a number n of elements, each having a radial width dy_j, then

$$T = N\sum_{j=1}^{n}(dL_j\cos\phi_j - dD\sin\phi_j), \tag{2.4.66}$$

$$Q = N\sum_{j=1}^{n}(dL_j\sin\phi_j + dD\cos\phi_j)y_j dy_j, \tag{2.4.67}$$

$$P = N\sum_{j=1}^{n}(dL_j\sin\phi_j + dD\cos\phi_j)\,\Omega y_j dy_j, \tag{2.4.68}$$

with the forces evaluated from Eq. 2.4.61 and Eq. 2.4.62. All the quantities appearing in the aerodynamic forces change with the radial position, including the chord (for complex geometries), the density (for high-speed flows), and the C_L and C_D (depending on local angle of attack α, Reynolds number Re and Mach number M). The problem is to find the inflow angle ϕ, the resultant angle of attack α and the actual inflow velocity U at each blade section.

For the solution of the problem we still have too many unknowns and too few equations. One way of finding the induced velocity ratio is to introduce the axial momentum theory. This combination eliminates the unknowns discussed above, namely the radial distribution of induced velocity v_i.

With the approximation introduced by Eq. 2.4.55, the element of thrust can be written as $dT \simeq dL$. Since the results of both theories must be the same, we are led to the equivalence between Eq. 2.4.56 and Eq. 2.4.61

$$\begin{aligned} dL &= \frac{N}{2}\rho c C_L U^2 dy = \frac{N}{2}\rho c C_{L_\alpha}\alpha U^2 dy = dT \\ &= \frac{n}{2}\rho c C_{L_\alpha}\alpha U^2 dy = 4\pi\rho\lambda\lambda_i(\Omega R)^2 y dy. \end{aligned} \tag{2.4.69}$$

Now simplify Eq. 2.4.69, by using: $U = \Omega y$, Eq. 2.4.60 (definition of angle of attack), Eq. 2.4.55 (small inflow angle approximation), and Eq. 2.4.22 (definition of solidity). After this algebra, the result (Problem 8) is

$$\frac{1}{8}\sigma C_{L_\alpha}(\theta r - \lambda) = \lambda(\lambda - \lambda_c). \tag{2.4.70}$$

Equation 2.4.70 must be solved for the unknown inflow velocity ratio λ. It is a quadratic equation with a positive and a negative root. The meaningful solution is

$$\lambda = \left[\left(\frac{\sigma C_{L_\alpha}}{16} - \frac{\lambda_c}{2}\right)^2 - \frac{\sigma C_{L_\alpha}}{8}\theta r\right]^{1/2} + \left(\frac{\sigma C_{L_\alpha}}{16} - \frac{\lambda_c}{2}\right). \tag{2.4.71}$$

The calculation of the rotor power is done in the same way. All the quantities are known from the previous discussion, and the power can be calculated alongside the thrust, from the equation

$$P = N\sum_{j=1}^{n}(dL_j\phi_j + dD)\,\Omega y_j dy_j. \tag{2.4.72}$$

Computational Procedure

1. Read propeller's geometry (radius, chord, twist distributions).
2. Read operational parameters (rpm, U, density altitude).
3. Read the two-dimensional C_D, and C_L,C_{L_α}, distributions.
4. Divide the radius in a number n of blade elements.
5. Set the radial position (*start loop*).
6. At current r solve Eq. 2.4.70 to find λ, Eq. 2.4.71.
7. Solve Eq. 2.4.55 to find ϕ.
8. Solve Eq. 2.4.60 to find α.
9. Solve Eq. 2.4.69 to find the thrust dT.
10. Update the thrust.
11. End loop.

A correction is usually applied to the above procedure, because it implies that the aerodynamic flow around the blade section follows that of the airfoil. However, both the tip and the root are affected by three-dimensional effects, which lead to loss of lift, and hence loss of thrust. One of these corrections is due to Prandtl, and is written as

$$k(r) = \left(\frac{2}{\pi}\right)\cos^{-1}e^{-f(r)}, \tag{2.4.73}$$

with

$$f(r) = \frac{N}{2}\frac{1-r}{r\phi}. \tag{2.4.74}$$

The corrected thrust will be $dTk(r)$. A plot of $k(r)$ shows that it is equal to one for most of the span, but it tends rapidly to zero near the tip, with a rate depending on the number of blades and the inflow angle. The alternative to this method is to consider an effective blade radius $R_e < R$, and proceed with the calculations as described above.

The procedure described neglects a number of difficulties. These are related to the format in which the aerodynamic data are available. Ideally one would have data like $C_D = C_D(\alpha, M)$, $C_L = C_L(\alpha, M)$ in a matrix form.

From the performance point of view, it is important to know how to operate the propeller engines, so as to minimize losses and increase fuel efficiency. For a given propeller installation, the parameters that can be changed are the pitch, the rotational speed and the flight speed. Other important aspects of propeller performance are the interaction with the wing and the fuselage[174,175], compressibility effects[176,177], and the propeller noise.

Problems

1. A turbofan engine has a specific fuel consumption between 0.368 and 0.385 lb/lb/h. Convert these data into units kg/N/s. Is this a conversion into international units? Finally convert the data into kg/kN/s.
2. A certain aircraft has a fuel flow at cruise conditions estimated at 1,400 US gallons per hour. Convert this data into *N/s*. Also, assume that the aircraft has two engines, and that the total thrust delivered at cruise conditions is $T = 210$ kN. What is the specific fuel consumption of the engines in *N/N/s*? (For aviation fuel use the data in Table 2.4-2).
3. The Boeing 747-400 is known to burn fuel at the rate of about 3% of its take-off gross weight every hour. Calculate the amount of fuel in metric tons per hour, in liters per hour, and the number of liters required to fly 1 km at a cruise altitude $h = 11{,}000$ m.
4. Discuss the parameters affecting the mass flow rate into a jet engine. Provide some qualitative plots of $\dot{m}_a$ as a function of flight altitude, Mach number, engine diameter, and rpm. Use the St Venant equation to calculate the chocked mass flow rate per unit nozzle area as a function of flight altitude.
5. The Boeing B-747/400 has its power plant re-engineered. As a result, its TSFC is decreased by 3%. The other characteristics of the aircraft do not change significantly; therefore, its weight can be considered constant. The aircraft is to have an annual block time of 700 hours. Calculate the saving in fuel consumption over the year. Also, considering that the fuel consumption represents 9% of the direct operating costs of the aircraft, calculate the money that can be saved by the improved engines. Consider constant aviation fuel price, equal to 0.70 US dollars/liter. Repeat the analysis for the SAAB 340, and draw a conclusion regarding the effects of fuel efficiency on the operation of the two aircraft. (*Please note that aviation fuel prices change greatly over short times, and depend on the airport and the quantity purchased.*)
6. As an example of pollution created in the high atmosphere by commercial air traffic, we want to calculate the heat released per unit of time by aircraft model A flying at its cruise altitude (10,800 m) at its cruise speed ($M = 0.80$) at ISA conditions. The Mach number of the jets from the nozzle of its two jet engines is $M = 1.3$, and the nozzle cross-sectional area is $A_n = 1.8$ m^2, and the average jet temperature is estimated at $T_j = 850$ degrees.
7. Consider a four-bladed propeller, whose performance charts are given in Figures 2.4-14 and 2.4-15. This propeller has a diameter $d = 1.8$ m, and is coupled to an engine that delivers a maximum shaft power $P = 300$ kW. Its maximum tip Mach number at sea level is $M_{tip} = 0.70$. Determine if this propeller can be operated with an efficiency $\eta = 0.90$.
8. Derive Eq. 2.4.70 from Eq. 2.4.69 and calculate its only physical solution, Eq. 2.4.71.

Aerospace Engineering Desk Reference; ISBN 9781856175753

2.5 Rotorcraft

Antonio Filippone

Whereas the airplane was developed by young men not previously known as inventors … the fortunes of the helicopter have been more in the hands of those with reputation previously acquired elsewhere.

E.P. Warner, 1922

This chapter introduces the direct-lift vehicles, e.g. vehicles that generate lift by the action of one or more rotors. This category includes helicopters, autogiros and con-vertiplanes.

The flight conditions of a conventional helicopter include hover, axial climb and descent, forward (inclined) climb and descent, level forward flight, sideways, backward and yawed flight, turning and maneuvering, autorotation, and flight conversion. The flight condition between hover and level flight at moderate speeds is called flight *transition.* Besides these conditions, we must consider off-design flight operations, such as the vortex ring state, autorotation and cornering flight. Therefore, the number of possible flight conditions of a rotary-wing aircraft exceeds that of a fixed-wing aircraft.

A full performance analysis, aimed at selecting or designing a helicopter, planning a mission, or upgrading a vehicle, must include at least range, endurance, hover in- and out-of-ground, climb rates at full power and one engine inoperative (OEI), fuel consumption, and performance in non-standard atmosphere (cold or hot day).

2.5.1 Fundamentals

It often happens that the rotor technology dominates any discussion on the helicopter. Indeed, some essential geometrical parameters, such as the airfoil section, the shape of the blade tips, the blade twist, and the articulation of the hub, are very important in terms of overall rotorcraft performance.

Rotary-wing vehicles are far more complex than fixed-wing aircraft, because of the presence of rotating parts, their aerodynamics and the interference with the airframe. There are additional stability, control, vibration and noise problems, and a variety of flight conditions peculiar to the rotary-wing aircraft. Furthermore, the blades are highly flexible, and are required to perform complicated movements as they travel around the hub (flapping, pitching/feathering, lagging).

The original idea of the helicopter pioneers was to develop a direct-lift aircraft that would be able to hover and to convert its flight from vertical to horizontal, and vice versa. This concept, simple as it looked, required over 30 years of experiments by engineers and scientists around the world. The history of the development of the practical helicopter is one of the most fascinating and controversial subjects in aviation. One of the contentious issues is whether stable hovering is enough to grant anyone the right to the invention. Individuals who have claimed to have invented the helicopter include *at least* Etienne Oehmichen, Louis Breguet, Corradino D'Ascanio, Nicolas Florine, Anton Flettner, Heinrich Focke and Igor Sikorsky. However, the helicopter was at a dead end until Juan de la Cierva invented the cyclic pitch control. His system provided the means for handling the structural loads, and was a major step forward in the development of both the helicopter and the autogiro.

Maneuverability takes forms not seen in any other flight vehicle. The helicopter can turn around itself on the spot, pull up, do a split-S maneuver, and fly backward and sideways at great speeds. For example, the Kamov Ka-52 Alligator claims a maximum backward speed of 90 km/h, and the Boeing-Sikorsky RAH-66 Comanche would be capable of a backward speed of 130 km/h. Therefore, the helicopter is a highly versatile flight vehicle that has reached phenomenal levels of sophistication.

For reference, Jane's publishes detailed listings on all helicopters in service, under development and upgrading, with plenty of technical data. A wealth of information is made available by the FAA/JAA to support the helicopter certification standards and procedures (FAR 29 of the Airworthiness Standards: transport category rotorcraft). These standards are becoming increasingly important, because they are used also for the certification of military helicopters. The proof of compliance is specified by FAR.

2.5.2 Helicopter Configurations

Before introducing a general description of the helicopter types we need to address a fundamental flight problem: torque reaction. As the rotor spins in one direction, it does work on the air, and needs a torque to drive it. This torque is provided by the engine, and it must be reacted by the airframe. Torque reaction causes the airframe to spin in the opposite direction, unless the torque is balanced by other means: a direct moment or a force through an offset with respect to the rotor shaft. In the first case a counter-rotating rotor is required; in the latter case full control is achieved with a tail rotor, as shown in Figure 2.5-1.

Helicopters are primarily classified on the basis of their rotor system, on their gross weight and the type of operation they perform (civil, commercial, military). There is also a convention regarding the weight of the aircraft. A light helicopter has a maximum take-off weight (MTOW) below 12,000 lb (5,450 kg); a medium helicopter has an weight up to 45,000 lb (10,000 kg), although this category includes several heavy helicopters; heavy-lift helicopters have an MTOW above this value. It is estimated that 95% of the helicopters operating worldwide are light helicopters; most of them have the conventional tail rotor configuration, as in Figure 2.5-1.

Figure 2.5-2 shows the most common configurations. Most helicopters have a single main rotor, slightly inclined forward, and a tail rotor, as illustrated in Figure 2.5-1. This is nearly vertical and nearly aligned with the longitudinal axis of the airframe for torque balance, stability, control and maneuver. Under this class there are all the general utility vehicles, heavy-lift vehicles, light helicopters and most of the high-performance military helicopters.

Another class of vehicle is characterized by two main rotors: these can be coaxial and counter-rotating, tandem (two separate axes), and intermeshing (with the rotor disks partially overlapping and the shafts inclined outward). Other helicopters may lack the tail rotor and owe their control and stability to other technologies, such as NOTAR, which is based on jet propulsion. Table 2.5-1 is a summary of helicopter types based on the torque reaction system.

For all types of helicopters vertical control is achieved by means of the rotor thrust. Longitudinal control is achieved with the collective pitch. Lateral control is achieved by a combination of cyclic pitch and tail rotor thrust. A summary of control systems is reported in Table 2.5-2.

There are other types of helicopters, mostly of experimental nature, but these will not be discussed here. These include the compound helicopter, which derives part of its lift from conventional fixed wings. Compound aircraft that have been designed, tested and built include the McDonnell XV-1 (1954), the Fairey Rotodyne (1957), the Hiller X-18 (1959), the Lockheed AH-56 Cheyenne, the Kamov Ka-22 and more recently the Sikorsky-Piasecki X-49. See Deal and Jenkins[331] for some flight research on this type of aircraft, and Stepniewski and Keys[110] for a preliminary design analysis.

Figure 2.5-3 shows NASA's research aircraft RSRA (1980s). The aircraft was designed to study the effects arising from stopping the main rotor in flight. The lift would then be provided with the conventional wings mounted low on the fuselage. This concept gave the aircraft the vertical stability of a helicopter, and the horizontal cruise capability of a fixed-wing aircraft, as documented by Erickson and co-workers. A critical problem in flight stability and vibration is the off-loading of the rotor and the reduced tip speed.

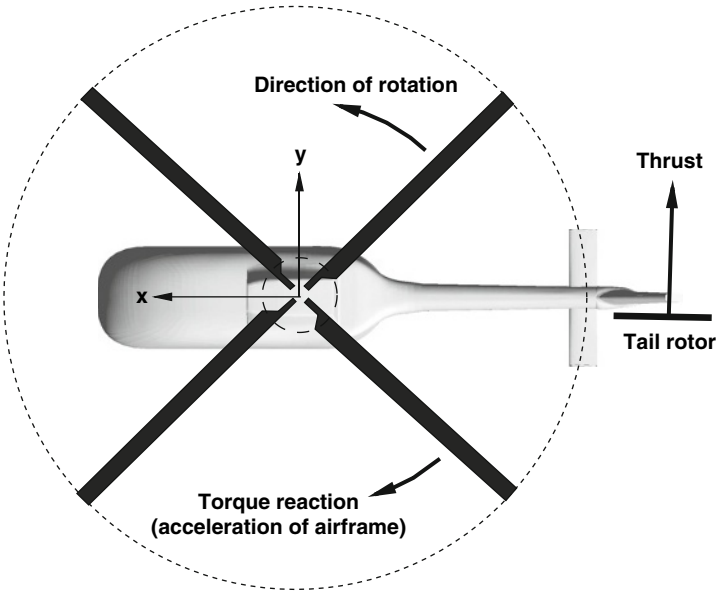


Figure 2.5-1 Torque reaction phenomenon; torque reaction is in the opposite direction.

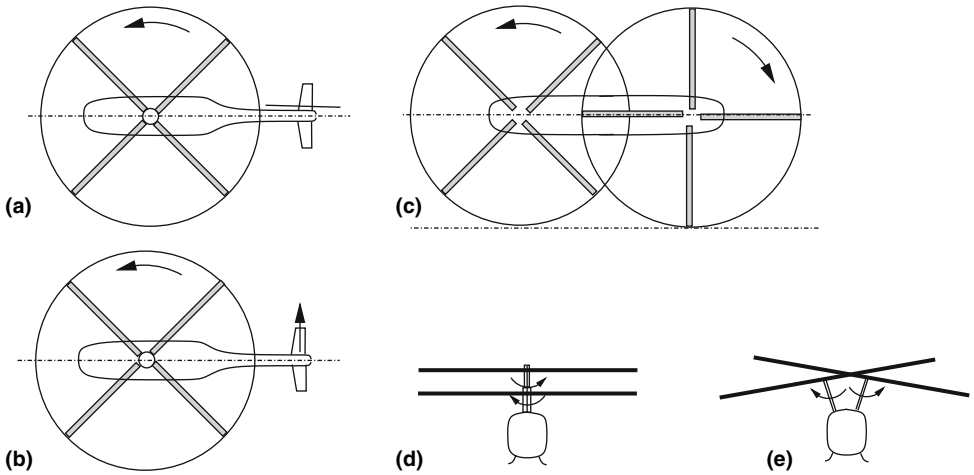


Figure 2.5-2 Helicopter configurations, with various anti-torque systems: (a) conventional helicopter; (b) tandem; (c) NOTAR; (d) coaxial; (e) intermeshing.

Table 2.5-1 Summary of main helicopter configurations, based on yaw control systems.

Vehicle	Rotor type	Tail rotor
Eurocopter EC-365	Single rotor	Fenestron
Sikorsky S-76	Single rotor	Yes
Boeing CH-47	Tandem	No
Kamov Ka-52	Counter-rotating	No
Kaman K-max	Intermeshing	No
Boeing-Sikorsky V-22	Tiltrotor	No
Boeing MD-500	NOTAR	No

Table 2.5-2 Summary of helicopter control methods; *p*, *q*, *r* are pitch/roll/yaw rates, respectively.

Configuration	*x*	*y*	*z*	*p*	*q*	*r*
Single rotor	MR tilt fore/aft	MR tilt lateral	MR thrust collective	MR tilt fore/aft	MR tilt lateral	TR thrust
Tandem	Rotors tilt fore/aft	Rotor tilt lateral	Rotor thrust collective	Rotors differential	Tilt lateral	Rotors differential
Coaxial	Rotors tilt fore/aft	Rotor tilt lateral	Rotor thrust collective	Tilt fore/aft	Tilt lateral	Rotors differential
Intermeshing	Rotors tilt fore/aft	Rotor tilt lateral	Rotor thrust collective	Tilt fore/aft	Tilt lateral	Rotors differential

The autogiro is another hybrid vehicle that derives its lift from a rotating blade and its forward thrust from a conventional propeller. This type of aircraft (the original idea of Juan de la Cierva) has some flaws, and it has never been successful beyond the realm of aviation enthusiasts. Nevertheless, the autogiro continues to stir attention in the aviation community. For a given GTOW, modern autogiros require the same power levels as conventional helicopters at intermediate flight speeds.

2.5.3 Mission Profiles

The mission profiles of the helicopter include many of the operations of the fixed-wing airplane. Due to its peculiar flight characteristics, the helicopter can carry a wide range and type of payloads. This is true both in the civil and military arena. Civilian operations include: scheduled flight services between airports and heliports, search and rescue,

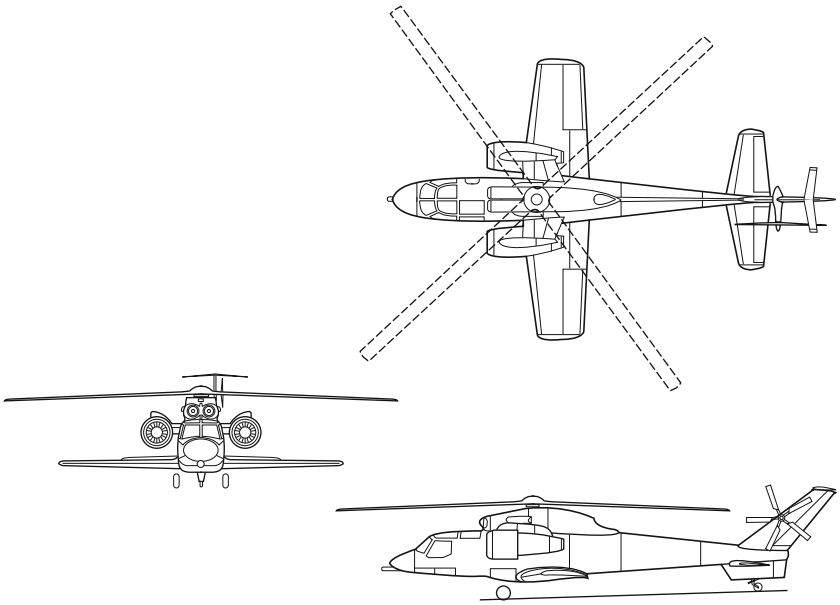


Figure 2.5-3 NASA's RSRA compound helicopter configuration.

disaster relief, traffic monitoring, policing, and executive services. A review of the roles of the civil helicopter has been provided by Fay.

Many civil helicopters are approved for operation from/to challenging heliports, such as off-shore platforms (even in severe weather conditions), aircraft carriers, and rooftops of buildings. They are allowed to fly over areas where there are no emergency landing sites (Category A). For such operations the helicopter must be able to take off with OEI, and achieve at least a 100 feet-per-minute (0.5 m/s) rate of climb. Helicopters that do not fall under this category are certified to operate in safer areas, where landing is always possible (Category B). In any case, the certification authorities (see, for example, FAR Part 29) require a height/velocity diagram, both with single and dual engine operation, to be included in the flight manual. The height/velocity diagram is a flight envelope in which the operation of the aircraft is possible and safe. Take-off operations are affected by the atmospheric conditions, helipad size and position, and community constraints.

Many helicopters currently operate within large metropolitan areas and have access to a good network of helipads. Back in 1951, when the first helicopter operations started at the heart of Manhattan, the press was stunned by helicopters "taking off and landing from 16-storey skyscrapers".

Manufacturers offer different helicopters in the same nominal series, depending on applications. Each of them has slightly different performance charts and different flight manuals. An example of this is the Sikorsky S-76C, available in utility/off-shore configuration, or executive configuration.

Helicopters for sea and naval operations include Anti-Submarine Warfare (ASW), anti-surface ship warfare, medical evacuation *(medevac)*, logistics, search and rescue, and mine-sweeping*. Other military operations include long-range missions in hostile territory, to be conducted day and night, in adverse weather, and extreme environments.

Mission planning for commercial and passenger traffic can be slightly more complicated than a flight mission of a fixed-wing aircraft, essentially because helicopters tend to fly at lower altitudes around congested and built-up corridors. Flight planning in these cases needs permission to fly over these areas. In addition, it may require studying different flight corridors to minimize community noise, which depends heavily on local weather conditions. The flight altitude for each segment should be the maximum allowed by the ATC, in order to minimize noise disturbances and maximize safety. In London this altitude is 2,500 ft (770 m).

2.5.4 Flight Envelopes

The flight envelopes circumscribe the entire issue of rotary-wing performance. All the rotary-wing aircraft in Figure 2.5-2 are capable of vertical take-off and landing; they are also capable of transitioning to low-speed forward

* The US Army has different designations for the same class of vehicles: OH is an observation helicopter; UH is a general utility, CH is a cargo/heavy-lift helicopter; and AH is an attack/combat helicopter.

flight and to turn around themselves. Therefore, the helicopter represents a unique flight vehicle that extends the spectrum of powered flight possibilities, although it is not in direct competition with the fixed-wing airplane, as far as speed is concerned.

In 1987, Drees wrote that *speed will certainly be increased to the 450 kt level.* That has not happened yet, nor does there seem to be serious attempts at reaching these flight performances with a rotary-wing aircraft. Nevertheless, research has continued at various institutions. An exhaustive report at Sikorsky concluded that for commercial transports, tilt-wings and variable diameter tilt-rotor concepts have a better performance. For a military/attack role, the variable-diameter helicopter was best. A design speed of 375 to 425 kt was found to be the maximum desirable for transport missions. Research in high-speed rotorcraft at McDonnell–Douglas is documented by Rutherford and co-workers. These authors discuss options for rotorcraft speeds in the range of 450 kt with rotor-wing and tilt-wing options. Alternative configuration options have been provided by Talbot and co-workers. It was concluded that, with the exception of autorotative performance and the large download at low speeds, the compound helicopter has several advantages at high speed.

The level flight speed record is owned by the Westland Lynx helicopter, with 400.87 km/h (216.3 kt), thanks to a combination of installed power and the helicopter's stability at high speeds. The actual speeds of the common helicopter are considerably lower, in the range of 250 to 280 km/h (e.g. up to 150 kt).

The maximum speed is limited by a large number of factors, not directly related to the installed engine power. These factors include the tip speed, the dynamic stall, vibration, stability, maximum transmission torque, noise emission, etc.

At present, high-altitude helicopter flight has a limited commercial market, although it is of some interest for mountain rescue operations, for military operations and atmospheric research. In fact, most helicopter flights take place at low altitudes, up to 1,000 m (3,000 ft) from ground level.

2.5.5 Definitions and Reference Systems

To properly understand the operation and performance of the helicopter, it is necessary to explain the nomenclature and the reference systems commonly used. Figure 2.5-4 shows the nomenclature in the conventional symmetry plane of the aircraft, x–z. We note that the aircraft is generally not symmetric on this plane for several reasons (presence of tail rotor, asymmetric tail planes and control surfaces, blade orientations, windows, probes, etc.). To start with, we will have a reference system that moves with the helicopter. This system is centered at the center of gravity (CG), which is a variable point. The x axis is on the symmetry plane, and runs from the tail to the nose of the aircraft. The y axis is at 90 degrees with the y axis and pointing to starboard (from the point of view of the pilot); the z axis forms a right-hand orthogonal reference with x and y.

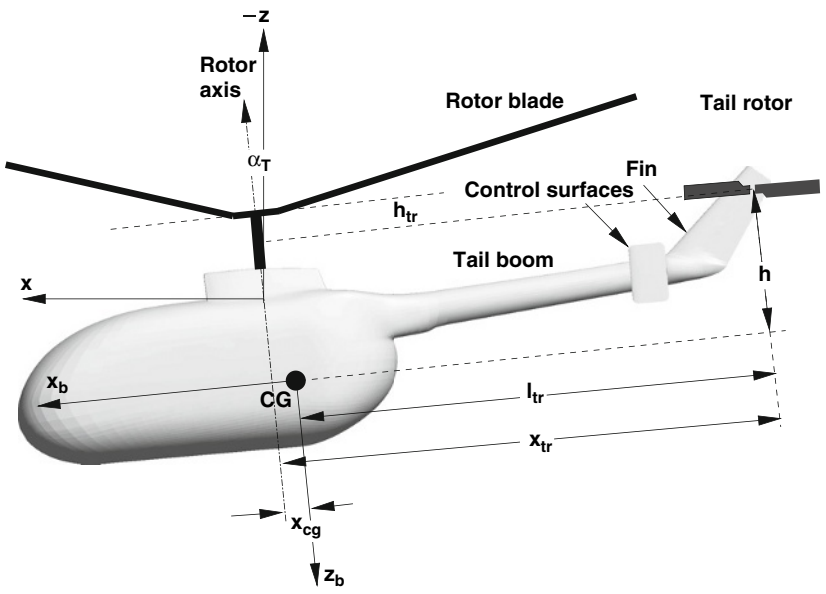


Figure 2.5-4 Helicopter nomenclature and conventions.

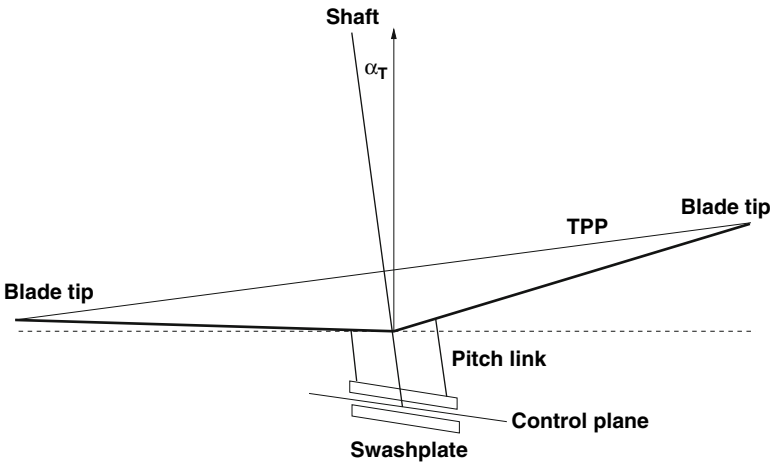


Figure 2.5-5 Tip Path Plane (TPP), or rotor disk.

Separate reference systems are required for the rotors. The main rotor is mounted on a shaft, slightly tilted forward. The tilt angle (or mast angle), calculated with respect to the vertical axis on the ground, is called α_T. As they rotate, the blade tips travel along a flight path that lies on a plane, generally called Tip Path Plane (TPP). There can be considerable differences between the TPP and the rotor disk. However, in the performance calculations the TPP is assumed to coincide with the rotor disk plane. Figure 2.5-5 shows the plane traced by the blade tips in their rotation around the shaft.

Another important plane from the control point of view is the swashplate plane. This is the reference plane for the cyclic pitch commanded by the pilot. The swashplate is a mechanical system consisting of two disks placed as a sandwich between some roll bearings. The upper disk rotates around the lower disk, and both can assume an arbitrary orientation, commanded by some actuators; the vertical movement of the swashplate commands the cyclic pitch of the blades via a series of pitch links.

The rotor blades are flexible and are subject to periodic loads that create lead/lag, flapping and torsion. For reference, in normal operations these movements are of the order of a few degrees, and are important for the blade's dynamics. In addition, there can be a complex hinge system (in articulated rotors), as sketched in Figure 2.5-6, that allows for a number of degrees of freedom.

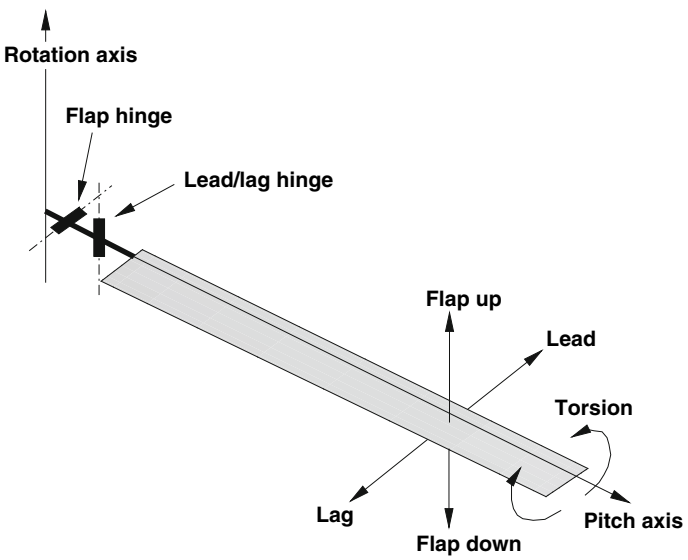


Figure 2.5-6 Articulated rotor with flap and lead hinges. The actual sequence of hinges depends on the particular rotorcraft.

The tail rotor is nearly vertical, and nearly aligned with the conventional symmetry plane. In fact, the rotor plane is inclined forward and upward by a small angle. A separate reference system centered at the hub of the tail rotor is required for detailed tail rotor calculations (aerodynamics, aeroelastic response, noise emission).

The balance of forces and moments is a more complex matter, since the main rotor thrust does not act at the CG. Furthermore, there is a set of moments acting on the aircraft, namely the rolling moment and the hub moment; additional global forces are created by the horizontal tail plane. The full balance equations and the practical means to achieve the balance in flight is called *trim*. Figure 2.5-4 shows the relevant distances required for preliminary performance and stability calculations.

2.5.5.1 Rotor Parameters

The nomenclature used for the blades and blade sections is similar to the propellers. In particular, the blade pitch θ, blade chord c, rotor solidity σ, tip speed U_{tip} and tip Mach number M_{tip} have the same definition. The definitions of thrust, power, and torque coefficients are, respectively,

$$C_T = \frac{T}{\rho A(\Omega R)^2},\ C_P = \frac{P}{\rho A(\Omega R)^3},\ C_Q = \frac{Q}{\rho A(\Omega R)^2 R}. \tag{2.5.1}$$

Since the power is $P = Q\Omega$, from the definition it follows that $C_Q = C_P$. Parameters that are specific to the helicopter include: the blade azimuth ψ, the coning angle β, the advance ratio μ, and the weight coefficient C_W. This parameter is defined as

$$C_W = \frac{W}{\rho A(\Omega R)^2}. \tag{2.5.2}$$

This is a normalized weight, just as the C_L is a normalized lift force. The *azimuth angle* ψ is required to identify the position of the blade during its rotation, and the corresponding loading. A blade oriented along the principal axis of the aircraft, from nose to tail, will have zero azimuth. After 90 degrees, it will be in the so-called advancing position; at $\psi = 180$ degrees it will be aligned with the longitudinal axis, tail to nose; at $\psi = 270$ the blade will be *retreating*.

The *coning angle* β is the angle between the blade's mean line and the rotor plane. Coning is due to a cross-combination of geometric settings, flapping and pitching of the blades. The coning angle changes as the blades rotate, because the equilibrium between centrifugal forces and aerodynamic forces is a differential equation depending on the azimuth angle.

The angle of attack α in forward flight is the angle between the rotor disk and the free stream velocity. The angle of attack of the fuselage is defined in the same way as an airplane; this can be different from the rotor's angle of attack, depending on flight conditions. The angle $-\alpha_T$ is the angle of inclination of the thrust in the direction of the force of gravity.

The *advance ratio* is the ratio between the air speed parallel to the TPP and the tip speed, i.e.

$$\mu = \frac{U\cos\alpha}{U_{tip}} = \frac{U\cos\alpha}{\Omega R} = \frac{U_T}{\Omega R}, \tag{2.5.3}$$

where $U_T = U\cos\alpha$ is the in-plane velocity component and α is the angle of attack of the rotor. The effective rotor loading is defined in terms of the blade loading coefficient C_T/σ,

$$\frac{C_T}{\sigma} = \frac{1}{\sigma}\frac{2T}{\rho A U_{tip}^2}. \tag{2.5.4}$$

This is an average measure of blade loading. A parameter that influences the stall limit is the propulsive force coefficient,

$$\frac{C_{D_e}}{\sigma} = \frac{1}{\sigma}\frac{2D_e}{\rho A U_{tip}^2}, \tag{2.5.5}$$

where D_e is the effective drag in the flight direction, and C_{D_e} is the effective drag coefficient. Estimated values of the rotor solidity are shown in Figure 2.5-7. These have been elaborated for sets of rotor data. These data can also be plotted as a function of the MTOW. We conclude that most values of σ fall within the range 0.04 and 0.14.

Helicopter rotor rpm are constant. The main reason why the rpm is kept constant is that even a small change over the design rotational speed can lead some subsystems into resonance problems. In fact, it can be proved that there is

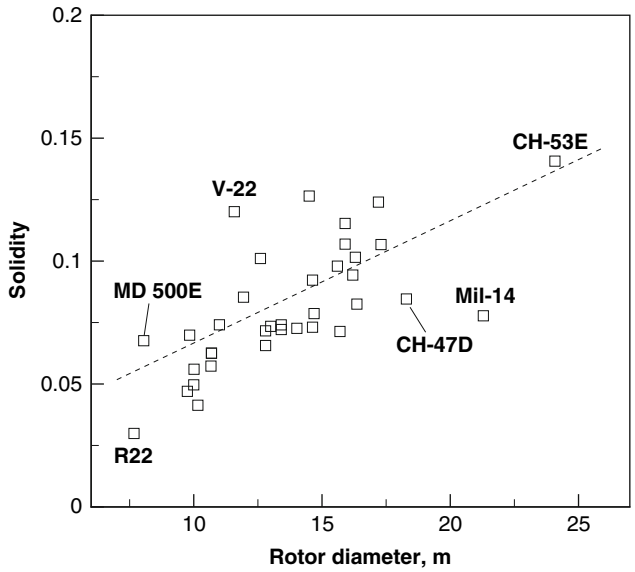


Figure 2.5-7 Estimated rotor solidity for selected helicopters. Note the scatter among various helicopters.

only a narrow range over which the rpm can be varied without operating within the natural bending and flapping frequencies of the blades. Until effective methods for avoiding resonance are found, constant rotor speeds will be commonplace in helicopter engineering[†].

2.5.6 Non-Dimensional Parameters

An effective method for first-order analysis of helicopter performance is the use of a few non-dimensional parameters, as first proposed by Knowles. Cooke and Fitzpatrick use extensively this method for flight test analysis. The power required by the helicopter in any flight condition is a function of a number of parameters, namely

$$P = f(W,\ R,\ v,\ v_c,\ h,\ \rho,\ \mathcal{T},\ \Omega), \tag{2.5.6}$$

with the usual meaning of the symbols. Helicopter performance must be calculated and tested over a range of weights, rotor speeds, forward speeds, climb rates, altitudes and atmospheric conditions. Therefore, the parametric space is large.

The non-dimensional rotor speed will be $\omega = \Omega/\Omega_o$, with Ω_o the nominal rotor speed. The relative temperature is $\theta = \mathcal{T}/\mathcal{T}_o$, and the relative density is $\sigma_1 = \rho/\rho_o$.[‡] One suitable normalization for the power is

$$\frac{P}{\rho A U_{tip}^3} = \frac{P}{\rho A \Omega^3 R^3} = \frac{P}{\rho_o \sigma_1 \pi \Omega^3 \omega^3 R^3} \propto \frac{P}{\sigma_1 \omega^3 R^5}. \tag{2.5.7}$$

Note that the non-dimensional power in Eq. 2.5.7 (also called *referred power)* is proportional to the C_p, defined in Eq. 2.5.1. The weight W can be normalized according to

$$\frac{W}{\rho A U_{tip}^2} = \frac{W}{\rho A \Omega^2 R^2} = \frac{W}{\rho \pi \Omega^2 R^4} \propto \frac{W}{\sigma_1 \omega^2 R^4}. \tag{2.5.8}$$

Note that the non-dimensional weight in Eq. 2.5.8 (also called *referred weight)* is proportional to the C_w, defined in Eq. 2.5.2. For a helicopter with a fixed rotor radius, the parameter R can be dropped from the functional in Eq. 2.5.6. The forward speed can be normalized with the tip speed, to yield the advance ratio. However, another expression is often used in flight performance analysis: $U / \Omega R \propto U/\omega$. The same normalization is carried out for the climb rate v_c.

[†] As a matter of fact, the rpm does change within a limited range, due to engine power transients. These changes, however, are usually contained within 10 rpm.

[‡] We use σ_1 in place of the standard σ to avoid confusion with the symbol used earlier in the fixed-wing aircraft analysis; θ is not to be confused with the blade pitch.

In summary, we have

$$\frac{P}{\sigma_1\omega^3} = f\left(\frac{W}{\sigma_1\omega^2}, \frac{U}{\omega}, \frac{v_c}{\omega}, h, \theta, \omega\right). \tag{2.5.9}$$

The parameters U/ω and v_c/ω are called *referred true air speed* and *referred climb rate*, respectively. Since the direct measurement of density is not possible, σ_1 is generally replaced by δ/θ. An alternative expression of the referred weight is

$$\frac{P}{\delta/\sqrt{\theta}} = f\left(\frac{W}{\delta}, \frac{U}{\omega}, \frac{v_c}{\omega}, h, \frac{\omega}{\sqrt{\theta}}\right). \tag{2.5.10}$$

The parameter $\omega/\sqrt{\theta}$ gives the effects of the air temperature on the rotor speed. In conclusion, the helicopter performance can be referred to a limited number of non-dimensional parameters. Not all of them will be present at all times. For example, the hovering analysis requires that $U/\omega = v_c/\omega = 0$. For a forward flight analysis $v_c/\omega = 0$. A forward climb flight will depend on all the parameters. A flight procedure requiring a constant referred weight W/δ requires aircraft to climb as fuel is burned. In fact, as the weight decreases, a constant referred weight is achieved by climbing to a lower pressure altitude.

2.5.7 Methods for Performance Calculations

The fundamentals of propulsion and aerodynamics of the rotor are discussed in a number of textbooks. Glauert was one of the first to treat the aerodynamics of the air-screw from a general point of view, followed by Theodorsen, although the original ideas go back to Rankine (1865) and Froude (1889). Since the publication of Glauert's theory, a number of other developments have taken place. For example, a comprehensive treatment of the momentum theory for helicopter performance (hover, climb, descent, forward flight, flight restrictions) was published by Heyson.

The first relevant studies on helicopter performance were due to Bailey and Gessow and Tapscott. These authors, along with Tanner and others, show the rotorcraft performance in terms of the rotor's non-dimensional coefficients. Gessow and Myers have included chapters on performance in their classic book on helicopter aerodynamics, and they are the authors of several technical papers on the subject, dating from the 1940s. More modern books on the subject include those by Stepniewski and Keys, Prouty and Leishman. All of these textbooks are geared in part toward rotorcraft performance. Cooke and Fitzpatrick include a discussion on helicopter performance in their book devoted to helicopter test and evaluation. Stepniewski and Keys extensively discuss a case study of helicopter preliminary design and performance analysis. This code uses a six-degree of freedom model for the aircraft, with rigid-body airframe and rotor dynamics. The rotor aerodynamics are modeled with the blade element theory. A similar approach is followed by Eurocopter's STAN flight mechanics code. Some standard methods for the calculation of helicopter performance are available in the ESDU publications, particularly on hover and forward flight.

Since all the calculation methods are necessarily complicated by the unsteady effects of the rotating blades, and by the interference between various components, we need to address a key point regarding suitable methods for the evaluation of the rotorcraft performance. We will show how simple methods based on the combined momentum and blade element theories yield steady state solutions suitable for most performance calculations. Obviously, we can improve on these methods, but the step forward requires taking into account the unsteady aerodynamics of the rotor, the full rotor dynamics, the aeroelastic response, and the interaction between rotors and airframe. Therefore, we find it appropriate to start with the basic momentum theory to calculate the vertical flight performance of the helicopter.

Problems

1. A helicopter rotor in hover encounters aerodynamic forces that oppose its rotation. When a torque is applied from the engines to maintain a rotation at a constant speed, the airframe starts rotating in the opposite direction. Describe methods to counteract this torque and sketch the forces and moments involved in the control of the helicopter.
2. In a conventional helicopter stability control can be achieved with both a pusher or a tractor tail rotor. A pusher tail rotor produces thrust away from the vertical fin; a tractor rotor produces thrust against the fin. Discuss, by further investigations, the merits of each configuration *problem-based learning: additional research is required.*

Aerospace Engineering Desk Reference; ISBN 9781856175753

2.6 V/STOL

Antonio Filippone

An efficient, operational V/STOL aircraft has proven to be an elusive target over the last thirty years.

B. McCormick, 1978

V/STOL is an acronym for Vertical/Short Take-off and Landing Aircraft. This category includes aircraft that are capable of performing most of the flight conditions of fixed-wing aircraft. However, take-off, landing and ground performance can be considerably different from CTOL aircraft. Other peculiar problems of V/STOL vehicles are lift augmentation in the terminal phases, transition from vertical to forward flight (and vice versa), operation at high angles of attack, and engine installation. There are, of course, additional performance requirements on aerodynamics, propulsion, structures and flight control systems.

V/STOL aircraft make up the most exotic and unfortunate category of flight vehicle, including several aircraft that never took off, many that flew only experimental tests, and a handful that have entered production. According to McCormick the history of V/STOL aircraft is full of *disillusionment*.

This chapter will only focus on those peculiar performance indices of real-life V/STOL aircraft, and a few essential concepts for the future. It will not be possible to review the various design options in this class of vehicle. Essential concepts in V/STOL performance are available in AGARD R-710. In addition, McCormick published an interesting short history of V/STOL flight research, and a book on the aerodynamics of V/STOL. Hirschberg published another historical account of V/STOL research. Maisel and co-workers reviewed the research on V/STOL leading to the experimental tilt-rotor XV-15 (precursor of the modern Boeing-Sikorsky V-22). The analysis of some modern design concepts is discussed by Talbot and co-workers. The effects of small vector/thrust angles and their optimal values on subsonic commercial jets were discussed by Gilyard and Bolonkin.

2.6.1 Hover Characteristics

There are two types of hover used by a V/STOL: by the action of a rotor/propeller, or by the action of a jet. The methods presented for estimating hover performance (thrust, power, end endurance) are essentially the same. Transition to airplane mode is different and requires unsteady aerodynamic analysis.

The second case, pertinent to some V/STOL aircraft, is hover capability by means of one or more jets (for example, BAe Harrier, Lockheed Joint Strike Fighter (JSF), and a number of experimental airplanes). This capability, and the related stability and control issues, are strongly dependent on the ground effect. The technical means for producing these jets and the aerodynamic problems associated with propulsion are the subject of V/STOL aerodynamics and aircraft design.

For an aircraft out of ground, lift is created by a free jet flow into stationary surroundings. A single jet flow spreads out according to physical mechanisms that depend on the Reynolds number and on the turbulence entrainment. Substantial research exists in this area. The reader can start with Donaldson and Snedeker and Krotha-palli and co-workers, where some fundamental experiments and correlations are shown. Various semi-empirical formulas exist for estimating the jet development (diameter, average speed, turbulent development).

From our point of view, the jet flow must generate enough thrust to overcome the aircraft's weight. The jet must produce a thrust equivalent to the weight to maintain a stable hover condition. With a one-dimensional analysis, if U_j is the average speed of the jet at the engine nozzle and $\dot{m}_j$ is the jet mass flow rate, then the resulting thrust equation must be

$$T = W \simeq (\dot{m}U)_{jet} = 2(\rho A U^2)_{jet}. \tag{2.6.1}$$

The thrust required is generated by a suitable combination of high mass flow rate and high jet speed. The amounts required can be easily calculated. For a $GTOW = 10{,}000$ kg, and a jet Mach number limited to $M = 1.5$ at sea level, the mass flow requirements are about 195 kg/s, or 160 m^3/s of air at atmospheric conditions. This mass flow has to be drawn from the surroundings of the aircraft, as shown in Figure 2.6-1. The corresponding disk loading (defined as the weight of the aircraft divided by the total area of the nozzles) is at least two orders of magnitude higher than the conventional helicopters. A maximum disk loading of about 22,000 kg/m^2 is estimated for the BAe Sea Harrier.

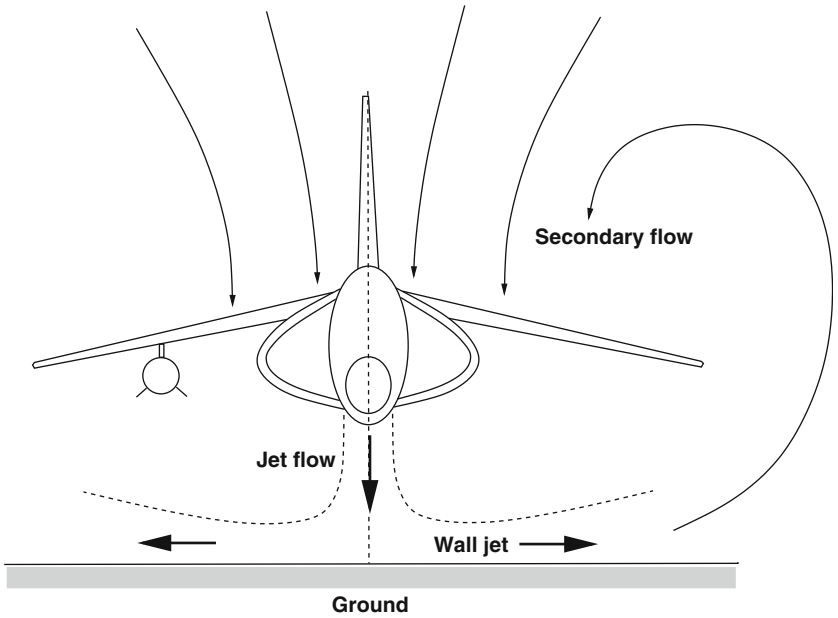


Figure 2.6-1 Single-jet V/STOL aircraft hover in ground effect.

Another way to look at direct lift is to consider the thrust requirements on the engine. Clearly, an OGE hover requires $T/W > 1$, albeit not much higher than 1. This is called *thrust margin.* By comparison, a CTOL jet aircraft can operate with $T/W = 0.20$.

The jet speed has to be limited for several reasons, including noise, vibrations, and heat fatigue. Equation 2.6.1 gives only a rough estimate of the thrust generated by a vertical jet. Thrust requirements for hover are in excess of the aircraft gross weight, due to losses in lift attributable to a number of factors. The most important are: (1) jet-induced losses; (2) hot gas reingestion; and (3) margins required for maneuver and acceleration. These losses depend on the relative ground clearance, and are greater in proximity to the ground. The US Military standards require a minimum $0.1g$ vertical acceleration capability with a $T/W = 1.05$. Therefore, the installed thrust must be 5% in excess of the gross weight. In this context, *installed thrust* is the difference between the gross thrust and the jet-induced losses.

2.6.2 Jet-Induced Lift

The effects created by impinging jets on V/STOL vehicle performance are significant, and potentially dangerous if not properly understood. Skifstad reviewed the physics of the jets for V/STOL applications and highlighted several aerodynamic aspects, including the induced aerodynamic field, the free jet, the impinging jet and the multiple jets. Consider first the case of a single-jet V/STOL aircraft. The jet-induced losses are due to a downwash created by the jet, because this draws air from the surroundings. This field of velocities creates a further downward push, particularly around the wings, which are most exposed. Jet nozzles under the wings compound this problem.

The presence of the ground changes the characteristics of the flow, because impingement is followed by lateral spreading of the jet, rising of hot gases, and return of the gases back onto the airplane. One such case is sketched in Figure 2.6-1. Depending on various geometrical conditions between the aircraft and the ground, three main cases are identified:

1. The jet spreads radially as a thin sheet with rapidly decaying speed; the reduced pressure creates a suck-down effect. Clearly, this is to be avoided. Calculation of the spreading depends on a number of factors, including the aircraft configuration, the geometry of the jet, and the position of the aircraft above the ground.
2. The jet spreads radially and creates a peripheric vortex ring. The vortex ring augments the jet lift. This is the *hovercraft* principle.
3. The jet spreads out, but it creates an upward radial flow bouncing off the ground (fountain effect), which may either be favorable (cushion) or adverse (suck-down). The conditions under which these events occur depend

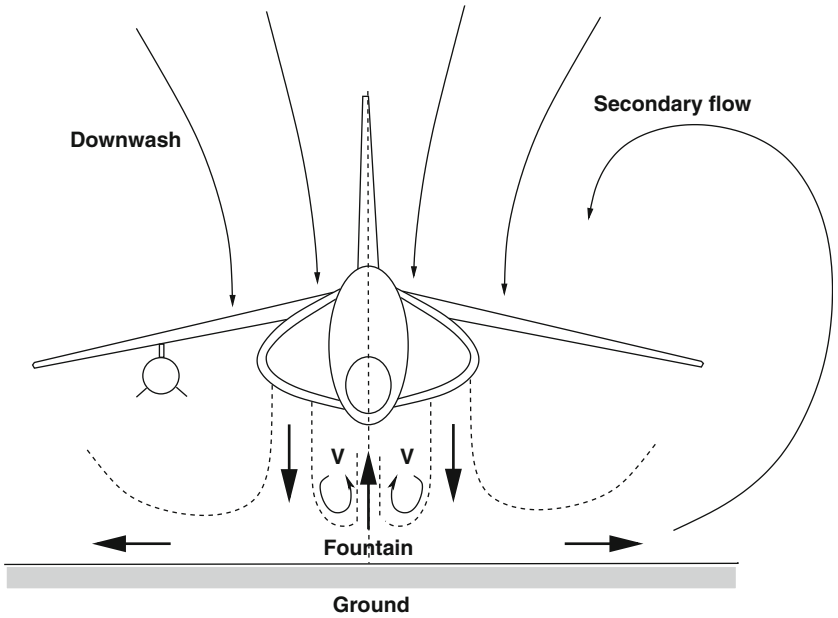


Figure 2.6-2 Twin-jet V/STOL aircraft hover in ground effect showing the fountain effect.

on the geometry of the aircraft. They can be controlled by using strakes (BAe Sea Harrier) at appropriate positions. The cushion effect cannot be used to reduce the engine thrust installed, because the thrust margin must always be based on the worst possible scenario.

A strong suck-down effect is likely to occur in the single-jet V/STOL, while a twin jet, with closely spaced jets (Figure 2.6-2), reduces this effect to a few percent of the total thrust. Properly configured twin jets produce favorable fountain effects that almost cancel the suck-down. A detailed investigation on side-by-side jets has been provided by Louisse and Marshall.

Hot-gas reingestion arises from the fact that exhaust gases bouncing off the ground and spreading around the aircraft mix with the surrounding area thanks to the buoyancy effect. They increase the local temperatures (by a few degrees) and may return to the engine's inlets reducing the engine efficiency. This leads to a thrust loss of about 4% on the BAe Sea Harrier. Another phenomenon is the hot-gas fountain, similar to the one described, in which the mixing of hot gases with the surrounding air is more limited; the engines may breathe more polluted air, with temperatures considerably higher than the ambient temperature. Thrust losses in this case can be as much as 8%. A typical example of jet lift losses due to the various causes is shown in Figure 2.6-3, elaborated and adapted from a number of publications (AGARD CP-242, AGARD R-710).

Two cases are shown in Figure 2.6-3: a case of *clean* aircraft: and an aircraft with strakes. The use of strakes below the aircraft, at appropriate locations, can inhibit some of the processes discussed. The strakes are capable of limiting the overall lift loss to about *2%* of the engine thrust. For lower hover positions of the aircraft, there are jet flow instabilities that are responsible for the non-linear effects indicated by the lift curves. The maximum losses are limited to 4% or 5%, and are gradually reduced to about 1.5% in OGE operation. Therefore, a thrust margin of 5% is required, as indicated earlier.

2.6.2.1 Estimation of Jet-Induced Fountain Lift and Suck-Down

According to Bellavia and co-workers, the total loss of lift, relative to the net thrust, can be estimated from an expression like

$$\frac{\Delta L}{T} = \frac{\Delta L_{OGE}}{T} + \frac{\Delta L_f}{T} + \frac{\Delta L_{sf}}{T} + \frac{\Delta L_{sr}}{T}, \tag{2.6.2}$$

where ΔL_{OGE} is the loss of lift out-of-ground effect (free air), ΔL_f is the fountain lift, and ΔL_{sf} and ΔL_{sr} are the forward and rear suck-down lift, respectively. The expressions for these losses depend on a number of semi-empirical parameters, in addition to several specific parameters, such as the jet diameters, the distance between jets, and the

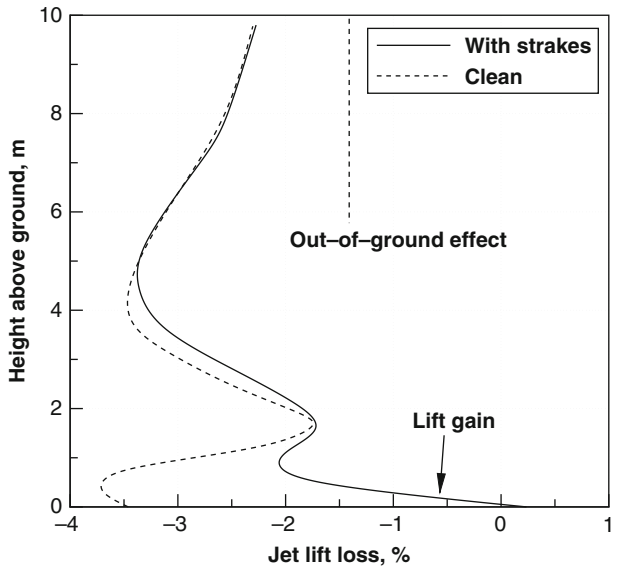


Figure 2.6-3 Jet lift loss due to ground interference, for two aircraft configurations.

planform area. Various models exist for the evaluation of the terms in Eq. 2.6.2, as summarized by Walters and Henderson.

For an aircraft with two jets, the loss of OGE lift can be written as

$$\frac{\Delta L}{T} = k\sqrt{\frac{A_p}{A_j}}\left(\frac{2\pi d}{d_e}\right)^{1.58} p_r^{-1/2}, \tag{2.6.3}$$

where A_p is the planform area of the aircraft, A_j is the cross-sectional area of the jet at the exit of the nozzle, d is the jet diameter, d_e the equivalent diameter, p_r is the pressure ratio of the nozzle; finally, $k = 10^{-4}$ is a constant. Plotting of this function reveals that the OGE jet-induced losses are in the range of 0.5% for realistic values of the nozzle pressure ratio ($p_r = 2$ to 6).

2.6.3 Lift Augmentation

Circulation control technology is a means of increasing the lift without flaps and slats, and is particularly useful for V/STOL operations of conventional aircraft. One such control is a jet flap. This is a stream of air injected near the trailing edge to energize weak boundary layers prone to separation. Basically, the jet flaps simulate the effect of the flap, without the external mechanisms of the flaps.

The use of a conventional jet diverted downward is an intermediate case between pure jet lift and propeller lift. It is aimed at improving the ground performance of the aircraft in conjunction with aerodynamic lift. The methods giving a lift augmentation by diverting the jet are called *vectored-thrust* systems. In Chapter 2.3 we briefly discussed a number of high-lift systems, which are used for augmenting the lift during the terminal phases and maneuvers of the aircraft. These systems, and other more complex powered systems, can be an integral part of the aircraft, and provide better performance near the ground.

Lift augmentation by deviation of the engine thrust is a common means of increasing performance of V/STOL vehicles. Upper-Surface Blowing (USB) uses the concept of lift augmentation by reduction of the upper side pressure. This is achieved by diverting the hot gas from the engine nozzle over the wing, see Figure 2.6-4. One such application is identified on the STOL vehicles Antonov An-72 and An-74, and on some experimental aircraft, such as the Boeing YC-14 and NASA's QSRA. These propulsion systems allow the aircraft to operate at double the C_L of that of CTOL aircraft (see also Figure 2.3-7).

The Externally Blown Flap (EBF) uses the concept of diverting the jet from the lower side of the wing. In practice, this is done by using the jet in coordination with a multi-element wing system, which serves to divert the jet, see Figure 2.6-5. One such application is identified in the Lockheed YC-15 and Lockheed C-17 military transport aircraft.

Both concepts allow slow and precise landings on short runways and in extreme weather, such as in Arctic conditions. Additional benefits of the vectored thrust can be achieved in flight conditions other than take-off and landing. For

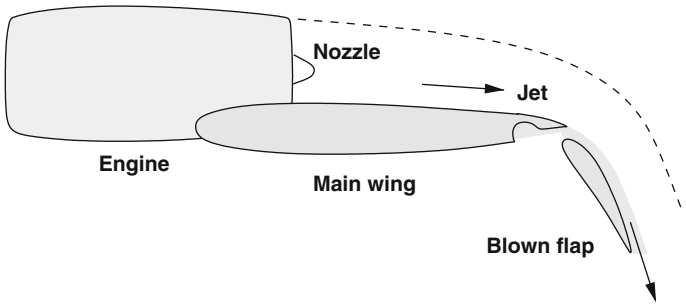


Figure 2.6-4 Sketch of the USB concept.

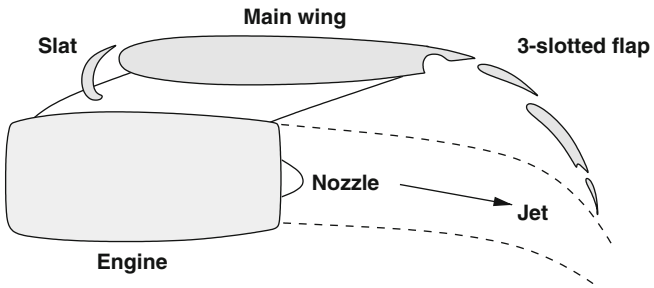


Figure 2.6-5 Sketch of the EBF concept.

example, a vectored nozzle can increase the turn rate and decrease the turn radius. Thus, the aircraft is capable of performing unrestricted turns at high speed (see also Problem 3).

In air combat situations, vectored nozzles can increase the deceleration rates, while applying extra lift. The additional deceleration is due to the forward component of the gross thrust. V/STOL aircraft such as the BAe Harrier can add up to 0.75 g in deceleration rates using this technique.

2.6.4 Calculation of Short Take-Off

The use of powered lift systems that provide thrust at an angle ϵ with the aircraft longitudinal axis leads to a change in the flight equations. We must note that only the gross thrust can be vectored, while the intake momentum drag operates in the direction of the flight path. With this caution, we use the symbol T for gross thrust, instead of net thrust, as we have done earlier; T_o denotes the thrust that is actually vectored or diverted. The momentum equation along the horizontal runway is

$$m\frac{\partial U}{\partial t} = T\cos\epsilon - D - R. \tag{2.6.4}$$

The lift is augmented by the vector thrust $T_o \sin\epsilon$; therefore, the sum of all forces in the vertical direction is

$$W - L - T_o \sin\epsilon, \tag{2.6.5}$$

and the rolling resistance is

$$R = \mu(W - L - T_o \sin\epsilon). \tag{2.6.6}$$

Solution of this equation is carried out in the same way as the CTOL, after noting that there is a constant factor $\cos\epsilon$ on the thrust term. The lift-off run equation can be written as

$$\frac{1}{2}U_{lo}^2 = \frac{1}{m}\int_o^x [T_o\cos\epsilon - D - \mu\,(W - T_o\sin\epsilon - L)]\,dx \tag{2.6.7}$$

or

$$\frac{1}{2}U_{lo}^2 = \frac{1}{m}\int_o^x [T_o\cos\epsilon - \mu\,(W - T_o\sin\epsilon) - D - \mu L)]\,dx \tag{2.6.8}$$

We can consider the thrust as a constant, and equal to the static thrust; therefore, we collect all the known terms in the coefficient c_1,

$$c_1 = T_o \cos\epsilon - \mu(W - T_o \sin\epsilon). \tag{2.6.9}$$

By replacing c_1 and the definition of C_D and C_L in Eq. 2.6.7, we find

$$\frac{1}{2} m U_{lo}^2 = \int_o^x \left[c_1 - \frac{1}{2}\rho A U^2 (C_D - \mu C_L) \right] dx. \tag{2.6.10}$$

Now we need a few approximations. The first is that the sum of the aerodynamic terms in the integrand function be a constant. In reality, the C_L is a combination of aerodynamics and propulsion characteristics of the aircraft, and cannot be calculated exactly without engine considerations.

The second approximation is that the acceleration is also constant. An estimate for the lift-off run of the STOL aircraft is

$$\frac{1}{2} m U_{lo}^2 = c_1 x_{lo} - \frac{1}{2}\rho A (C_D - \mu C_L) \frac{1}{2} U_{lo}^2 x_{lo}, \tag{2.6.11}$$

$$x_{lo} = \frac{m U_{lo}^2 / 2}{T_o \cos\epsilon - \mu(W - T_o \sin\epsilon) - \rho A (C_D - \mu C_L) U_{lo}^2 / 4}. \tag{2.6.12}$$

Equation 2.6.12 can be manipulated so as to eliminate the lift-off speed, and to show the thrust ratio *T/W*. The take-off run requires the calculation of the other phases of flight: flare, rotation and climb to clear a screen at 35 or 50 ft. If we neglect in first approximation the flare and rotation, we find the take-off run by summing Eq. 2.6.12 to the airborne distance

$$x = \frac{h}{\tan\gamma}, \quad \tan\gamma = \frac{T - D}{W}. \tag{2.6.13}$$

An analysis of interest is the comparison between the lift-off run with and without vectored thrust. This is shown in Figure 2.6-6, which is relative to the supersonic jet aircraft. We have assumed engine thrust and drag characteristics not dependent on the speed, and equal to those at $M = 0$. The other conditions are sea level, take-off speed $U_{to} = 300$ km/h (162 kt), and dry runway conditions at sea level.

The reduction of take-off field is about 25% with a 25 degree thrust angle. If an analysis is carried out into the weight effects on the take-off run, then we see a considerable difference in x_o. This is shown in Figure 2.6-7.

In the past, some aircraft were designed for very high lift. Notable in this field is the NASA QSRA, which was capable of achieving lift coefficients as high as 10, as shown in Figure 2.6-8, thanks to its USB flaps.

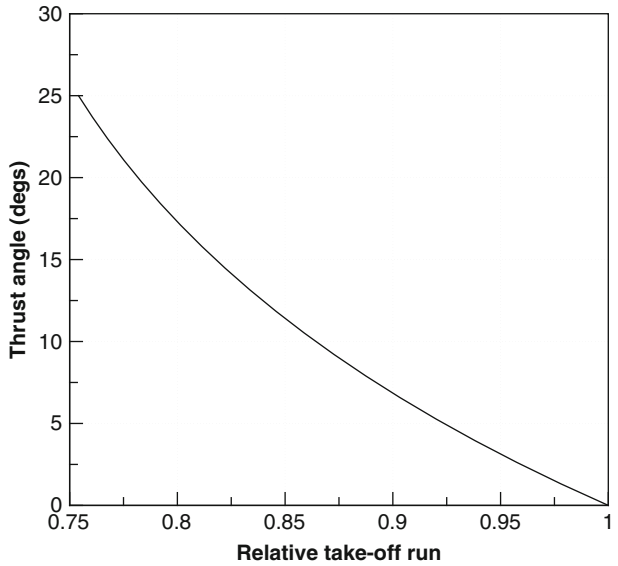


Figure 2.6-6 Relative take-off run for jet aircraft model C with vectored thrust.

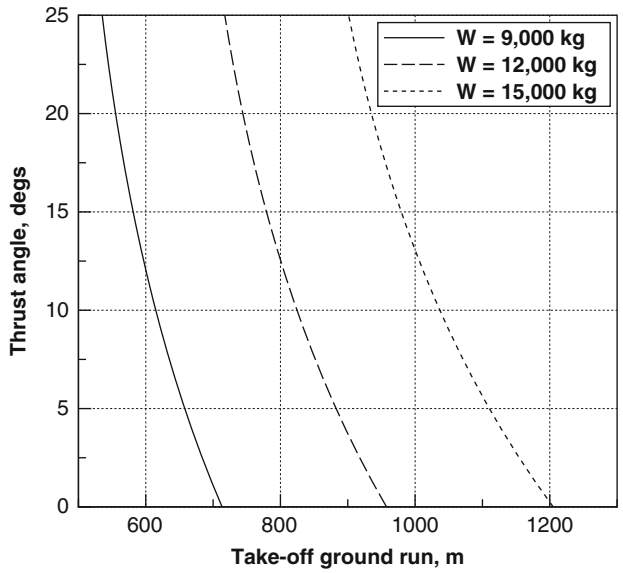


Figure 2.6-7 Take-off run for jet aircraft model C with vectored thrust, military thrust, sea level ISA conditions.

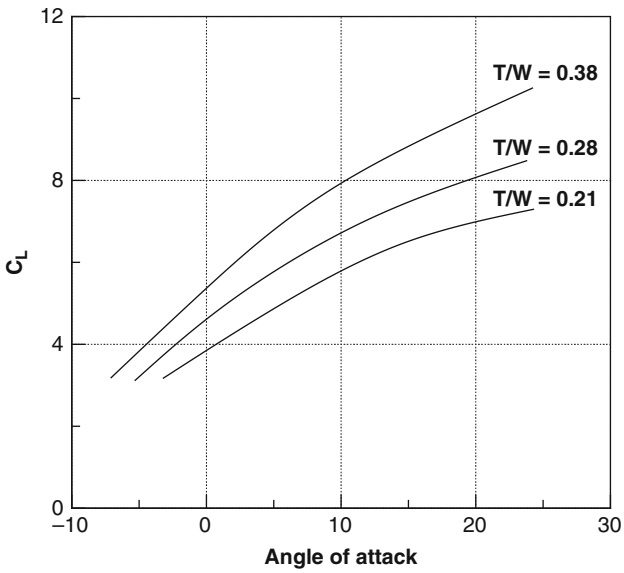


Figure 2.6-8 Estimated vectored thrust lift coefficients for the QSRA at three thrust ratios.

However, it is not true that high lift is always beneficial. Depending on the thrust ratio, the take-off run can in fact increase instead of decreasing. The take-off run is found from the sum of Eq. 2.6.12 and Eq. 2.6.13. The second segment (airborne distance) may offset the benefits of high C_L and vectored thrust. This is shown in Figure 2.6-9, which resulted from the solution of Eq. 2.6.12 and Eq. 2.6.13 with a fixed vector thrust angle $\epsilon = 10$ degrees.

2.6.5 Ski Jump

For aircraft based on a carrier ship there are several constraint factors on take-off and landing, because the shipboard flight deck is limited. In general, there are additional accelerating and decelerating gear, such as catapults, wires, and parachutes. However, the performance of these aircraft is clearly influenced by the wind conditions, and their performance is usually termed *wind-over-deck*. One type of take-off sometimes used by high-performance aircraft from an aircraft carrier is the *ski-jump*, as sketched in Figure 2.6-10. This take-off can be a catapult or non-catapult jump from an inclined deck.

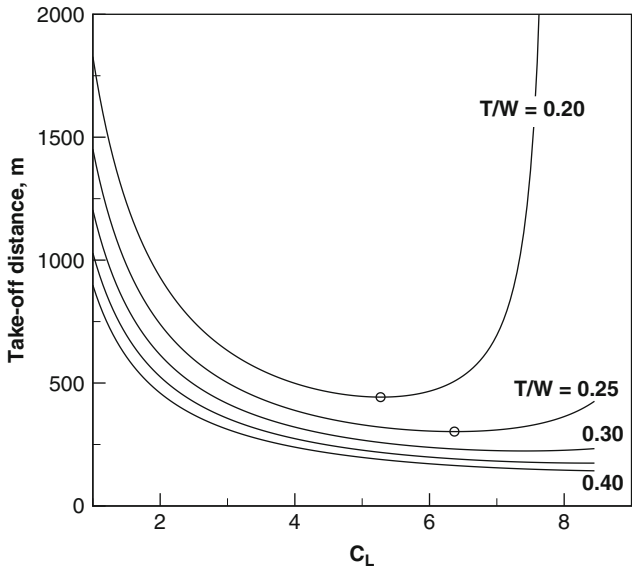


Figure 2.6-9 Effects of C_L on STOL aircraft take-off run, at selected thrust ratios; $\epsilon = 10$ degrees.

Clarke and Walters have simulated the ski jump of the Grumman F-14A and the McDonnell–Douglas F-18A with a three-degree-of-freedom model, and compared their results with some ramp-assisted take-offs. They included non-linear models for the thrust, for the landing gear loading (spring/strut system), and control systems dynamics. The maneuver relies on the fact that the aircraft is capable of gaining lift as soon as it leaves the deck, so that the maximum height drop from the deck is above sea level in the (statistically) worst ocean waves. The variation of height drop is considerably dependent on the launch speed. For the BAe Sea Harrier the maximum height drop is about 3 m once in 3.6 million take-off operations – by all means a safe operation. However, the safe limits of a ski jump are set by the minimum lift-off speed that provides zero climb rate.

The basic equations of the flight trajectory are

$$\dot{x} = U\cos\gamma, \tag{2.6.14}$$

$$v_c = \dot{h} = U\sin\gamma, \tag{2.6.15}$$

$$\dot{U} = \frac{1}{m}(T\cos\epsilon - D) - g\sin\gamma, \tag{2.6.16}$$

$$U\dot{\gamma} = \frac{1}{m}(T\sin\epsilon + L) - g\cos\gamma, \tag{2.6.17}$$

where γ is the climb angle and ϵ is the angle of thrust calculated with respect to the longitudinal axis of the aircraft. These equations are valid from lift-off. The acceleration of the aircraft from brake release to lift-off is given by equations similar to the ones discussed in § 2.6.4. However, There are some changes to be considered: (1) the

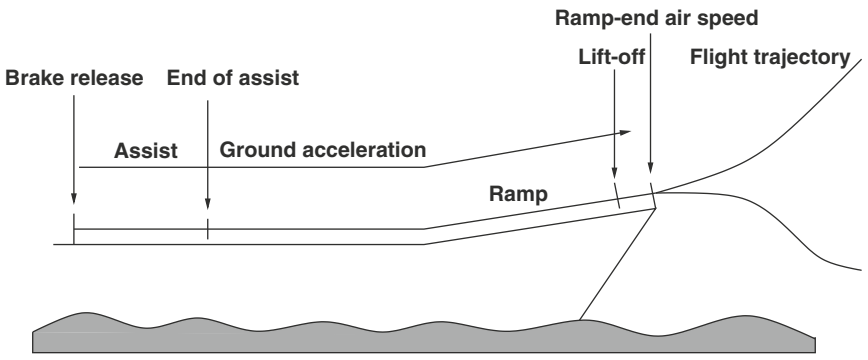


Figure 2.6-10 Ski jump from aircraft carrier and height drop.

presence of an inclined ramp; and (2) the acceleration provided by the catapult. The catapult has a nearly impulsive effect, and at the point of release from the aircraft there will be a rapid change of acceleration.

The ski jump simulation can be split in two parts: the ground run till the lift-off point and the initial airborne phase. Let us consider first the effect of the ramp angle γ_o on the take-off run, assuming that there is no assist to the acceleration. The momentum equation is written along the ramp direction, inclined by γ_o which for this purpose is assumed to be constant. The ramp adds a constant term in Eq. 2.6.4,

$$m\frac{\partial U}{\partial t} = T\cos\epsilon - D - R - W\sin\gamma_o, \tag{2.6.18}$$

where the rolling resistance is

$$R = \mu(W\cos\gamma_o - L - T\sin\epsilon). \tag{2.6.19}$$

We assume again that the thrust is not dependent on the speed, and is equal to the static thrust. Solution of Eq. 2.6.19, according to the methods previously described, is the following:

$$\frac{1}{2}U_{lo}^2 = \frac{1}{m}\int_o^x \left[T_o\cos\epsilon - \frac{1}{2}\rho A(C_{D_o} + kC_L^2)U^2 - \mu(W\cos\gamma_o - T_o\sin\epsilon) - \frac{1}{2}\mu\rho AC_LU^2 - W\sin\gamma_o\right]dx. \tag{2.6.20}$$

A number of constant parameters can be identified. We call

$$c_1 = \mu(W\cos\gamma_o - T_o\sin\epsilon) - W\sin\gamma_o, \tag{2.6.21}$$

so that Eq. 2.6.21 becomes

$$\frac{1}{2}U_{lo}^2 = \frac{1}{m}\int_o^x \left[c_1 - \frac{1}{2}\rho A(C_{D_o} + kC_L^2 + \mu C_L)U^2\right]dx. \tag{2.6.22}$$

If the C_L is constant, we also have

$$c_2 = -\frac{1}{2}\rho A(C_{D_o} + kC_L^2 + \mu C_L), \tag{2.6.23}$$

and the equation is reduced to

$$\frac{1}{2}U_{lo}^2 = \frac{1}{m}\int_o^x (c_1 + c_2U^2)dx, \tag{2.6.24}$$

which has a well-known form. Unfortunately, the C_L is not constant, and the integral must be solved numerically. We can now take into account the actual variation of thrust, lift coefficient, and ramp angle by a numerical solution of Eq. 2.6.21:

$$\frac{1}{2}U_{lo}^2 = \frac{1}{m}\sum_{i=1}^{n}(c_1 + c_2U^2)_i\Delta x_i, \tag{2.6.25}$$

with

$$c_{1_i} = \mu(W\cos\gamma_o - T_o\sin\epsilon)_i, \qquad c_{2_i} = \frac{1}{2}\rho A(C_{D_o} + kC_L^2 + \mu C_L)_i. \tag{2.6.26}$$

If there is no assist, the equation can be solved with a constant acceleration. The calculation of the airborne phase requires the integration of the differential equations, Eqs 2.6.14–2.6.17, with appropriate initial conditions. These conditions can be found from the calculation of the take-off run, Eq. 2.6.21, or with specified starting conditions. Figure 2.6-11 shows the effects of GTOW on the flight path, and indicates that if the aircraft is too heavily loaded there is the risk of crash landing in the ocean. For all the cases shown in Figure 2.6-11 the starting conditions are: Mach number at take-off $M = 0.22$; lift coefficient $C_L = 0.8$; ramp angle $\gamma_o = 9$ degrees; ramp elevation $h = 14$ m from nominal water level; thrust angle $\epsilon = 4$ degrees. Similar trade-off studies can be done in terms of ramp angle, lift coefficient (use of high-lift devices, vectored thrust), and take-off speed.

2.6.6 Convertiplanes or Tilt Rotors

A convertiplane is an aircraft that can change the direction of the thrust from vertical to horizontal and vice versa. Basically, this can be done by tilting the rotors/propellers. A considerable amount of literature is available on tilt-rotor concepts. Some basic performance analysis has been provided by Hafner.

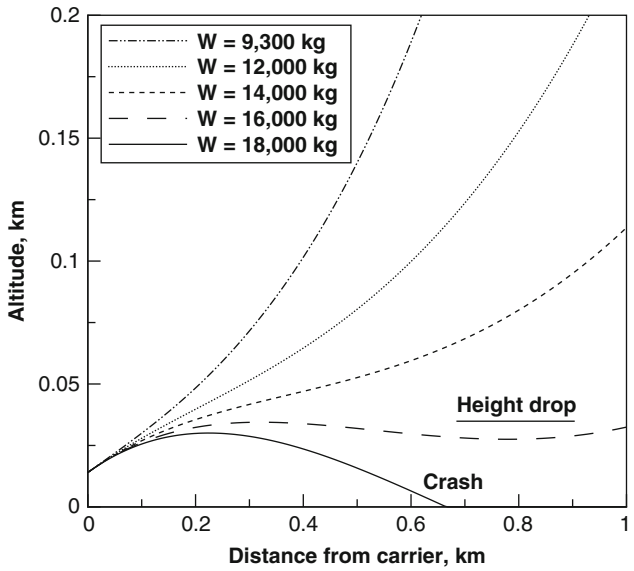


Figure 2.6-11 Weight effect on ski jump from aircraft carrier. Summary of calculated flight paths.

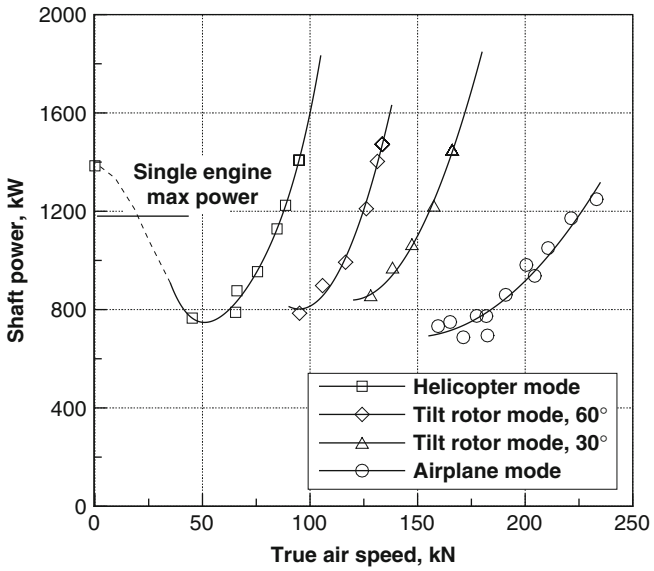


Figure 2.6-12 Forward flight power of the XV-15 tilt rotor.

Since the requirements for hover and cruise are substantially different, the optimal configuration of the aircraft is a compromise. Disk loadings are higher than conventional helicopters, and can exceed 100 kgf/m^2 (10^3 Pa). Figure 2.6-12 shows the power requirements for level flight of the experimental tilt-rotor XV-15 from flight testing. One curve shows the power curve in helicopter mode. At the other end of the chart there is the power required for level flight in airplane mode. Between these two operating modes there is a set of curves obtained with flight at intermediate tilt angles. A proper analysis of the conversion from helicopter to airplane mode requires an unsteady aerodynamics calculation.

Operation in helicopter mode is possible only at low forward speeds; in the airplane mode low speeds are not possible, due to stall problems on the lifting surfaces. Therefore, the aircraft has the possibility of extending its flight envelope, to include the characteristics of the helicopter and the propeller-driven airplane. The Boeing-Sikorsky V-22 and the Agusta-Bell BA609 work on the same principle.

The tilt rotor has a peculiar flight envelope relative to the angle of thrust. In fact, there is a restricted range of thrust angles at any given speed. Safe operation outside these boundaries is not possible. Height/velocity diagrams for

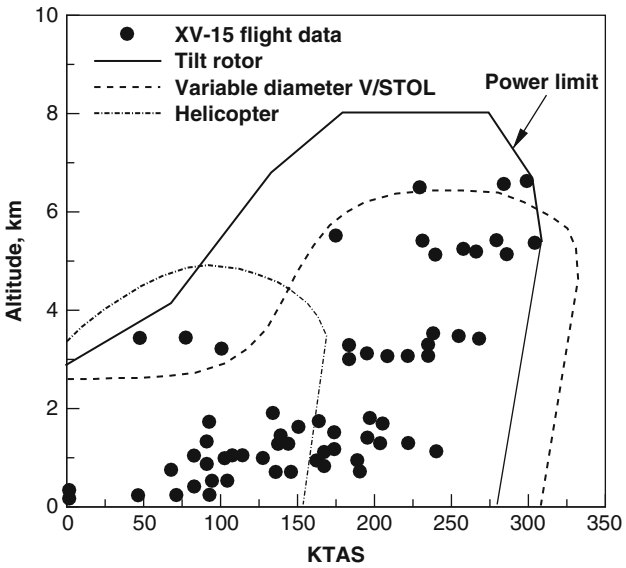


Figure 2.6-13 Estimated V/STOL flight envelopes for conceptual tilt-rotor and variable-diameter rotor aircraft.

the XV-15 in OEI operation near the ground, and optimal OEI take-off flight paths using optimal control theory, have been discussed by Carlson and Zhao.

2.6.7 V/STOL Flight Envelopes

The flight envelopes of V/STOL aircraft are different than the envelopes shown so far. First, V/STOL aircraft can maintain their altitude at zero ground speed by hovering. This is achieved either by jet lift or by rotor propulsion. Second, the maximum speeds obtainable are somewhat lower. Flight envelopes of various V/STOL concepts (tilt wing, variable diameter tilt rotor, shrouded rotor) are discussed by Scott.

Figure 2.6-13 shows the flight envelopes of three V/STOL vehicles. This figure also shows some of the limitations on the flight envelopes. The combination of vertical take-off/landing capability, in addition to conventional cruise performance, extends the flight envelopes and has the potential to increase the aircraft's productivity.

Transition for vertical lift to aerodynamic lift requires a smooth change of the thrust vector, so that ultimately the vectored portion of the thrust is limited or negligible. Thus, the aircraft starts the cruise with aerodynamic lift only. Experience has indicated that this is an extremely complicated flight condition, with flow instabilities difficult to control. For a jet-lift V/STOL aircraft in transition, there must be enough installed thrust to overcome the gross weight and the resistance to forward speed, at least to a point when enough wing lift has been developed.

Problems

1. Study the effects of runway conditions on the take-off run of a hypothetical vector thrust on the supersonic jet fighter of reference. Consider sea-level ISA conditions, and two take-off weights: MTOW and the combat weight.
2. For the same aircraft as in the previous problem, and the same take-off conditions, study the effect of after-burning thrust at take-off.
3. Consider a generic V/STOL aircraft with vectored thrust. The nozzle is capable of relatively large deflections. Assume that the aircraft is to perform a turn, with a bank angle ϕ. Make a sketch of the forces involved in the frontal plane y_b, z_b (body axes) and write the momentum equations for the aircraft in such a situation. Solve and discuss how the turn performance is affected if the thrust is deflected by an angle ϵ on the longitudinal axis. (If necessary, do additional research to find suitable papers on the subject.)
4. Consider a V/STOL aircraft, capable of hover and climb with pure jet lift. The aircraft is powered by a single turbofan engine that delivers a maximum static thrust at sea level $T = 95$ kN. The engine has four nozzles that can be vectored by up to 110 degrees. Each nozzle has a diameter of 0.38 m. The air flow at maximum thrust is 78 kgf/s, is at an average temperature $\theta = 900$ K, and roughly atmospheric pressure. The thrust-specific fuel

consumption is 0.344 N/h/N. Calculate the disk loading of the aircraft, the average speed of the jet, and the fuel required to hover for 1 minute.

5. The combination of the fixed-wing and the rotary-wing flight envelopes (Figure 2.6-13) show a gap in the *h*/*V* chart where flight is not possible. In particular, low-speed flight (less than 100 kt) is not achieved by any of the aircraft discussed at the troposphere. Flight in this region could allow detailed measurements of the atmosphere (in particular, humidity, concentration of pollutants, wind shear, etc.). Discuss critically the technical means for extending the flight envelope in this area.

This page intentionally left blank

SECTION 3

Aircraft Structures

This page intentionally left blank

Aerospace Engineering Desk Reference; ISBN 9781856175753

3.1 Materials

T.H.G. Megson

With the present chapter we begin the purely aeronautical section of the book, where we consider structures peculiar to the field of aeronautical engineering. These structures are typified by arrangements of thin, load-bearing skins, frames and stiffeners, fabricated from lightweight, high strength materials of which aluminium alloys are the most widely used examples.

As a preliminary to the analysis of the basic aircraft structural forms presented in subsequent chapters we shall discuss the materials used in aircraft construction.

Several factors influence the selection of the structural material for an aircraft, but amongst these strength allied to lightness is probably the most important. Other properties having varying, though sometimes critical significance are stiffness, toughness, resistance to corrosion, fatigue and the effects of environmental heating, ease of fabrication, availability and consistency of supply and, not least important, cost.

The main groups of materials used in aircraft construction have been wood, steel, aluminium alloys with, more recently, titanium alloys, and fibre-reinforced composites. In the field of engine design, titanium alloys are used in the early stages of a compressor while nickel-based alloys or steels are used for the hotter later stages. As we are concerned primarily with the materials involved in the construction of the airframe, discussion of materials used in engine manufacture falls outside the scope of this book.

3.1.1 Aluminium Alloys

Pure aluminium is a relatively low strength extremely flexible metal with virtually no structural applications. However, when alloyed with other metals its properties are improved significantly. Three groups of aluminium alloy have been used in the aircraft industry for many years and still play a major role in aircraft construction. In the first of these aluminium is alloyed with copper, magnesium, manganese, silicon and iron, and has a typical composition of 4% copper, 0.5% magnesium, 0.5% manganese, 0.3% silicon and 0.2% iron with the remainder being aluminium. In the wrought, heat-treated, naturally aged condition this alloy possesses a 0.1% proof stress not less than 230 N/mm^2, a tensile strength not less than 390 N/mm^2 and an elongation at fracture of 15%. Artificial ageing at a raised temperature of, for example, 170°C increases the proof stress to not less than 370 N/mm^2 and the tensile strength to not less than 460 N/mm^2 with an elongation of 8%.

The second group of alloys contain, in addition to the above, 1–2% of nickel, a higher content of magnesium and possible variations in the amounts of copper, silicon and iron. The most important property of these alloys is their retention of strength at high temperatures which makes them particularly suitable for aero engine manufacture. A development of these alloys by Rolls-Royce and High Duty Alloys Ltd replaced some of the nickel by iron and reduced the copper content; these RR alloys, as they were called, were used for forgings and extrusions in aero engines and airframes.

The third group of alloys depends upon the inclusion of zinc and magnesium for their high strength and have a typical composition of 2.5% copper, 5% zinc, 3% magnesium and up to 1 % nickel with mechanical properties of 0.1 % proof stress 510 N/mm^2, tensile strength 585 N/mm^2 and an elongation of 8%. In a modern development of this alloy nickel has been eliminated and provision made for the addition of chromium and further amounts of manganese.

Alloys from each of the above groups have been used extensively for airframes, skins and other stressed components, the choice of alloy being influenced by factors such as strength (proof and ultimate stress), ductility, ease of manufacture (e.g. in extrusion and forging), resistance to corrosion and amenability to protective treatment, fatigue strength, freedom from liability to sudden cracking due to internal stresses and resistance to fast crack propagation under load. Clearly, different types of aircraft have differing requirements. A military aircraft, for instance, having a relatively short life measured in hundreds of hours, does not call for the same degree of fatigue and corrosion resistance as a civil aircraft with a required life of 30 000 hours or more.

Unfortunately, as one particular property of aluminium alloys is improved, other desirable properties are sacrificed. For example, the extremely high static strength of the aluminium–zinc–magnesium alloys was accompanied for many years by a sudden liability to crack in an unloaded condition due to the retention of internal stresses in bars, forgings and sheet after heat treatment. Although variations in composition have eliminated this problem to a considerable extent other deficiencies showed themselves. Early post-war passenger aircraft experienced large numbers of

stress-corrosion failures of forgings and extrusions. The problem became so serious that in 1953 it was decided to replace as many aluminium–zinc–manganese components as possible with the aluminium–4 per cent copper alloy L65 and to prohibit the use of forgings in zinc-bearing alloy in all future designs. However, improvements in the stress-corrosion resistance of the aluminium–zinc–magnesium alloys have resulted in recent years from British, American and German research. Both British and American opinions agree on the benefits of including about 1 per cent copper but disagree on the inclusion of chromium and manganese, while in Germany the addition of silver has been found extremely beneficial. Improved control of casting techniques has brought further improvements in resistance to stress corrosion. The development of aluminium–zinc–magnesium–copper alloys has largely met the requirement for aluminium alloys possessing high strength, good fatigue crack growth resistance and adequate toughness. Further development will concentrate on the production of materials possessing higher specific properties, bringing benefits in relation to weight saving rather than increasing strength and stiffness.

The first group of alloys possess a lower static strength than the above zinc-bearing alloys, but are preferred for portions of the structure where fatigue considerations are of primary importance such as the undersurfaces of wings where tensile fatigue loads predominate. Experience has shown that the naturally aged version of these alloys has important advantages over the fully heat-treated forms in fatigue endurance and resistance to crack propagation. Furthermore, the inclusion of a higher percentage of magnesium was found, in America, to produce, in the naturally aged condition, mechanical properties between those of the normal naturally aged and artificially aged alloy. This alloy, designated 2024 (aluminium–copper alloys form the 2000 series), has the nominal composition (4.5 per cent copper, 1.5 per cent magnesium, 0.6 per cent manganese, with the remainder aluminium) and appears to be a satisfactory compromise between the various important, but sometimes conflicting, mechanical properties.

Interest in aluminium–magnesium–silicon alloys has recently increased, although they have been in general use in the aerospace industry for decades. The reasons for this renewed interest are that they are potentially cheaper than aluminium–copper alloys and, being weldable, are capable of reducing manufacturing costs. In addition, variants, such as the ISO 6013 alloy, have improved property levels and, generally, possess a similar high fracture toughness and resistance to crack propagation as the 2000 series alloys.

Frequently, a particular form of an alloy is developed for a particular aircraft. An outstanding example of such a development is the use of Hiduminium RR58 as the basis for the main structural material, designated CM001, for Concorde. Hiduminium RR58 is a complex aluminium–copper–magnesium–nickel–iron alloy developed during the 1939–1945 war specifically for the manufacture of forged components in gas turbine aero engines. The chemical composition of the version used in Concorde was decided on the basis of elevated temperature, creep, fatigue and tensile testing programmes and has the detailed specification of:

	%Cu	%Mg	%Si	%Fe	%Ni	%Ti	%Al
Minimum	2.25	1.35	0.18	0.90	1.0	–	Remainder
Maximum	2.70	1.65	0.25	1.20	1.30	0.20	

Generally, CM001 is found to possess better overall strength/fatigue characteristics over a wide range of temperatures than any of the other possible aluminium alloys.

The latest aluminium alloys to find general use in the aerospace industry are the aluminium–lithium alloys. Of these, the aluminium–lithium–copper–manganese alloy, 8090, developed in the UK, is extensively used in the main fuselage structure of GKN Westland Helicopters' design EH101; it has also been qualified for Eurofighter 2000 (now named the Typhoon) but has yet to be embodied. In the USA the aluminium–lithium–copper alloy, 2095, has been used in the fuselage frames of the F16 as a replacement for 2124, resulting in a fivefold increase in fatigue life and a reduction in weight. Aluminium–lithium alloys can be successfully welded, possess a high fracture toughness and exhibit a high resistance to crack propagation.

3.1.2 Steel

The use of steel for the manufacture of thin-walled, box-section spars in the 1930s has been superseded by the aluminium alloys described in Section 3.1.1. Clearly, its high specific gravity prevents its widespread use in aircraft construction, but it has retained some value as a material for castings for small components demanding high tensile strengths, high stiffness and high resistance to wear. Such components include undercarriage pivot brackets, wing-root attachments, fasteners and tracks.

Although the attainment of high and ultra-high tensile strengths presents no difficulty with steel, it is found that other properties are sacrificed and that it is difficult to manufacture into finished components. To overcome some of

these difficulties, types of steel known as *maraging* steels were developed in 1961, from which carbon is either eliminated entirely or present only in very small amounts. Carbon, while producing the necessary hardening of conventional high tensile steels, causes brittleness and distortion; the latter is not easily rectifiable as machining is difficult and cold forming impracticable. Welded fabrication is also almost impossible or very expensive. The hardening of maraging steels is achieved by the addition of other elements such as nickel, cobalt and molybdenum. A typical maraging steel would have these elements present in the proportions: nickel 17–19 per cent, cobalt 8–9 per cent, molybdenum 3–3.5 per cent, with titanium 0.15–0.25 per cent. The carbon content would be a maximum of 0.03 per cent, with traces of manganese, silicon, sulphur, phosphorus, aluminium, boron, calcium and zirconium. Its 0.2 per cent proof stress would be nominally 1400 N/mm^2 and its modulus of elasticity 180 000 N/mm^2.

The main advantages of maraging steels over conventional low alloy steels are: higher fracture toughness and notched strength, simpler heat treatment, much lower volume change and distortion during hardening, very much simpler to weld, easier to machine and better resistance to stress corrosion/hydrogen embrittlement. On the other hand, the material cost of maraging steels is three or more times greater than the cost of conventional steels, although this may be more than offset by the increased cost of fabricating a complex component from the latter steel.

Maraging steels have been used in: aircraft arrester hooks, rocket motor cases, helicopter undercarriages, gears, ejector seats and various structural forgings.

In addition to the above, steel in its stainless form has found applications primarily in the construction of super- and hypersonic experimental and research aircraft, where temperature effects are considerable. Stainless steel formed the primary structural material in the Bristol 188, built to investigate kinetic heating effects, and also in the American rocket aircraft, the X-15, capable of speeds of the order of Mach 5–6.

3.1.3 Titanium

The use of titanium alloys increased significantly in the 1980s, particularly in the construction of combat aircraft as opposed to transport aircraft. This increase continued in the 1990s to the stage where, for combat aircraft, the percentage of titanium alloy as a fraction of structural weight is of the same order as that of aluminium alloy. Titanium alloys possess high specific properties, have a good fatigue strength/tensile strength ratio with a distinct fatigue limit, and some retain considerable strength at temperatures up to 400–500°C. Generally, there is also a good resistance to corrosion and corrosion fatigue although properties are adversely affected by exposure to temperature and stress in a salt environment. The latter poses particular problems in the engines of carrier-operated aircraft. Further disadvantages are a relatively high density so that weight penalties are imposed if the alloy is extensively used, coupled with high primary and high fabrication costs, approximately seven times those of aluminium and steel.

In spite of this, titanium alloys were used in the airframe and engines of Concorde, while the Tornado wing carry-through box is fabricated from a weldable medium strength titanium alloy. Titanium alloys are also used extensively in the F15 and F22 American fighter aircraft and are incorporated in the tail assembly of the Boeing 777 civil airliner. Other uses include forged components such as flap and slat tracks and undercarriage parts.

New fabrication processes (e.g. superplastic forming combined with diffusion bonding) enable large and complex components to be produced, resulting in a reduction in production man-hours and weight. Typical savings are 30 per cent in man-hours, 30 per cent in weight and 50 per cent in cost compared with conventional riveted titanium structures. It is predicted that the number of titanium components fabricated in this way for aircraft will increase significantly and include items such as access doors, sheet for areas of hot gas impingement, etc.

3.1.4 Plastics

Plain plastic materials have specific gravities of approximately unity and are therefore considerably heavier than wood although of comparable strength. On the other hand, their specific gravities are less than half those of the aluminium alloys so that they find uses as windows or lightly stressed parts whose dimensions are established by handling requirements rather than strength. They are also particularly useful as electrical insulators and as energy absorbing shields for delicate instrumentation and even structures where severe vibration, such as in a rocket or space shuttle launch, occurs.

3.1.5 Glass

The majority of modern aircraft have cabins pressurized for flight at high altitudes. Windscreens and windows are therefore subjected to loads normal to their midplanes. Glass is frequently the material employed for this purpose in the form of plain or laminated plate or heat-strengthened plate. The types of plate glass used in aircraft have

a modulus of elasticity between 70 000 and 75 000 N/mm^2 with a modulus of rupture in bending of 45 N/mm^2 . Heat-strengthened plate has a modulus of rupture of about four and a half times this figure.

3.1.6 Composite Materials

Composite materials consist of strong fibres such as glass or carbon set in a matrix of plastic or epoxy resin, which is mechanically and chemically protective. The fibres may be continuous or discontinuous but possess a strength very much greater than that of the same bulk materials. For example, carbon fibres have a tensile strength of the order of 2400 N/mm^2 and a modulus of elasticity of 400 000 N/mm^2.

A sheet of fibre-reinforced material is anisotropic, i.e. its properties depend on the direction of the fibres. Generally, therefore, in structural form two or more sheets are sandwiched together to form a *lay-up* so that the fibre directions match those of the major loads.

In the early stages of the development of composite materials glass fibres were used in a matrix of epoxy resin. This glass-reinforced plastic (GRP) was used for radomes and helicopter blades but found limited use in components of fixed wing aircraft due to its low stiffness. In the 1960s, new fibrous reinforcements were introduced; Kevlar, for example, is an aramid material with the same strength as glass but is stiffer. Kevlar composites are tough but poor in compression and difficult to machine, so they were used in secondary structures. Another composite, using boron fibre and developed in the USA, was the first to possess sufficient strength and stiffness for primary structures.

These composites have now been replaced by carbon-fibre-reinforced plastics (CFRP), which have similar properties to boron composites but are very much cheaper. Typically, CFRP has a modulus of the order of three times that of GRP, one and a half times that of a Kevlar composite and twice that of aluminium alloy. Its strength is three times that of aluminium alloy, approximately the same as that of GRP, and slightly less than that of Kevlar composites. CFRP does, however, suffer from some disadvantages. It is a brittle material and therefore does not yield plastically in regions of high stress concentration. Its strength is reduced by impact damage which may not be visible and the epoxy resin matrices can absorb moisture over a long period which reduces its matrix-dependent properties, such as its compressive strength; this effect increases with increase of temperature. Further, the properties of CFRP are subject to more random variation than those of metals. All these factors must be allowed for in design. On the other hand, the stiffness of CFRP is much less affected than its strength by the above and it is less prone to fatigue damage than metals. It is estimated that replacing 40% of an aluminium alloy structure by CFRP would result in a 12% saving in total structural weight.

CFRP is included in the wing, tailplane and forward fuselage of the latest Harrier development, is used in the Tornado taileron and has been used to construct a complete Jaguar wing and engine bay door for testing purposes. The use of CFRP in the fabrication of helicopter blades has led to significant increases in their service life, where fatigue resistance rather than stiffness is of primary importance. Figure 3.1-1 shows the structural complexity of a Sea

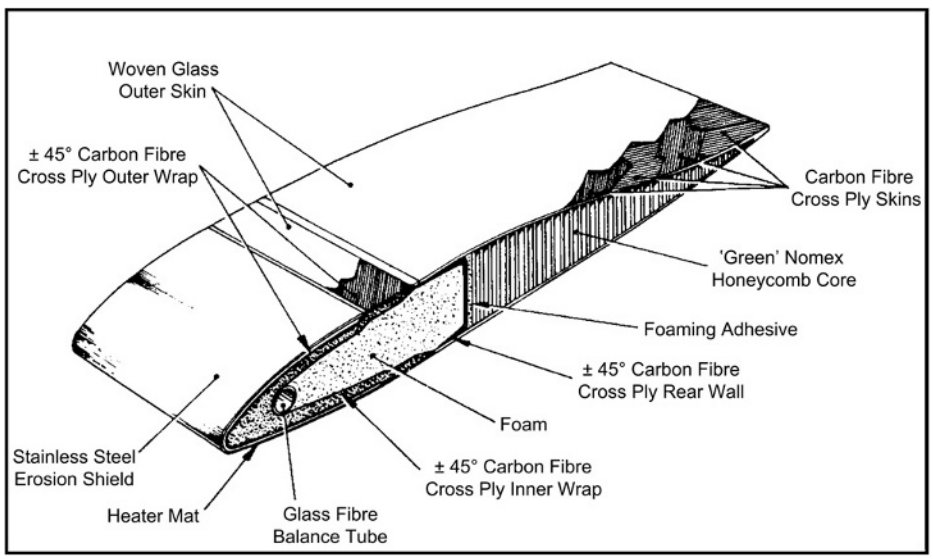


Figure 3.1-1 Sectional view of helicopter main rotor blade (courtesy Royal Aeronautical Society, *Aerospace* magazine).

King helicopter rotor blade which incorporates CFRP, GRP, stainless steel, a honeycomb core and foam filling. An additional advantage of the use of composites for helicopter rotor blades is that the moulding techniques employed allow variations of cross-section along the span, resulting in substantial aerodynamic benefits. This approach is being employed in the fabrication of the main rotor blades of the GKN Westland Helicopters EH101.

A composite (fibreglass and aluminium) is used in the tail assembly of the Boeing 777 while the leading edge of the Airbus A310–300/A320 fin assembly is of conventional reinforced glass fibre construction, reinforced at the nose to withstand bird strikes. A complete composite airframe was produced for the Beechcraft Starship turboprop executive aircraft which, however, was not a commercial success due to its canard configuration causing drag and weight penalties.

The development of composite materials is continuing with research into the removal of strength-reducing flaws and local imperfections from carbon fibres. Other matrices such as polyetheretherketone, which absorbs much less moisture than epoxy resin, has an indefinite shelf life and performs well under impact, are being developed; fabrication, however, requires much higher temperatures. Metal matrix composites such as graphite–aluminium and boron–aluminium are lightweight and retain their strength at higher temperatures than aluminium alloys, but are expensive to produce.

Generally, the use of composites in aircraft construction appears to have reached a plateau, particularly in civil subsonic aircraft where the fraction of the structure comprising composites is approximately 15%. This is due largely to the greater cost of manufacturing composites compared with aluminium alloy structures since composites require hand crafting of the materials and manual construction processes. These increased costs are particularly important in civil aircraft construction and are becoming increasingly important in military aircraft.

3.1.7 Properties of Materials

In Sections 3.1.1–3.1.6 we discussed the various materials used in aircraft construction and listed some of their properties. We shall now examine in more detail their behaviour under load and also define different types of material.

Ductility

A material is said to be *ductile* if it is capable of withstanding large strains under load before fracture occurs. These large strains are accompanied by a visible change in cross-sectional dimensions and therefore give warning of impending failure. Materials in this category include mild steel, aluminium and some of its alloys, copper and polymers.

Brittleness

A brittle material exhibits little deformation before fracture, the strain normally being below 5%. Brittle materials therefore may fail suddenly without visible warning. Included in this group are concrete, cast iron, high strength steel, timber and ceramics.

Elastic materials

A material is said to be *elastic* if deformations disappear completely on removal of the load. All known engineering materials are, in addition, *linearly elastic* within certain limits of stress so that strain, within these limits, is directly proportional to stress.

Plasticity

A material is perfectly *plastic* if no strain disappears after the removal of load. Ductile materials are *elastoplastic* and behave in an elastic manner until the *elastic limit* is reached after which they behave plastically. When the stress is relieved the elastic component of the strain is recovered but the plastic strain remains as *a permanent set*.

Isotropic materials

In many materials the elastic properties are the same in all directions at each point in the material although they may vary from point to point, such a material is known as *isotropic*. An isotropic material having the same properties at all points is known as *homogeneous* (e.g. mild steel).

Anisotropic materials

Materials having varying elastic properties in different directions are known as *anisotropic*.

Orthotropic materials

Although a structural material may possess different elastic properties in different directions, this variation may be limited, as in the case of timber which has just two values of Young's modulus, one in the direction of the grain and one perpendicular to the grain. A material whose elastic properties are limited to three different values in three mutually perpendicular directions is known as *orthotropic*.

3.1.7.1 Testing of Engineering Materials

The properties of engineering materials are determined mainly by the mechanical testing of specimens machined to prescribed sizes and shapes. The testing may be static or dynamic in nature depending on the particular property being investigated. Possibly the most common mechanical static tests are tensile and compressive tests which are carried out on a wide range of materials. Ferrous and non-ferrous metals are subjected to both forms of test, while compression tests are usually carried out on many non-metallic materials. Other static tests include bending, shear and hardness tests, while the toughness of a material, in other words its ability to withstand shock loads, is determined by impact tests.

Tensile tests

Tensile tests are normally carried out on metallic materials and, in addition, timber. Test pieces are machined from a batch of material, their dimensions being specified by Codes of Practice. They are commonly circular in cross-section, although flat test pieces having rectangular cross-sections are used when the batch of material is in the form of a plate. A typical test piece would have the dimensions specified in Fig. 3.1-2. Usually the diameter of a central portion of the test piece is fractionally less than that of the remainder to ensure that the test piece fractures between the gauge points.

Before the test begins, the mean diameter of the test piece is obtained by taking measurements at several sections using a micrometer screw gauge. Gauge points are punched at the required gauge length, the test piece is placed in the testing machine and a suitable strain measuring device, usually an extensometer, is attached to the test piece at the gauge points so that the extension is measured over the given gauge length. Increments of load are applied and the corresponding extensions recorded. This procedure continues until yield occurs, when the extensometer is removed as a precaution against the damage which would be caused if the test piece fractured unexpectedly. Subsequent extensions are measured by dividers placed in the gauge points until, ultimately, the test piece fractures. The final gauge length and the diameter of the test piece in the region of the fracture are measured so that the percentage elongation and percentage reduction in area may be calculated. These two parameters give a measure of the ductility of the material.

A stress–strain curve is drawn (see Figs 3.1-9 and 3.1-13), the stress normally being calculated on the basis of the original cross-sectional area of the test piece, i.e. a *nominal stress* as opposed to an *actual stress* (which is based on the actual area of cross-section).

For ductile materials there is a marked difference in the latter stages of the test as a considerable reduction in cross-sectional area occurs between yield and fracture. From the stress–strain curve the ultimate stress, the yield stress and Young's modulus, E, are obtained.

There are a number of variations on the basic tensile test described above. Some of these depend upon the amount of additional information required and some upon the choice of equipment. There is a wide range of strain measuring devices to choose from, extending from different makes of mechanical extensometer, e.g. Huggenberger, Lindley, Cambridge, to the electrical resistance strain gauge. The last would normally be used on flat test pieces, one on each face to eliminate the effects of possible bending. At the same time a strain gauge could be attached in a direction

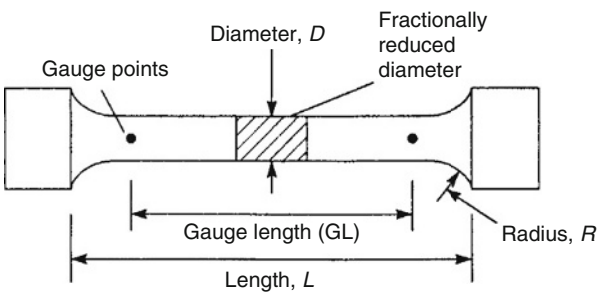


Figure 3.1-2 Standard cylindrical test piece.

perpendicular to the direction of loading so that lateral strains are measured. The ratio lateral strain/longitudinal strain is Poisson's ratio, ν.

Testing machines are usually driven hydraulically. More sophisticated versions employ load cells to record load and automatically plot load against extension or stress against strain on a pen recorder as the test proceeds, an advantage when investigating the distinctive behaviour of mild steel at yield.

Compression tests

A compression test is similar in operation to a tensile test, with the obvious difference that the load transmitted to the test piece is compressive rather than tensile. This is achieved by placing the test piece between the platens of the testing machine and reversing the direction of loading. Test pieces are normally cylindrical and are limited in length to eliminate the possibility of failure being caused by instability. Again contractions are measured over a given gauge length by a suitable strain measuring device.

Variations in test pieces occur when only the ultimate strength of the material in compression is required. For this purpose concrete test pieces may take the form of cubes having edges approximately 10 cm long, while mild steel test pieces are still cylindrical in section but are of the order of 1 cm long.

Bending tests

Many structural members are subjected primarily to bending moments. Bending tests are therefore carried out on simple beams constructed from the different materials to determine their behaviour under this type of load.

Two forms of loading are employed the choice depending upon the type specified in Codes of Practice for the particular material. In the first a simply supported beam is subjected to a 'two-point' loading system as shown in Fig. 3.1-3(a). Two concentrated loads are applied symmetrically to the beam, producing zero shear force and constant bending moment in the central span of the beam (Fig. 3.1-3(b) and (c)). The condition of pure bending is therefore achieved in the central span.

The second form of loading system consists of a single concentrated load at mid-span (Fig. 3.1-4(a)) which produces the shear force and bending moment diagrams shown in Fig. 3.1-4(b) and (c).

The loads may be applied manually by hanging weights on the beam or by a testing machine. Deflections are measured by a dial gauge placed underneath the beam. From the recorded results a load–deflection diagram is plotted.

For most ductile materials the test beams continue to deform without failure and fracture does not occur. Thus plastic properties, for example the ultimate strength in bending, cannot be determined for such materials. In the case of brittle materials, including cast iron, timber and various plastics, failure does occur, so that plastic properties can

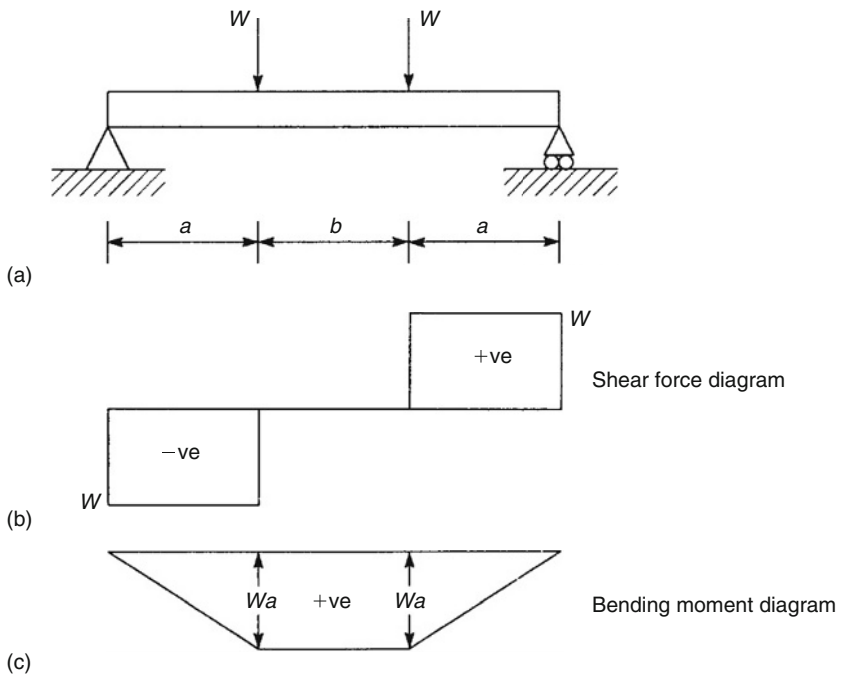


Figure 3.1-3 Bending test on a beam, 'two-point' load.

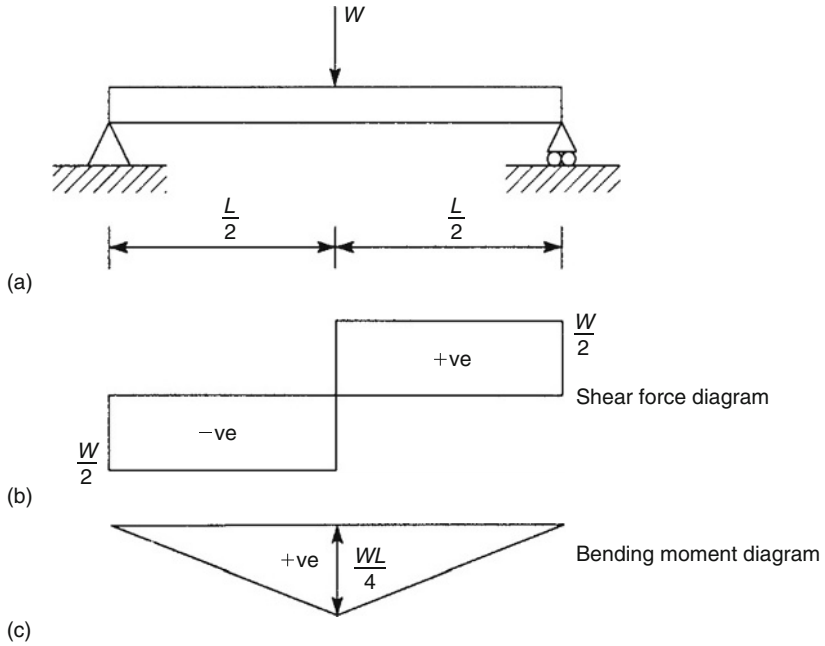


Figure 3.1-4 Bending test on a beam, single load.

be evaluated. For such materials the ultimate strength in bending is defined by the *modulus of rupture*. This is taken to be the maximum direct stress in bending, $\sigma_{x,u}$, corresponding to the ultimate moment M_u, and is assumed to be related to M_u by the elastic relationship

$$\sigma_{x,u} = \frac{M_u}{I} y_{max}$$

Other bending tests are designed to measure the ductility of a material and involve the bending of a bar round a pin. The angle of bending at which the bar starts to crack is then taken as an indication of its ductility.

Shear tests

Two main types of shear test are used to determine the shear properties of materials. One type investigates the direct or transverse shear strength of a material and is used in connection with the shear strength of bolts, rivets and beams. A typical arrangement is shown diagrammatically in Fig. 3.1-5 where the test piece is clamped to a block and the load is applied through the shear tool until failure occurs. In the arrangement shown the test piece is subjected to double shear, whereas if it is extended only partially across the gap in the block it would be subjected to single shear. In either case the average shear strength is taken as the maximum load divided by the shear resisting area.

The other type of shear test is used to evaluate the basic shear properties of a material, such as the shear modulus, G, the shear stress at yield and the ultimate shear stress. In the usual form of test a solid circular-section test piece is placed in a torsion machine and twisted by controlled increments of torque. The corresponding angles of twist are recorded and torque–twist diagrams plotted from which the shear properties of the material are obtained. The method is similar to that used to determine the tensile properties of a material from a tensile test.

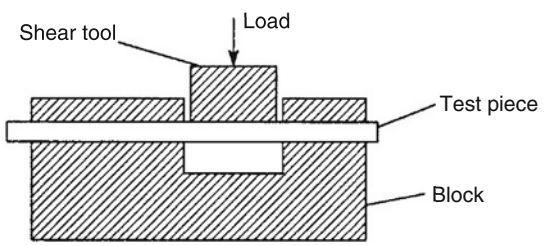


Figure 3.1-5 Shear test.

Hardness tests

The machinability of a material and its resistance to scratching or penetration are determined by its 'hardness'. There also appears to be a connection between the hardness of some materials and their tensile strength so that hardness tests may be used to determine the properties of a finished structural member where tensile and other tests would be impracticable. Hardness tests are also used to investigate the effects of heat treatment, hardening and tempering and of cold forming. Two types of hardness test are in common use: *indentation tests* and *scratch and abrasion tests.*

Indentation tests may be subdivided into two classes: static and dynamic. Of the static tests the *Brinell* is the most common. In this a hardened steel ball is pressed into the material under test by a static load acting for a fixed period of time. The load in kg divided by the spherical area of the indentation in mm^2 is called the *Brinell hardness number* (BHN). In Fig. 3.1-6, if D is the diameter of the ball, F the load in kg, h the depth of the indentation and d the diameter of the indentation, then

$$\text{BHN} = \frac{F}{\pi D h} = \frac{2F}{\pi D[D - \sqrt{D^2 - d^2}]}$$

In practice, the hardness number of a given material is found to vary with F and D so that for uniformity the test is standardized. For steel and hard materials $F = 3000$ kg and $D = 10$ mm while for soft materials $F = 500$ kg and $D = 10$ mm; in addition the load is usually applied for 15 s.

In the *Brinell* test the dimensions of the indentation are measured by means of a microscope. To avoid this rather tedious procedure, direct reading machines have been devised of which the *Rockwell* is typical. The indenting tool, again a hardened sphere, is first applied under a definite light load. This indenting tool is then replaced by a diamond cone with a rounded point which is then applied under a specified indentation load. The difference between the depth of the indentation under the two loads is taken as a measure of the hardness of the material and is read directly from the scale.

A typical dynamic hardness test is performed by the *Shore Scleroscope* which consists of a small hammer approximately 20 mm long and 6 mm in diameter fitted with a blunt, rounded, diamond point. The hammer is guided by a vertical glass tube and allowed to fall freely from a height of 25 cm onto the specimen, which it indents before rebounding. A certain proportion of the energy of the hammer is expended in forming the indentation so that the height of the rebound, which depends upon the energy still possessed by the hammer, is taken as a measure of the hardness of the material.

A number of tests have been devised to measure the 'scratch hardness' of materials. In one test, the smallest load in grams which, when applied to a diamond point, produces a scratch visible to the naked eye on a polished specimen of material is called its hardness number. In other tests the magnitude of the load required to produce a definite width of scratch is taken as the measure of hardness. Abrasion tests, involving the shaking over a period of time of several specimens placed in a container, measure the resistance to wear of some materials. In some cases, there appears to be a connection between wear and hardness number although the results show no level of consistency.

Impact tests

It has been found that certain materials, particularly heat-treated steels, are susceptible to failure under shock loading whereas an ordinary tensile test on the same material would show no abnormality. Impact tests measure the ability of materials to withstand shock loads and provide an indication of their *toughness.* Two main tests are in use, the *Izod* and the *Charpy.*

Both tests rely on a striker or weight attached to a pendulum. The pendulum is released from a fixed height, the weight strikes a notched test piece and the angle through which the pendulum then swings is a measure of the toughness of the material. The arrangement for the Izod test is shown diagrammatically in Fig. 3.1-7(a). The

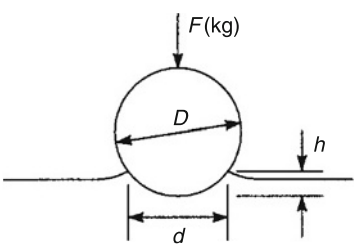


Figure 3.1-6 *Brinell* hardness test.

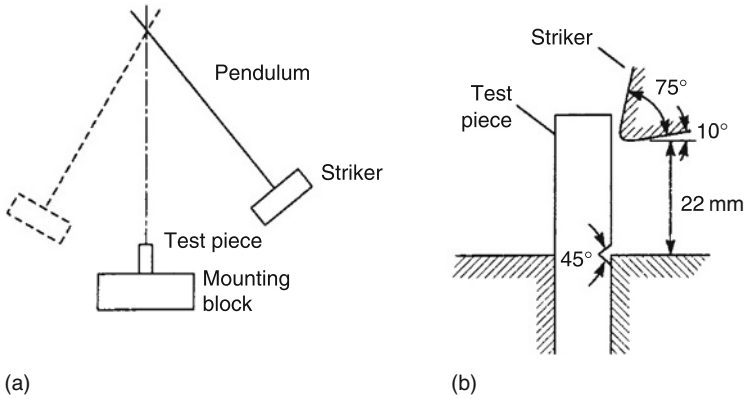


Figure 3.1-7 Izod impact test.

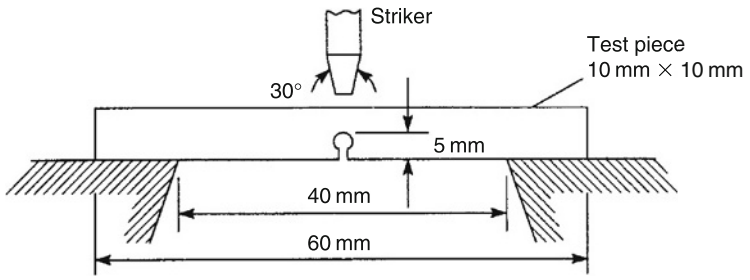


Figure 3.1-8 Charpy impact test.

specimen and the method of mounting are shown in detail in Fig. 3.1-7(b). The Charpy test is similar in operation except that the test piece is supported in a different manner as shown in the plan view in Fig. 3.1-8.

3.1.7.2 Stress–Strain Curves

We shall now examine in detail the properties of the different materials from the viewpoint of the results obtained from tensile and compression tests.

Low carbon steel (mild steel)

A nominal stress–strain curve for mild steel, a ductile material, is shown in Fig. 3.1-9 From 0 to 'a' the stress–strain curve is linear, the material in this range obeying Hooke's law. Beyond 'a', the *limit of proportionality,* stress is no longer proportional to strain and the stress–strain curve continues to 'b', the *elastic limit,* which is defined as the maximum

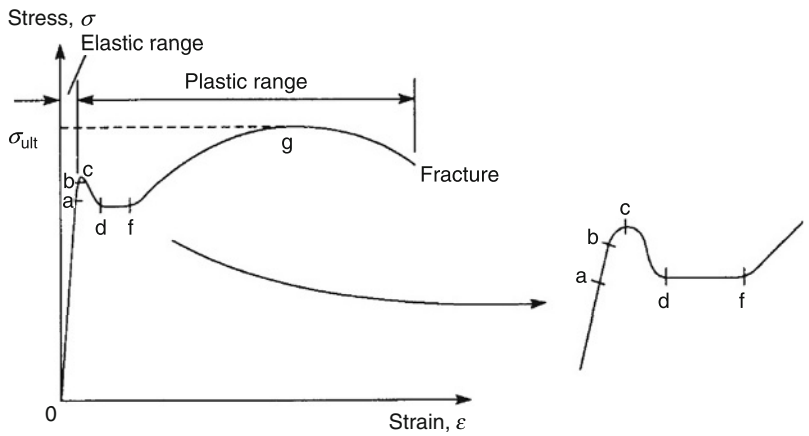


Figure 3.1-9 Stress–strain curve for mild steel.

stress that can be applied to a material without producing a permanent plastic deformation or *permanent set* when the load is removed. In other words, if the material is stressed beyond 'b' and the load then removed, a residual strain exists at zero load. For many materials it is impossible to detect a difference between the limit of proportionality and the elastic limit. From 0 to 'b' the material is said to be in the *elastic range* while from 'b' to fracture the material is in the *plastic range*. The transition from the elastic to the plastic range may be explained by considering the arrangement of crystals in the material. As the load is applied, slipping occurs between the crystals which are aligned most closely to the direction of load. As the load is increased, more and more crystals slip with each equal load increment until appreciable strain increments are produced and the plastic range is reached.

A further increase in stress from 'b' results in the mild steel reaching its *upper yield point* at 'c' followed by a rapid fall in stress to its *lower yield point* at 'd'. The existence of a lower yield point for mild steel is a peculiarity of the tensile test wherein the movement of the ends of the test piece produced by the testing machine does not proceed as rapidly as its plastic deformation; the load therefore decreases, as does the stress. From 'd' to 'f' the strain increases at a roughly constant value of stress until *strain hardening* again causes an increase in stress. This increase in stress continues, accompanied by a large increase in strain to 'g', the *ultimate stress*, σ_{ult}, of the material. At this point the test piece begins, visibly, to 'neck' as shown in Fig. 3.1-10. The material in the test piece in the region of the 'neck' is almost perfectly plastic at this stage and from this point, onwards to fracture, there is a reduction in nominal stress.

For mild steel, yielding occurs at a stress of the order of 300 N/mm^2. At fracture the strain (i.e. the elongation) is of the order of 30%. The gradient of the linear portion of the stress–strain curve gives a value for Young's modulus in the region of 200 000 N/mm^2.

The characteristics of the fracture are worthy of examination. In a cylindrical test piece the two halves of the fractured test piece have ends which form a 'cup and cone' (Fig. 3.1-11). The actual failure planes in this case are inclined at approximately 45° to the axis of loading and coincide with planes of maximum shear stress. Similarly, if a flat tensile specimen of mild steel is polished and then stressed, a pattern of fine lines appears on the polished surface at yield. These lines, which were first discovered by Lüder in 1854, intersect approximately at right angles and are inclined at 45° to the axis of the specimen, thereby coinciding with planes of maximum shear stress. These forms of yielding and fracture suggest that the crystalline structure of the steel is relatively weak in shear with yielding taking the form of the sliding of one crystal plane over another rather than the tearing apart of two crystal planes.

The behaviour of mild steel in compression is very similar to its behaviour in tension, particularly in the elastic range. In the plastic range it is not possible to obtain ultimate and fracture loads since, due to compression, the area of cross-section increases as the load increases producing a 'barrelling' effect as shown in Fig. 3.1-12. This increase in

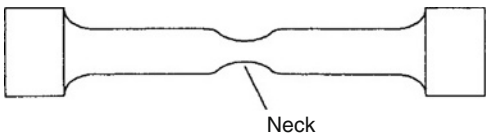


Figure 3.1-10 'Necking' of a test piece in the plastic range.

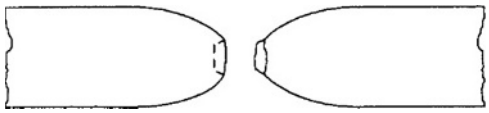


Figure 3.1-11 'Cup-and-cone' failure of a mild steel test piece.

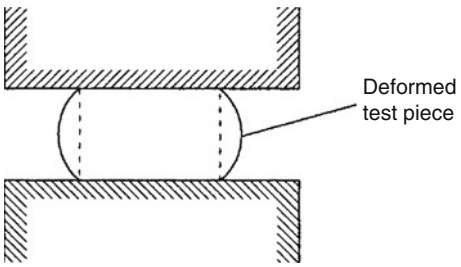


Figure 3.1-12 'Barrelling' of a mild steel test piece in compression.

cross-sectional area tends to decrease the true stress, thereby increasing the load resistance. Ultimately a flat disc is produced. For design purposes the ultimate stresses of mild steel in tension and compression are assumed to be the same.

Higher grades of steel have greater strengths than mild steel but are not as ductile. They also possess the same Young's modulus so that the higher stresses are accompanied by higher strains.

Aluminium

Aluminium and some of its alloys are also ductile materials, although their stress–strain curves do not have the distinct yield stress of mild steel. A typical stress–strain curve is shown in Fig. 3.1-13. The points 'a' and 'b' again mark the limit of proportionality and elastic limit, respectively, but are difficult to determine experimentally. Instead a *proof stress* is defined which is the stress required to produce a given permanent strain on removal of the load. In Fig. 3.1-13, a line drawn parallel to the linear portion of the stress–strain curve from a strain of 0.001 (i.e. a strain of 0.1%) intersects the stress–strain curve at the 0.1% proof stress. For elastic design this, or the 0.2% proof stress, is taken as the working stress.

Beyond the limit of proportionality the material extends plastically, reaching its ultimate stress, σ_{ult}, at 'd' before finally fracturing under a reduced nominal stress at 'f'.

A feature of the fracture of aluminium alloy test pieces is the formation of a 'double cup' as shown in Fig. 3.1-14, implying that failure was initiated in the central portion of the test piece while the outer surfaces remained intact. Again considerable 'necking' occurs.

In compression tests on aluminium and its ductile alloys similar difficulties are encountered to those experienced with mild steel. The stress–strain curve is very similar in the elastic range to that obtained in a tensile test but the ultimate strength in compression cannot be determined; in design its value is assumed to coincide with that in tension.

Aluminium and its alloys can suffer a form of corrosion particularly in the salt laden atmosphere of coastal regions. The surface becomes pitted and covered by a white furry deposit. This can be prevented by an electrolytic process called *anodizing* which covers the surface with an inert coating. Aluminium alloys will also corrode if they are placed in direct contact with other metals, such as steel. To prevent this, plastic is inserted between the possible areas of contact.

Brittle materials

These include cast iron, high strength steel, concrete, timber, ceramics, glass, etc. The plastic range for brittle materials extends to only small values of strain. A typical stress–strain curve for a brittle material under tension is shown in Fig. 3.1-15. Little or no yielding occurs and fracture takes place very shortly after the elastic limit is reached.

The fracture of a cylindrical test piece takes the form of a single failure plane approximately perpendicular to the direction of loading with no visible 'necking' and an elongation of the order of 2–3%.

In compression the stress–strain curve for a brittle material is very similar to that in tension except that failure occurs at a much higher value of stress; for concrete the ratio is of the order of 10 : 1. This is thought to be due to

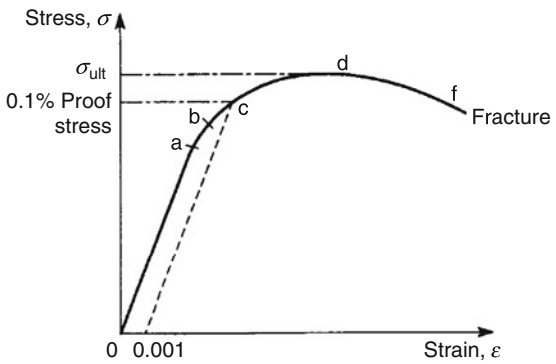


Figure 3.1-13 Stress–strain curve for aluminium

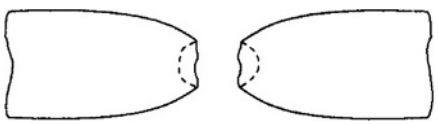


Figure 3.1-14 'Double-cup' failure of an aluminium alloy test piece.

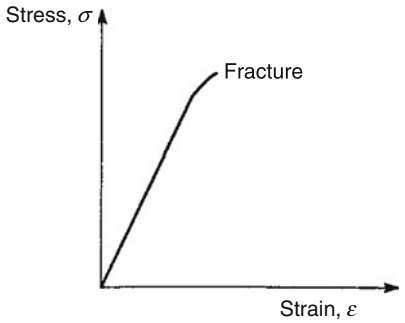


Figure 3.1-15 Stress–strain curve for a brittle material.

the presence of microscopic cracks in the material, giving rise to high stress concentrations which are more likely to have a greater effect in reducing tensile strength than compressive strength.

Composites

Fibre composites have stress–strain characteristics which indicate that they are brittle materials (Fig. 3.1-16). There is little or no plasticity and the modulus of elasticity is less than that of steel and aluminium alloy. However, the fibres themselves can have much higher values of strength and modulus of elasticity than the composite. For example, carbon fibres have a tensile strength of the order 2400 N/mm^2 and a modulus of elasticity of 400 000 N/mm^2.

Fibre composites are highly durable, require no maintenance and can be used in hostile chemical and atmospheric environments; vinyls and epoxy resins provide the best resistance.

All the stress–strain curves described in the preceding discussion are those produced in tensile or compression tests in which the strain is applied at a negligible rate. A rapid strain application would result in significant changes in the apparent properties of the materials giving possible variations in yield stress of up to 100%.

3.1.7.3 Strain Hardening

The stress–strain curve for a material is influenced by the *strain history*, or the loading and unloading of the material, within the plastic range. For example, in Fig. 3.1-17 a test piece is initially stressed in tension beyond the yield stress at, 'a', to a value at 'b'. The material is then unloaded to 'c' and reloaded to 'f' producing an increase in yield stress from the value at 'a' to the value at 'd'. Subsequent unloading to 'g' and loading to 'j' increases the yield stress still further to the value at 'h'. This increase in strength resulting from the loading and unloading is known as *strain hardening*. It can be seen from Fig. 3.1-17 that the stress–strain curve during the unloading and loading cycles forms loops (the shaded areas in Fig. 3.1-17). These indicate that strain energy is lost during the cycle, the energy being dissipated in the form of heat produced by internal friction. This energy loss is known as *mechanical hysteresis* and the loops as *hysteresis loops*. Although the ultimate stress is increased by strain hardening it is not influenced to the same extent as yield stress. The increase in strength produced by strain hardening is accompanied by decreases in toughness and ductility.

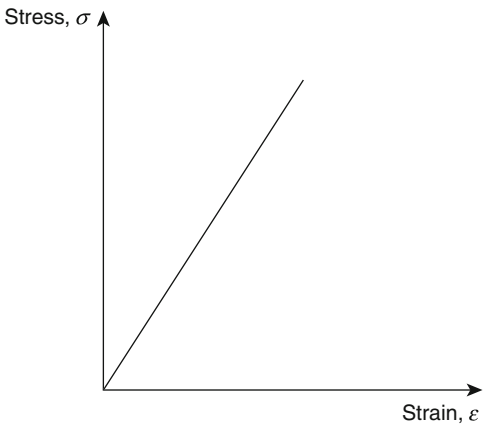


Figure 3.1-16 Stress–strain curve for a fibre composite.

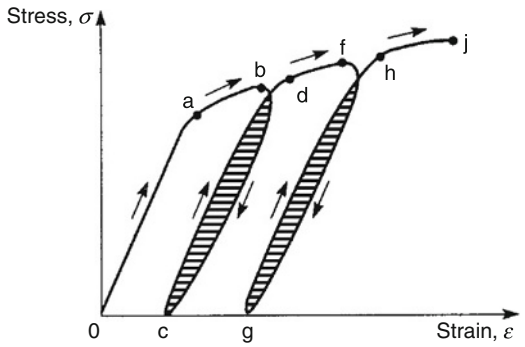


Figure 3.1-17 Strain hardening of a material.

3.1.7.4 Creep and Relaxation

We have seen earlier that a given load produces a calculable value of stress in a structural member and hence a corresponding value of strain once the full value of the load is transferred to the member. However, after this initial or 'instantaneous' stress and its corresponding value of strain have been attained, a great number of structural materials continue to deform slowly and progressively under load over a period of time. This behaviour is known as *creep*. A typical creep curve is shown in Fig. 3.1-18.

Some materials, such as plastics and rubber, exhibit creep at room temperatures but most structural materials require high temperatures or long-duration loading at moderate temperatures. In some 'soft' metals, such as zinc and lead, creep occurs over a relatively short period of time, whereas materials such as concrete may be subject to creep over a period of years. Creep occurs in steel to a slight extent at normal temperatures but becomes very important at temperatures above 316°C.

Closely related to creep is *relaxation*. Whereas creep involves an increase in strain under constant stress, relaxation is the decrease in stress experienced over a period of time by a material subjected to a constant strain.

3.1.7.5 Fatigue

Structural members are frequently subjected to repetitive loading over a long period of time. For example, the members of a bridge structure suffer variations in loading possibly thousands of times a day as traffic moves over the bridge. In these circumstances a structural member may fracture at a level of stress substantially below the ultimate stress for non-repetitive static loads; this phenomenon is known as *fatigue*.

Fatigue cracks are most frequently initiated at sections in a structural member where changes in geometry, e.g. holes, notches or sudden changes in section, cause *stress concentrations*. Designers seek to eliminate such areas by

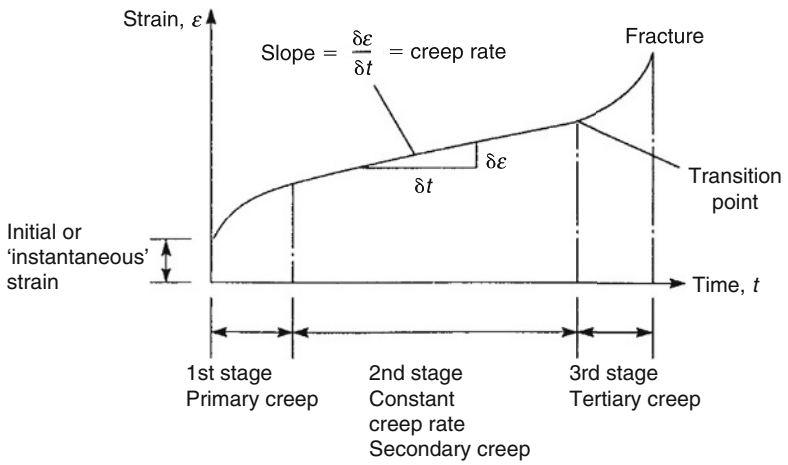


Figure 3.1-18 Typical creep curve.

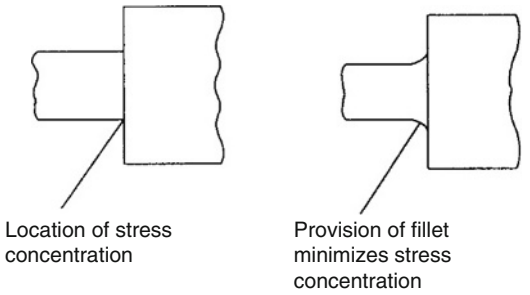


Figure 3.1-19 Stress concentration location.

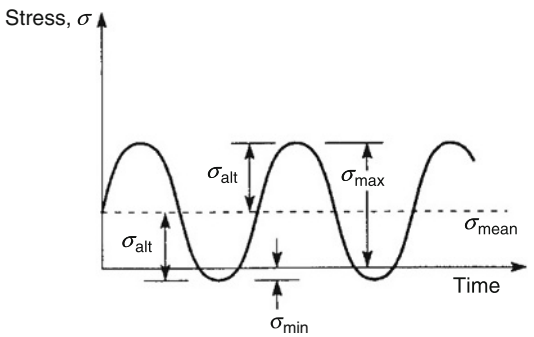


Figure 3.1-20 Alternating stress in fatigue loading.

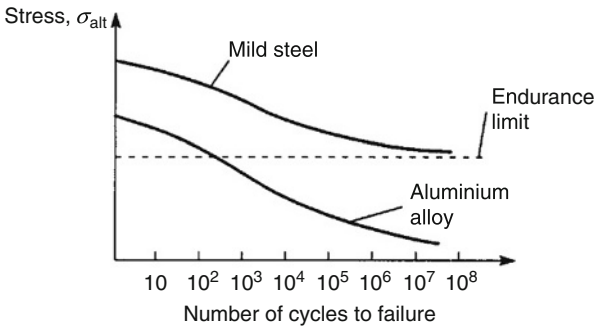


Figure 3.1-21 Stress–endurance curves.

ensuring that rapid changes in section are as smooth as possible. At re-entrant corners for example, fillets are provided as shown in Fig. 3.1-19.

Other factors which affect the failure of a material under repetitive loading are the type of loading (fatigue is primarily a problem with repeated tensile stresses due, probably, to the fact that microscopic cracks can propagate more easily under tension), temperature, the material, surface finish (machine marks are potential crack propagators), corrosion and residual stresses produced by welding.

Frequently in structural members an alternating stress, σ_{alt}, is superimposed on a static or mean stress, σ_{mean}, as illustrated in Fig. 3.1-20. The value of σ_{alt} is the most important factor in determining the number of cycles of load that produce failure. The stress σ_{alt} that can be withstood for a specified number of cycles is called the *fatigue strength* of the material. Some materials, such as mild steel, possess a stress level that can be withstood for an indefinite number of cycles. This stress is known as the *endurance limit* of the material; no such limit has been found for aluminium and its alloys. Fatigue data are frequently presented in the form of an *S–n* curve or stress–endurance curve as shown in Fig. 3.1-21.

In many practical situations the amplitude of the alternating stress varies and is frequently random in nature. The *S–n* curve does not, therefore, apply directly and an alternative means of predicting failure is required. *Miner's cumulative damage theory* suggests that failure will occur when

$$\frac{n_1}{N_1}+\frac{n_2}{N_2}+\cdots+\frac{n_r}{N_r}=1 \tag{3.1.1}$$

where n_1, n_2, …, n_r are the number of applications of stresses σ_{alt}, σ_{mean} and N_1, N_2,…, N_r are the number of cycles to failure of stresses σ_{alt}, σ_{mean}.

Problems

P.3.1.1 Describe a simple tensile test and show, with the aid of sketches, how measures of the ductility of the material of the specimen may be obtained. Sketch typical stress–strain curves for mild steel and an aluminium alloy showing their important features.

P.3.1.2 A bar of metal 25 mm in diameter is tested on a length of 250 mm. In tension the following results were recorded (Table P.3.1-2(a)).

Table P.3.1-2(a)

Load (kN)	10.4	31.2	52.0	72.8
Extension (mm)	0.036	0.089	0.140	0.191

A torsion test gave the following results (Table P.3.1-2(b)).

Table P.3.1-2(b)

Torque (kN m)	0.051	0.152	0.253	0.354
Angle of twist (degrees)	0.24	0.71	1.175	1.642

Represent these results in graphical form and hence determine Young's modulus, *E*, the modulus of rigidity, G, Poisson's ratio, ν, and the bulk modulus, *K*, for the metal.

Ans. $E \simeq 205\,000$ N/mm^2, $G \simeq 80\,700$ N/mm^2, $\nu \simeq 0.27^2$, $K \simeq 148\,500$ N/mm^2.

P.3.1.3 The actual stress–strain curve for a particular material is given by $\sigma = C\epsilon^n$ where *C* is a constant. Assuming that the material suffers no change in volume during plastic deformation, derive an expression for the nominal stress–strain curve and show that this has a maximum value when $\epsilon = n/(1-n)$.

Ans. $\sigma_{nom} = C\epsilon^n/(1+\epsilon)$.

P.3.1.4 A structural member is to be subjected to a series of cyclic loads which produce different levels of alternating stress as shown in Table P.3.1-4. Determine whether or not a fatigue failure is probable.

Ans. Not probable ($n_1/N_1 + n_2/N_2 + \cdots = 0.39$).

Table P.3.1-4

Loading	Number of cycles	Number of cycles to failure
1	10^4	5×10^4
2	10^5	10^6
3	10^6	24×10^7
4	10^7	12×10^7

Aerospace Engineering Desk Reference; ISBN 9781856175753

3.2 Structural Components

T.H.G. Megson

Aircraft are generally built up from the basic components of wings, fuselages, tail units and control surfaces. There are variations in particular aircraft, for example, a delta wing aircraft would not necessarily possess a horizontal tail although this is present in a canard configuration such as that of the Eurofighter (Typhoon). Each component has one or more specific functions and must be designed to ensure that it can carry out these functions safely. In this chapter we shall describe the various loads to which aircraft components are subjected, their function and fabrication and also the design of connections.

3.2.1 Loads on Structural Components

The structure of an aircraft is required to support two distinct classes of load: the first, termed *ground loads*, includes all loads encountered by the aircraft during movement or transportation on the ground such as taxiing and landing loads, towing and hoisting loads; while the second, *air loads*, comprises loads imposed on the structure during flight by manoeuvres and gusts. In addition, aircraft designed for a particular role encounter loads peculiar to their sphere of operation. Carrier born aircraft, for instance, are subjected to catapult take-off and arrested landing loads: most large civil and practically all military aircraft have pressurized cabins for high altitude flying; amphibious aircraft must be capable of landing on water and aircraft designed to fly at high speed at low altitude, e.g. the Tornado, require a structure of above average strength to withstand the effects of flight in extremely turbulent air.

The two classes of loads may be further divided into *surface forces* which act upon the surface of the structure, e.g. aerodynamic and hydrostatic pressure, and *body forces* which act over the volume of the structure and are produced by gravitational and inertial effects. Calculation of the distribution of aerodynamic pressure over the various surfaces of an aircraft's structure is presented in numerous texts on aerodynamics and will therefore not be attempted here. We shall, however, discuss the types of load induced by these various effects and their action on the different structural components.

Basically, all air loads are the resultants of the pressure distribution over the surfaces of the skin produced by steady flight, manoeuvre or gust conditions. Generally, these resultants cause direct loads, bending, shear and torsion in all parts of the structure in addition to local, normal pressure loads imposed on the skin.

Conventional aircraft usually consist of fuselage, wings and tailplane. The fuselage contains crew and payload, the latter being passengers, cargo, weapons plus fuel, depending on the type of aircraft and its function; the wings provide the lift and the tailplane is the main contributor to directional control. In addition, ailerons, elevators and the rudder enable the pilot to manoeuvre the aircraft and maintain its stability in flight, while wing flaps provide the necessary increase of lift for take-off and landing. Figure 3.2-1 shows typical aerodynamic force resultants experienced by an aircraft in steady flight.

The force on an aerodynamic surface (wing, vertical or horizontal tail) results from a differential pressure distribution caused by incidence, camber or a combination of both. Such a pressure distribution, shown in Fig. 3.2-2(a), has vertical (lift) and horizontal (drag) resultants acting at a centre of pressure (CP). (In practice, lift and drag are measured perpendicular and parallel to the flight path, respectively.) Clearly the position of the CP changes as the pressure distribution varies with speed or wing incidence.

However, there is, conveniently, a point in the aerofoil section about which the moment due to the lift and drag forces remains constant. We therefore replace the lift and drag forces acting at the CP by lift and drag forces acting at the aerodynamic centre (AC) plus a constant moment M_0 as shown in Fig. 3.2-2(b). (Actually, at high Mach numbers the position of the AC changes due to compressibility effects.)

While the chordwise pressure distribution fixes the position of the resultant aerodynamic load in the wing cross-section, the spanwise distribution locates its position in relation, say, to the wing root. A typical distribution for a wing/fuselage combination is shown in Fig. 3.2-3. Similar distributions occur on horizontal and vertical tail surfaces.

We see therefore that wings, tailplane and the fuselage are each subjected to direct, bending, shear and torsional loads and must be designed to withstand critical combinations of these. Note that manoeuvres and gusts do not introduce different loads but result only in changes of magnitude and position of the type of existing loads shown in Fig. 3.2-1. Over and above these basic in-flight loads, fuselages may be pressurized and thereby support hoop stresses, wings may carry weapons and/or extra fuel tanks with resulting additional aerodynamic and body forces contributing

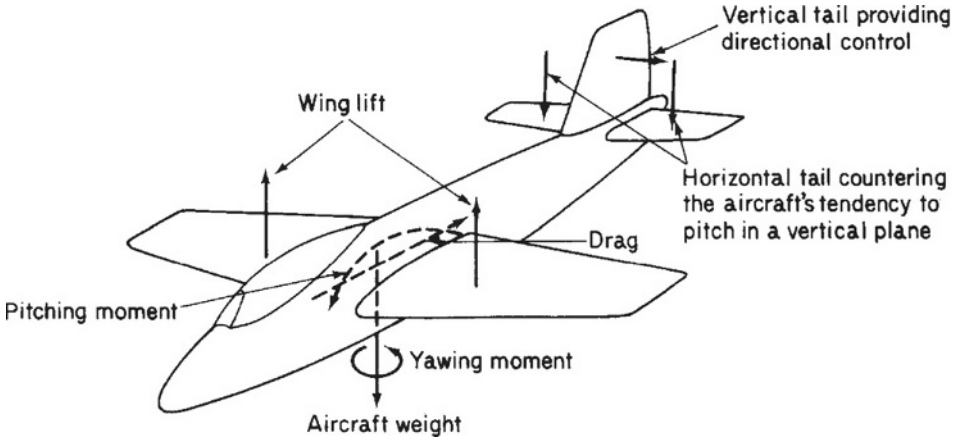


Figure 3.2-1 Principal aerodynamic forces on an aircraft during flight.

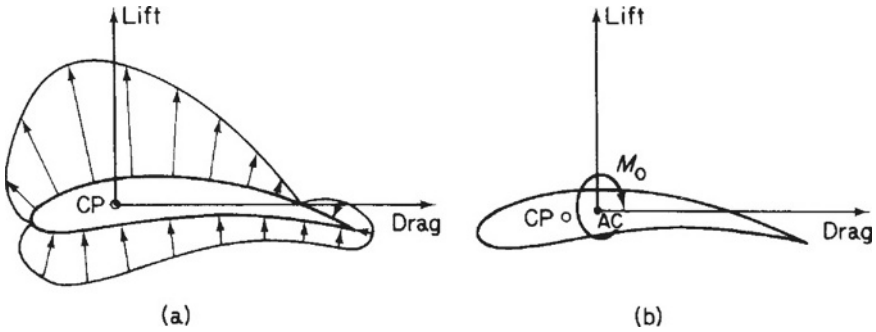


Figure 3.2-2 (a) Pressure distribution around an aerofoil; (b) transference of lift and drag loads to the AC.

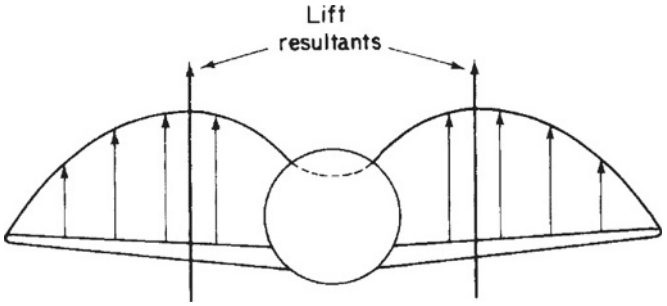


Figure 3.2-3 Typical lift distribution for a wing/fuselage combination.

to the existing bending, shear and torsion, while the thrust and weight of engines may affect either fuselage or wings depending on their relative positions.

Ground loads encountered in landing and taxiing subject the aircraft to concentrated shock loads through the undercarriage system. The majority of aircraft have their main undercarriage located in the wings, with a nosewheel or tailwheel in the vertical plane of symmetry. Clearly the position of the main undercarriage should be so as to produce minimum loads on the wing structure compatible with the stability of the aircraft during ground manoeuvres. This may be achieved by locating the undercarriage just forward of the flexural axis of the wing and as close to the wing root as possible. In this case the shock landing load produces a given shear, minimum bending plus torsion, with the latter being reduced as far as practicable by offsetting the torque caused by the vertical load in the undercarriage leg by a torque in an opposite sense due to braking.

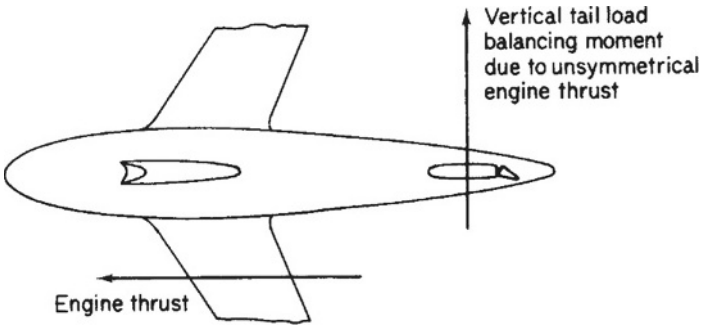


Figure 3.2-4 Fuselage and wing bending caused by an unsymmetrical engine load.

Other loads include engine thrust on the wings or fuselage which acts in the plane of symmetry but may, in the case of engine failure, cause severe fuselage bending moments, as shown in Fig. 3.2-4; concentrated shock loads during a catapult launch; and hydrodynamic pressure on the fuselages or floats of seaplanes.

In Chapter 3.3 we shall examine in detail the calculation of ground and air loads for a variety of cases.

3.2.2 Function of Structural Components

The basic functions of an aircraft's structure are to transmit and resist the applied loads; to provide an aerodynamic shape and to protect passengers, payload, systems, etc. from the environmental conditions encountered in flight. These requirements, in most aircraft, result in thin shell structures where the outer surface or skin of the shell is usually supported by longitudinal stiffening members and transverse frames to enable it to resist bending, compressive and torsional loads without buckling. Such structures are known as *semi-monocoque*, while thin shells which rely entirely on their skins for their capacity to resist loads are referred to as *monocoque*.

First, we shall consider wing sections which, while performing the same function, can differ widely in their structural complexity, as can be seen by comparing Figs 3.2-5 and 3.2-6. In Fig. 3.2-5, the wing of the small, light passenger aircraft, the De Havil-land Canada Twin Otter, comprises a relatively simple arrangement of two spars, ribs, stringers and skin, while the wing of the Harrier in Fig. 3.2-6 consists of numerous spars, ribs and skin. However, no matter how complex the internal structural arrangement the different components perform the same kind of function. The shape of the cross-section is governed by aerodynamic considerations and clearly must be maintained for all combinations of load; this is one of the functions of the ribs. They also act with the skin in resisting the distributed aerodynamic pressure loads; they distribute concentrated loads (e.g. undercarriage and additional wing store loads) into the structure and redistribute stress around discontinuities, such as undercarriage wells, inspection panels and fuel tanks, in the wing surface. Ribs increase the column buckling stress of the longitudinal stiffeners by providing end restraint and establishing their column length; in a similar manner they increase the plate buckling stress of the skin panels. The dimensions of ribs are governed by their spanwise position in the wing and by the loads they are required to support. In the outer portions of the wing, where the cross-section may be relatively small if the wing is tapered and the loads are light, ribs act primarily as formers for the aerofoil shape. A light structure is sufficient for this purpose whereas at sections closer to the wing root, where the ribs are required to absorb and transmit large concentrated applied loads, such as those from the undercarriage, engine thrust and fuselage attachment point reactions, a much more rugged construction is necessary. Between these two extremes are ribs which support hinge reactions from ailerons, flaps and other control surfaces, plus the many internal loads from fuel, armament and systems installations.

The primary function of the wing skin is to form an impermeable surface for supporting the aerodynamic pressure distribution from which the lifting capability of the wing is derived. These aerodynamic forces are transmitted in turn to the ribs and stringers by the skin through plate and membrane action. Resistance to shear and torsional loads is supplied by shear stresses developed in the skin and spar webs, while axial and bending loads are reacted by the combined action of skin and stringers.

Although the thin skin is efficient for resisting shear and tensile loads, it buckles under comparatively low compressive loads. Rather than increase the skin thickness and suffer a consequent weight penalty, stringers are attached to the skin and ribs, thereby dividing the skin into small panels and increasing the buckling and failing stresses. This stabilizing action of the stringers on the skin is, in fact, reciprocated to some extent although the effect normal to the surface of the skin is minimal. Stringers rely chiefly on rib attachments for preventing column action in this direction.

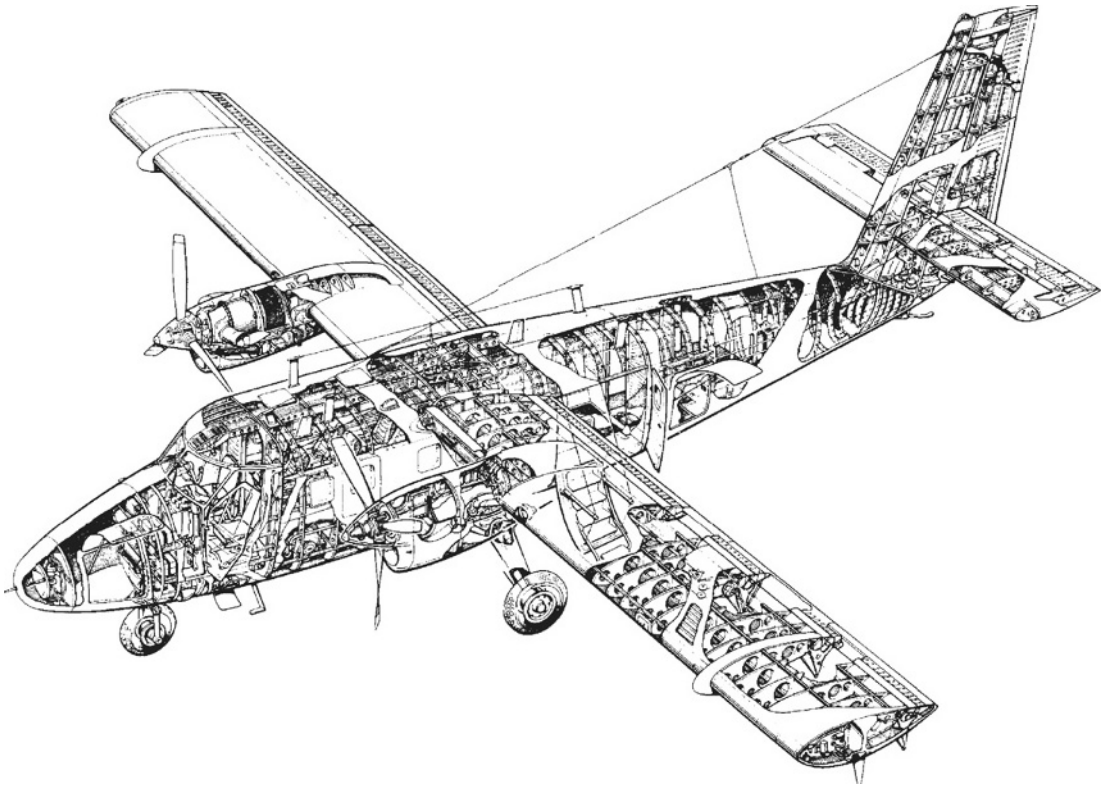


Figure 3.2-5 De Havilland Canada Twin Otter (courtesy of De Havilland Aircraft of Canada Ltd.).

We have noted in the previous paragraph the combined action of stringers and skin in resisting axial and bending loads.

The role of spar webs in developing shear stresses to resist shear and torsional loads has been mentioned previously; they perform a secondary but significant function in stabilizing, with the skin, the spar flanges or caps

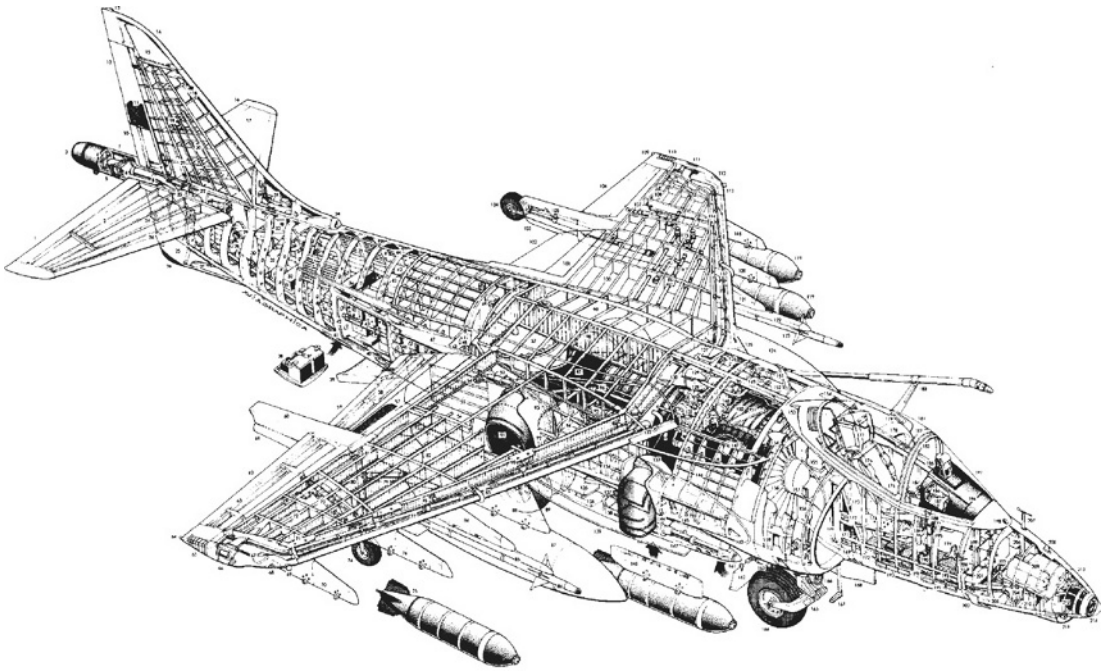


Figure 3.2-6 Harrier (courtesy of Pilot Press Ltd.).

1 Starboard all-moving tailplane
2 Tailplane composite construction
3 Tail radome
4 Military equipment
5 Tail pitch control air valve
6 Yaw control air valves
7 Tail 'bullet' fairing
8 Reaction control system air ducting
9 Trim tab actuator
10 Rudder trim tab
11 Rudder composite construction
12 Rudder
13 Antenna
14 Fin tip aerial fairing
15 Upper broad band communications antenna
16 Port tailplane
17 Graphite epoxy tailplane skin
18 Port side temperature probe
19 MAD compensator
20 Formation lighting strip
21 Fin construction
22 Fin attachment joint
23 Tailplane pivot sealing plate
24 Aerials
25 Ventral fin
26 Tail bumper
27 Lower broad band communications antenna
28 Tailplane hydraulic jack
29 Heat exchanger air exhaust
30 Aft fuselage frames
31 Rudder hydraulic actuator
32 Avionics equipment air conditioning plant
33 Avionics equipment racks
34 Heat exchanger ram air intake
35 Electrical system circuit breaker panels, port and starboard
36 Avionic equipment
37 Chaff and flare dispensers
38 Dispenser electronic control units
39 Ventral airbrake
40 Airbrake hydraulic jack
41 Formation lighting strip
42 Avionics bay access door, port and starboard
43 Avionics equipment racks
44 Fuselage frame and stringer construction
45 Rear fuselage fuel tank
46 Main undercarriage wheel bay
47 Wing root fillet
48 Wing spar/fuselage attachment joint
49 Water filler cap
50 Engine fire extinguisher bottle
51 Anti-collision light
52 Water tank
53 Flap hydraulic actuator
54 Flap hinge fitting
55 Nimonic fuselage heat shield
56 Main undercarriage bay doors (closed after cycling of mainwheels)
57 Flap vane composite construction
58 Flap composite construction
59 Starboard slotted flap, lowered
60 Outrigger wheel fairing
61 Outrigger leg doors
62 Starboard aileron
63 Aileron composite construction
64 Fuel jettison
65 Formation lighting panel
66 Roll control airvalve
67 Wing tip fairing
68 Starboard navigation light
69 Radar warning aerial
70 Outboard pylon
71 Pylon attachment joint
72 Graphite epoxy composite wing construction
73 Aileron hydraulic actuator
74 Starboard outrigger wheel
75 BL755 600-lb (272-kg) cluster bomb (CBU)
76 Intermediate pylon
77 Reaction control air ducting
78 Aileron control rod
79 Outrigger hydraulic retraction jack
80 Outrigger leg strut
81 Leg pivot fixing
82 Multi-spar wing construction
83 Leading-edge wing fence
84 Outrigger pylon
85 Missile launch rail
86 AIM-9L Sidewinder air-to-air missile
87 External fuel tank, 300 US gal (1 135 l)
88 Inboard pylon
89 Aft retracting twin mainwheels
90 Inboard pylon attachment joint
91 Rear (hot stream) swivelling exhaust nozzle
92 Position of pressure refuelling connection on port side
93 Rear nozzle bearing
94 Centre fuselage flank fuel tank
95 Hydraulic reservoir
96 Nozzle bearing cooling air duct
97 Engine exhaust divider duct
98 Wing panel centre rib
99 Centre section integral fuel tank
100 Port wing integral fuel tank
101 Flap vane
102 Port slotted flap, lowered
103 Outrigger wheel fairing
104 Port outrigger wheel
105 Torque scissor links
106 Port aileron
107 Aileron hydraulic actuator
108 Aileron/airvalve interconnection
109 Fuel jettison
110 Formation lighting panel
111 Port roll control air valve
112 Port navigation light
113 Radar warning aerial
114 Port wing reaction control air duct
115 Fuel pumps
116 Fuel system piping
117 Port wing leading-edge fence
118 Outboard pylon
119 BL755 cluster bombs (maximum load, seven)
120 Intermediate pylon
121 Port outrigger pylon
122 Missile launch rail
123 AIM-9L Sidewinder air-to-air missile
124 Port leading-edge root extension (LERX)
125 Inboard pylon
126 Hydraulic pumps
127 APU intake
128 Gas turbine starter/auxiliary power unit (APU)
129 Alternator cooling air exhaust
130 APU exhaust
131 Engine fuel control unit
132 Engine bay venting ram air intake
133 Rotary nozzle bearing
134 Nozzle fairing construction
135 Ammunition tank, 100 rounds
136 Cartridge case collector box
137 Ammunition feed chute
138 Fuel vent
139 Gun pack strake
140 Fuselage centreline pylon
141 Zero scarf forward (fan air) nozzle
142 Ventral gun pack (two)
143 Aden 25-mm cannon
144 Engine drain mast
145 Hydraulic system ground connectors
146 Forward fuselage flank fuel tank
147 Engine electronic control units
148 Engine accessory equipment gearbox
149 Gearbox driven alternator
150 Rolls-Royce Pegasus 11 Mk 105 vectored thrust turbofan
151 Formation lighting strips
152 Engine oil tank
153 Bleed air spill duct
154 Air conditioning intake scoops
155 Cockpit air conditioning system heat exchanger
156 Engine compressor/fan face
157 Heat exchanger discharge to intake duct
158 Nose undercarriage hydraulic retraction jack
159 Intake blow-in doors
160 Engine bay venting air scoop
161 Cannon muzzle fairing
162 Lift augmentation retractable cross-dam
163 Cross-dam hydraulic jack
164 Nosewheel
165 Nosewheel forks
166 Landing/taxiing lamp
167 Retractable boarding step
168 Nosewheel doors (closed after cycling of undercarriage)
169 Nosewheel door jack
170 Boundary layer bleed air duct
171 Nose undercarriage wheel bay
172 Kick-in boarding steps
173 Cockpit rear pressure bulkhead
174 Starboard side console panel
175 Martin-Baker Type 12 ejection seat
176 Safety harness
177 Ejection seat headrest
178 Port engine air intake
179 Probe hydraulic jack
180 Retractable in-flight refluling probe (bolt-on pack)
181 Cockpit canopy cover
182 Miniature detonating cord (MDC) canopy breaker
183 Canopy frame
184 Engine throttle and nozzle angle control levers
185 Pilot's head-up display
186 Instrument panel
187 Moving map display
188 Control column
189 Central warning system panel
190 Cockpit pressure floor
191 Underfloor control runs
192 Formation lighting strips
193 Aileron trim actuator
194 Rudder pedals
195 Cockpit section composite construction
196 Instrument panel shroud
197 One-piece wrap-around windscreen panel
198 Ram air intake (cockpit fresh air)
199 Front pressure bulkhead
200 Incidence vane
201 Air data computer
202 Pitot tube
203 Lower IFF aerial
204 Nose pitch control air valve
205 Pitch trim control actuator
206 Electrical system equipment
207 Yaw vane
208 Upper IFF aerial
209 Avionic equipment
210 ARBS heat exchanger
211 MIRLS sensors
212 Hughes Angle Rate Bombing System (ARBS)
213 Composite construction nose cone
214 ARBS glazed aperture

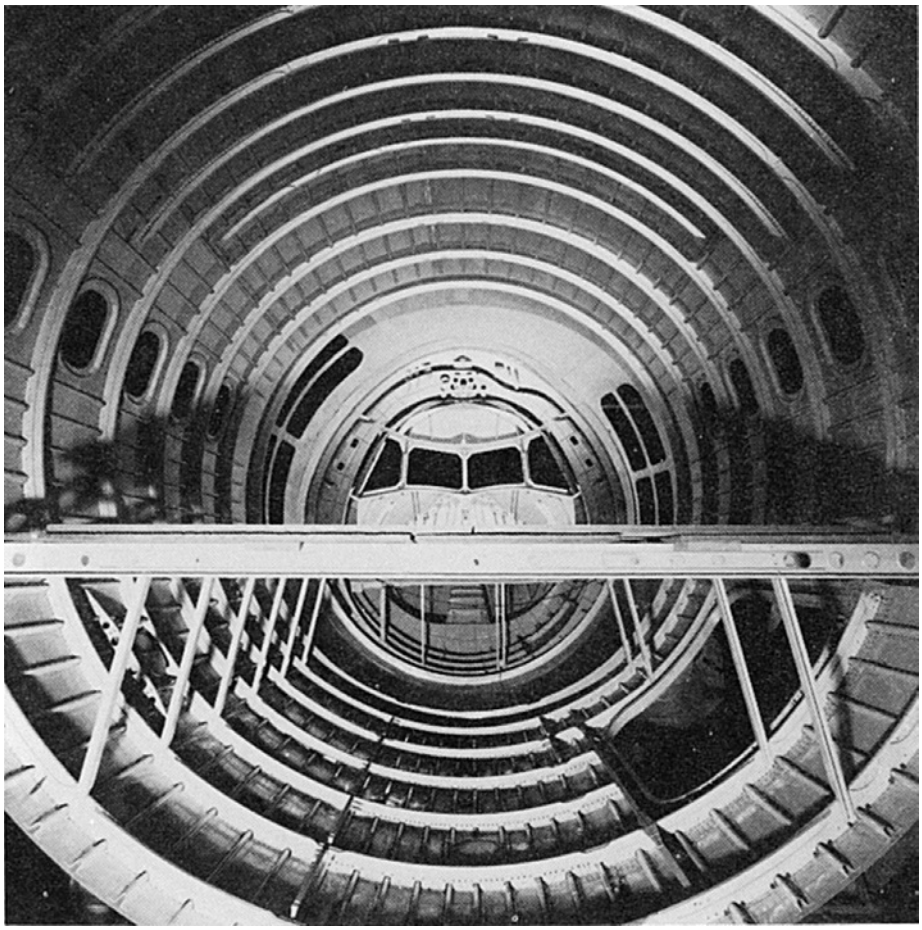


Figure 3.2-7 British Aerospace 146 (courtesy of British Aerospace).

which are therefore capable of supporting large compressive loads from axial and bending effects. In turn, spar webs exert a stabilizing influence on the skin in a similar manner to the stringers.

While the majority of the above remarks have been directed towards wing structures, they apply, as can be seen by referring to Figs 3.2-5 and 3.2-6, to all the aerodynamic surfaces, namely wings, horizontal and vertical tails, except in the obvious cases of undercarriage loading, engine thrust, etc.

Fuselages, while of different shape to the aerodynamic surfaces, comprise members which perform similar functions to their counterparts in the wings and tailplane. However, there are differences in the generation of the various types of load. Aerodynamic forces on the fuselage skin are relatively low; on the other hand, the fuselage supports large concentrated loads such as wing reactions, tailplane reactions, undercarriage reactions and it carries payloads of varying size and weight, which may cause large inertia forces. Furthermore, aircraft designed for high altitude flight must withstand internal pressure. The shape of the fuselage cross-section is determined by operational requirements. For example, the most efficient sectional shape for a pressurized fuselage is circular or a combination of circular elements. Irrespective of shape, the basic fuselage structure is essentially a single cell thin-walled tube comprising skin, transverse frames and stringers; transverse frames which extend completely across the fuselage are known as bulkheads. Three different types of fuselage are shown in Figs 3.2-5–3.2-7. In Fig. 3.2-5 the fuselage is unpressurized so that, in the passenger-carrying area, a more rectangular shape is employed to maximize space. The Harrier fuselage in Fig. 3.2-6 contains the engine, fuel tanks, etc. so that its cross-sectional shape is, to some extent, predetermined, while in Fig. 3.2-7 the passenger-carrying fuselage of the British Aerospace 146 is pressurized and therefore circular in cross-section.

3.2.3 Fabrication of Structural Components

The introduction of all-metal, stressed skin aircraft resulted in methods and types of fabrication which remain in use to the present day. However, improvements in engine performance and advances in aerodynamics have led to

higher maximum lift, higher speeds and therefore to higher wing loadings so that improved techniques of fabrication are necessary, particularly in the construction of wings. The increase in wing loading from about 350 N/m^2 for 1917–1918 aircraft to around 4800 N/m^2 for modern aircraft, coupled with a drop in the structural percentage of the total weight from 30–40 to 22–25 per cent, gives some indication of the improvements in materials and structural design.

For purposes of construction, aircraft are divided into a number of sub-assemblies. These are built in specially designed jigs, possibly in different parts of the factory or even different factories, before being forwarded to the final assembly shop. A typical breakdown into sub-assemblies of a medium-sized civil aircraft is shown in Fig. 3.2-8. Each sub-assembly relies on numerous minor assemblies such as spar webs, ribs, frames, and these, in turn, are supplied with individual components from the detail workshop.

Although the wings (and tailsurfaces) of fixed wing aircraft generally consist of spars, ribs, skin and stringers, methods of fabrication and assembly differ. The wing of the aircraft of Fig. 3.2-5 relies on fabrication techniques that have been employed for many years. In this form of construction the spars comprise thin aluminium alloy webs and flanges, the latter being extruded or machined and are bolted or riveted to the web. The ribs are formed in three parts from sheet metal by large presses and rubber dies and have flanges round their edges so that they can be riveted to the skin and spar webs; cut-outs around their edges allow the passage of spanwise stringers. Holes are cut in the ribs at positions of low stress for lightness and to accommodate control runs, fuel and electrical systems.

Finally, the skin is riveted to the rib flanges and longitudinal stiffeners. Where the curvature of the skin is large, for example at the leading edge, the aluminium alloy sheets are passed through 'rolls' to pre-form them to the correct shape. A further, aerodynamic, requirement is that forward chordwise sections of the wing should be as smooth as possible to delay transition from laminar to turbulent flow. Thus, countersunk rivets are used in these positions as opposed to dome-headed rivets nearer the trailing edge.

The wing is attached to the fuselage through reinforced fuselage frames, frequently by bolts. In some aircraft the wing spars are continuous through the fuselage depending on the demands of space. In a high wing aircraft (Fig. 3.2-5) deep spars passing through the fuselage would cause obstruction problems. In this case a short third spar provides an additional attachment point. The ideal arrangement is obviously where continuity of the structure is maintained over the entire surface of the wing. In most practical cases this is impossible since cut-outs in the wing surface are required for retracting undercarriages, bomb and gun bays, inspection panels, etc. The last are usually located on the undersurface of the wing and are fastened to stiffeners and rib flanges by screws, enabling them to resist direct and shear loads. Doors covering undercarriage wells and weapon bays are incapable of resisting wing stresses so that provision must be made for transferring the loads from skin, flanges and shear webs around the cut-out. This may be achieved by inserting strong bulkheads or increasing the spar flange areas, although, no matter the method employed, increased cost and weight result.

The different structural requirements of aircraft designed for differing operational roles lead to a variety of wing constructions. For instance, high-speed aircraft require relatively thin wing sections which support high wing

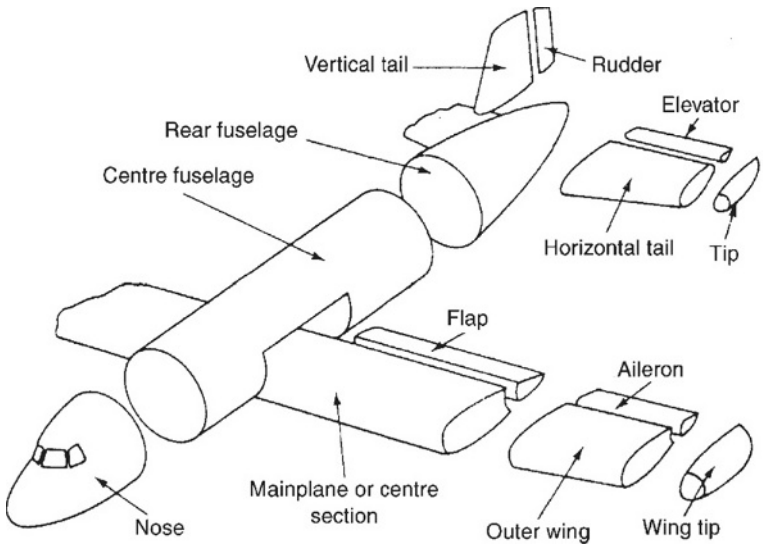


Figure 3.2-8 Typical sub-assembly breakdown.

loadings. To withstand the correspondingly high surface pressures and to obtain sufficient strength, much thicker skins are necessary. Wing panels are therefore frequently machined integrally with stringers from solid slabs of material, as are the wing ribs. Figure 3.2-9 shows wing ribs for the European Airbus in which web stiffeners, flanged lightness holes and skin attachment lugs have been integrally machined from solid. This integral method of construction involves no new design principles and has the advantages of combining a high grade of surface finish, free from irregularities, with a more efficient use of material since skin thicknesses are easily tapered to coincide with the spanwise decrease in bending stresses.

An alternative form of construction is the sandwich panel, which comprises a light honeycomb or corrugated metal core sandwiched between two outer skins of the stress-bearing sheet (see Fig. 3.2-10). The primary function of the core is to stabilize the outer skins, although it may be stress bearing as well. Sandwich panels are capable of developing high stresses, have smooth internal and external surfaces and require small numbers of supporting rings or frames. They also possess a high resistance to fatigue from jet efflux. The uses of this method of construction include lightweight 'planks' for cabin furniture, monolithic fairing shells generally having plastic facing skins, and the stiffening of flying control surfaces. Thus, for example, the ailerons and rudder of the British Aerospace Jaguar are fabricated from aluminium honeycomb, while fibreglass and aluminium faced honeycomb are used extensively in the wings and tail surfaces of the Boeing 747. Some problems, mainly disbonding and internal corrosion, have been encountered in service.

The general principles relating to wing construction are applicable to fuselages, with the exception that integral construction is not used in fuselages for obvious reasons. Figures 3.2-5, 3.2-6 and 3.2-7 show that the same basic method of construction is employed in aircraft having widely differing roles. Generally, the fuselage frames that support large concentrated floor loads or loads from wing or tailplane attachment points are heavier than lightly loaded frames and require stiffening, with additional provision for transmitting the concentrated load into the frame and hence the skin.

With the frames in position in the fuselage jig, stringers, passing through cut-outs, are riveted to the frame flanges. Before the skin is riveted to the frames and stringers, other subsidiary frames such as door and window frames are riveted or bolted in position. The areas of the fuselage in the regions of these cut-outs are reinforced by additional stringers, portions of frame and increased skin thickness, to react to the high shear flows and direct stresses developed.

On completion, the various sub-assemblies are brought together for final assembly. Fuselage sections are usually bolted together through flanges around their peripheries, while wings and the tailplane are attached to pick up points on the relevant fuselage frames. Wing spars on low wing civil aircraft usually pass completely through the fuselage, simplifying wing design and the method of attachment. On smaller, military aircraft, engine installations frequently

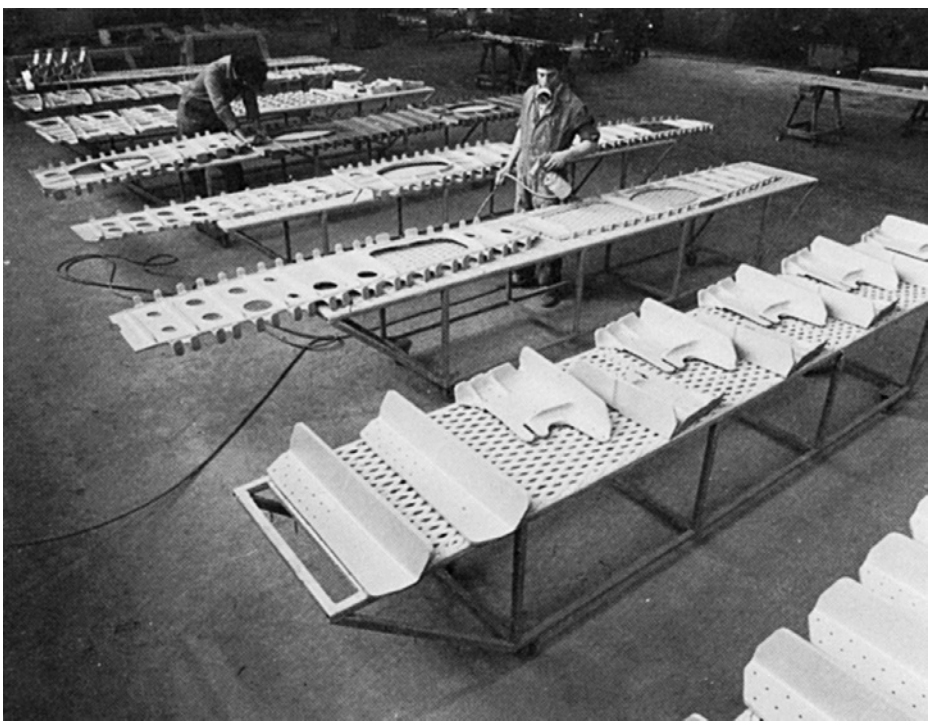


Figure 3.2-9 Wing ribs for the European Airbus (courtesy of British Aerospace).

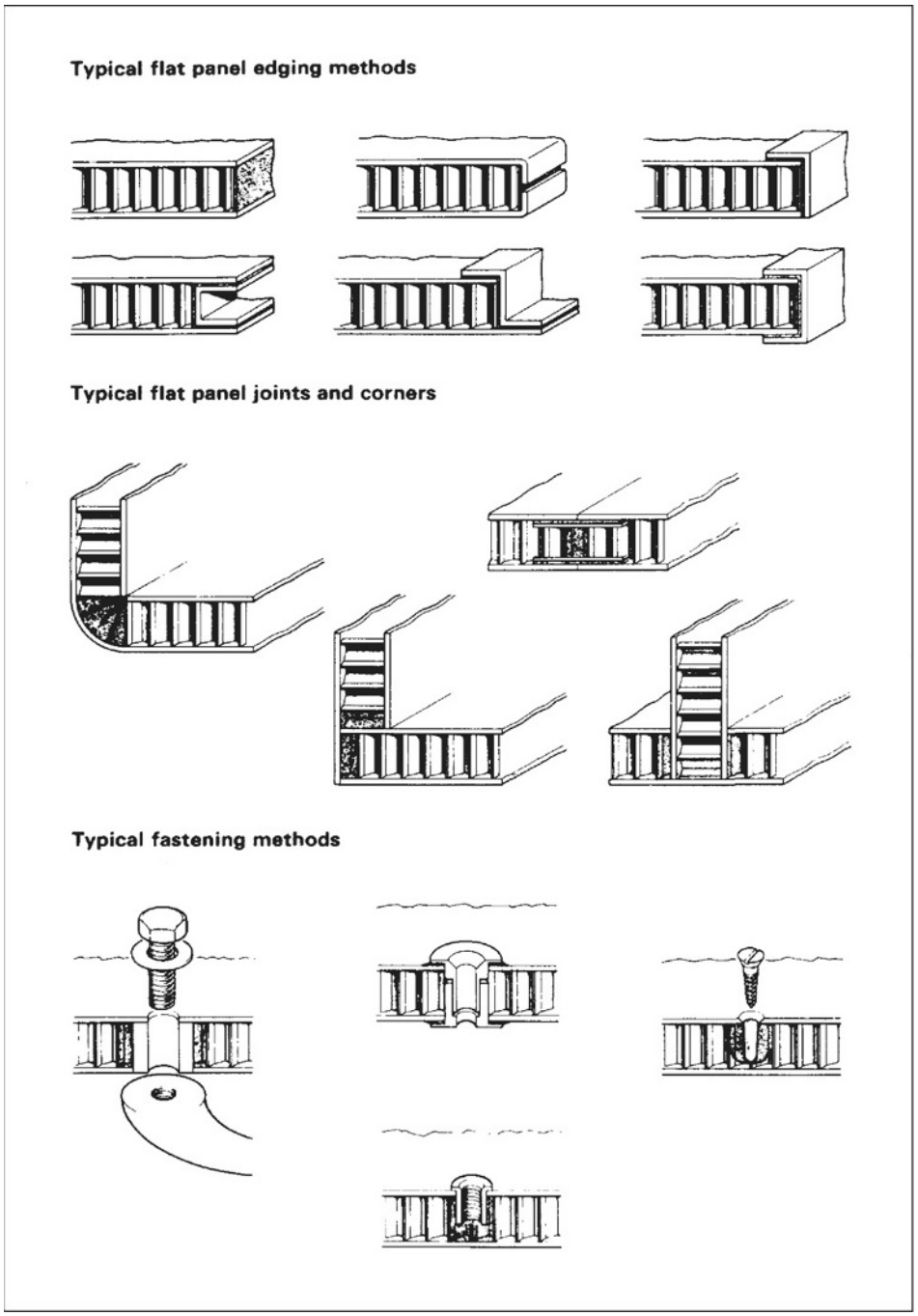


Figure 3.2-10 Sandwich panels (courtesy of Ciba-Geigy Plastics).

prevent this so that wing spars are attached directly to and terminate at the fuselage frame. Clearly, at these positions frame/stringer/skin structures require reinforcement.

3.2.4 Connections

The fabrication of aircraft components generally involves the joining of one part of the component to another. For example, fuselage skins are connected to stringers and frames while wing skins are connected to stringers and wing ribs unless, as in some military aircraft with high wing loadings, the stringers are machined integrally with the wing skin (see Section 3.2.3). With the advent of all-metal, i.e. aluminium alloy construction, riveted joints became the main form of connection with some welding although aluminium alloys are difficult to weld, and, in the modern era, some glued joints which use epoxy resin. In this section we shall concentrate on the still predominant method of connection, riveting.

In general riveted joints are stressed in complex ways and an accurate analysis is very often difficult to achieve because of the discontinuities in the region of the joint. Fairly crude assumptions as to joint behaviour are made but, when combined with experience, safe designs are produced.

3.2.4.1 Simple Lap Joint

Figure 3.2-11 shows two plates of thickness t connected together by a single line of rivets; this type of joint is termed a lap joint and is one of the simplest used in construction.

Suppose that the plates carry edge loads of P/unit width, that the rivets are of diameter d and are spaced at a distance b apart, and that the distance from the line of rivets to the edge of each plate is a. There are four possible modes of failure which must be considered as follows:

Rivet Shear

The rivets may fail by shear across their diameter at the interface of the plates. Then, if the maximum shear stress the rivets will withstand is τ_1 failure will occur when

$$Pb = \tau_1\left(\frac{\pi d^2}{4}\right)$$

which gives

$$P = \frac{\pi d^2 \tau_1}{4b} \tag{3.2.1}$$

Bearing pressure

Either the rivet or plate may fail due to bearing pressure. Suppose that p_b is this pressure then failure will occur when

$$\frac{Pb}{td} = p_b$$

so that

$$P = \frac{p_b t d}{b} \tag{3.2.2}$$

Plate failure in tension

The area of plate in tension along the line of rivets is reduced due to the presence of rivet holes. Therefore, if the ultimate tensile stress in the plate is σ_{ult} failure will occur when

$$\frac{P_b}{t(b-d)} = \sigma_{ult}$$

from which

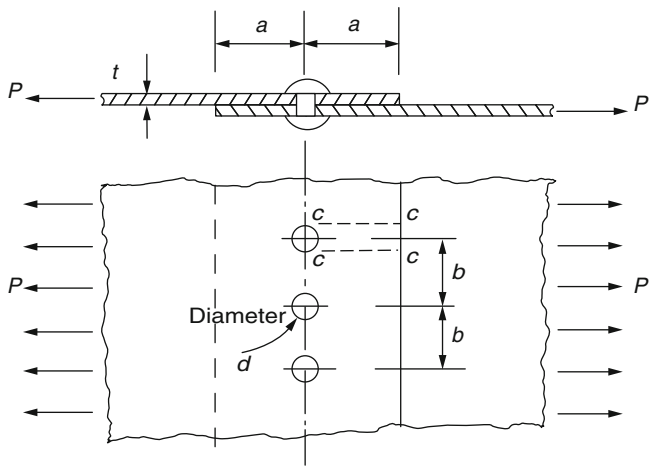


Figure 3.2-11 Simple riveted lap joint.

$$P = \frac{\sigma_{ult}\, t(b-d)}{b} \tag{3.2.3}$$

Shear failure in a plate

Shearing of the plates may occur on the planes cc resulting in the rivets being dragged out of the plate. If the maximum shear stress at failure of the material of the plates is τ_2 then a failure of this type will occur when

$$Pb = 2at\ \tau_2$$

which gives

$$P = \frac{2a\tau\ \tau_2}{b} \tag{3.2.4}$$

Example 3.2.1

A joint in a fuselage skin is constructed by riveting the abutting skins between two straps as shown in Fig. 3.2-12. The fuselage skins are 2.5 mm thick and the straps are each 1.2 mm thick; the rivets have a diameter of 4 mm. If the tensile stress in the fuselage skin must not exceed 125 N/mm^2 and the shear stress in the rivets is limited to 120N/mm^2 determine the maximum allowable rivet spacing such that the joint is equally strong in shear and tension.

A tensile failure in the plate will occur on the reduced plate cross-section along the rivet lines. This area is given by

$$A_p = (b-4) \times 2.5 \text{ mm}^2$$

The failure load/unit width P_f is then given by

$$P_f b = (b-4) \times 2.5 \times 125 \tag{i}$$

The area of cross-section of each rivet is

$$A_r = \frac{\pi \times 4^2}{4} = 12.6 \text{ mm}^2$$

Since each rivet is in double shear (i.e. two failure shear planes) the area of cross-section in shear is

$$2 \times 12.6 = 25.2 \text{ mm}^2$$

Then the failure load/unit width in shear is given by

$$P_f b = 25.2 \times 120 \tag{ii}$$

For failure to occur simultaneously in shear and tension, i.e. equating Eqs (i) and (ii)

$$25.2 \times 120 = (b-4) \times 2.5 \times 12.5$$

from which

$$b = 13.7 \text{ mm}$$

Say, a rivet spacing of 13 mm.

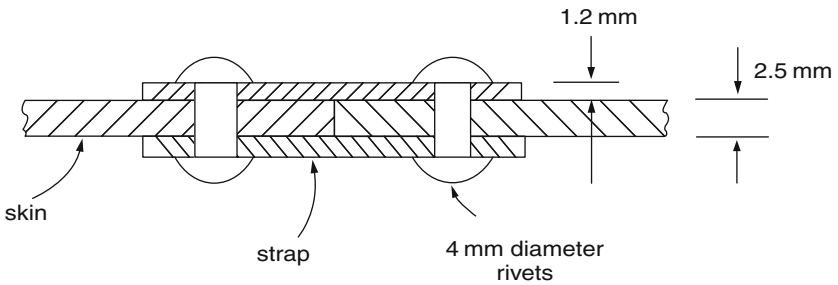


Figure 3.2-12 Joint of Example 3.2.1.

3.2.4.2 Joint Efficiency

The efficiency of a joint or connection is measured by comparing the actual failure load with that which would apply if there were no rivet holes in the plate. Then, for the joint shown in Fig. 3.2-11 the joint efficiency η is given by

$$\eta = \frac{\sigma_{ult} t(b-d)/b}{\sigma_{ult} t} = \frac{b-d}{b} \tag{3.2.5}$$

3.2.4.3 Group-Riveted Joints

Rivets may be grouped on each side of a joint such that the efficiency of the joint is a maximum. Suppose that two plates are connected as shown in Fig. 3.2-13 and that six rivets are required on each side. If it is assumed that each rivet is equally loaded then the single rivet on the line aa will take one-sixth of the total load. The two rivets on the line bb will then share two-sixths of the load while the three rivets on the line cc will share three-sixths of the load. On the line bb the area of cross-section of the plate is reduced by two rivet holes and that on the line cc by three rivet holes so that, relatively, the joint is as strong at these sections as at aa. Therefore, a more efficient joint is obtained than if the rivets were arranged in, say, two parallel rows of three.

3.2.4.4 Eccentrically Loaded Riveted Joints

The bracketed connection shown in Fig. 3.2-14 carries a load P offset from the centroid of the rivet group. The rivet group is then subjected to a shear load P through its centroid and a moment or torque Pe about its centroid.

It is assumed that the shear load P is distributed equally amongst the rivets causing a shear force in each rivet parallel to the line of action of P. The moment Pe is assumed to produce a shear force S in each rivet where S acts in a direction perpendicular to the line joining a particular rivet to the centroid of the rivet group. Furthermore, the value of S is assumed to be proportional to the distance of the rivet from the centroid of the rivet group. Then

$$Pe = \sum Sr$$

If $S = kr$ where k is a constant for all rivets then

$$Pe = k \sum r^2$$

from which

$$k = Pe / \sum r^2$$

and

$$S = \frac{Pe}{\sum r^2} r \tag{3.2.6}$$

The resultant force on a rivet is then the vector sum of the forces due to P and Pe.

Example 3.2.2

The bracket shown in Fig. 3.2-15 carries an offset load of 5 kN. Determine the resultant shear forces in the rivets A and B.

The vertical shear force on each rivet is $5/6 = 0.83$ kN. The moment (Pe) on the rivet group is $5 \times 75 = 375$ kN mm. The distance of rivet A (and B, G and H) from the centroid C of the rivet group is given by

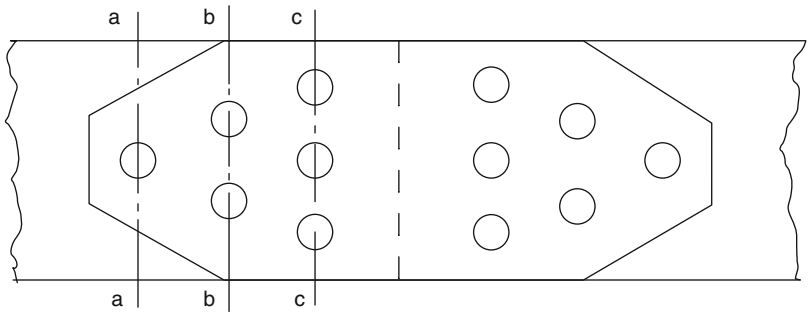


Figure 3.2-13 A group-riveted joint.

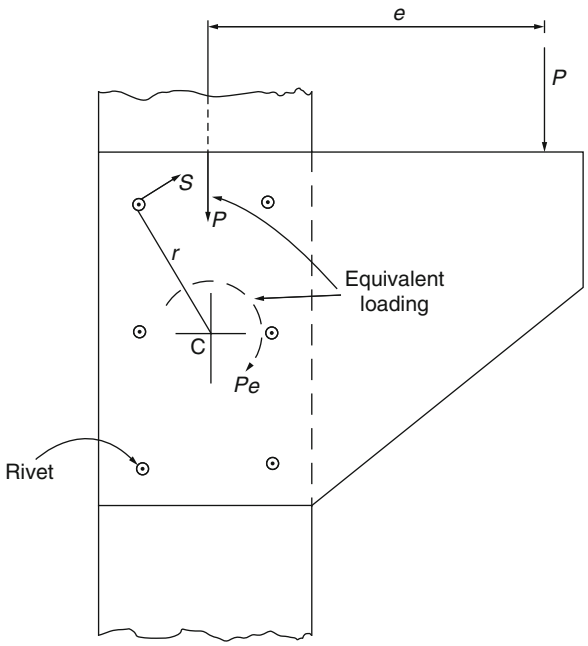


Figure 3.2-14 Eccentrically loaded joint.

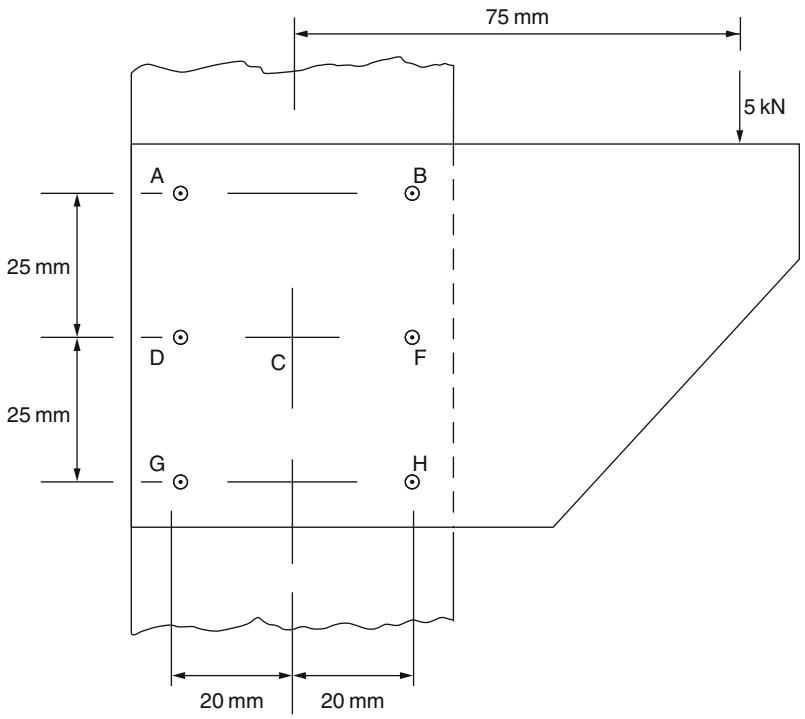


Figure 3.2-15 Joint of Example 3.2.2.

$$r = (20^2 + 25^2)^{1/2} = (1025)^{1/2} = 32.02\,\text{mm}$$

The distance of D (and F) from C is 20 mm. Therefore

$$\sum r^2 = 2 \times 400 + 4 \times 1025 = 4900$$

From Eq. (3.2.6) the shear forces on rivets A and B due to the moment are

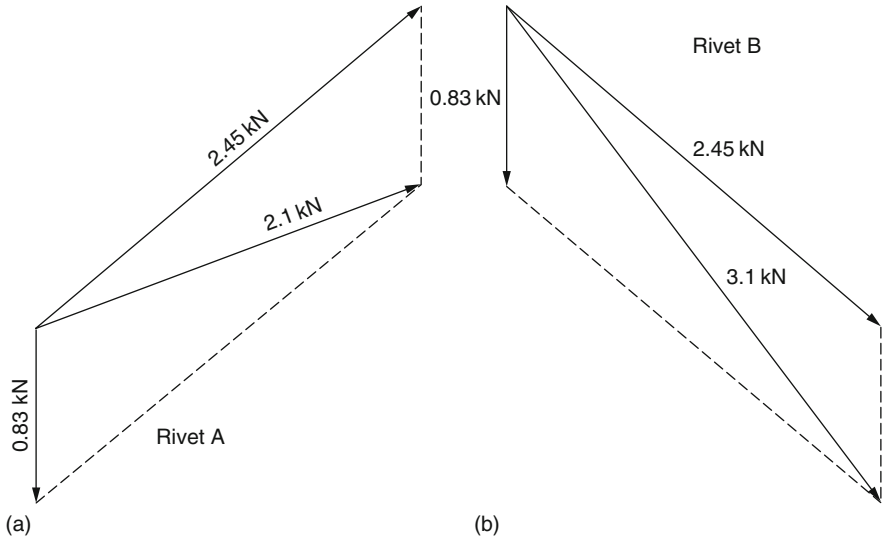


Figure 3.2-16 Force diagrams for rivets of Example 3.2.2.

$$S = \frac{375}{4900} \times 32.02 = 2.45\,\text{kN}$$

On rivet A the force system due to P and Pe is that shown in Fig. 3.2-16(a) while that on B is shown in Fig. 3.2-16(b).

The resultant forces may then be calculated using the rules of vector addition or determined graphically using the parallelogram of forces.[1]

The design of riveted connections is carried out in the actual design of the rear fuselage of a single-engined trainer/semi-aerobatic aircraft.

3.2.4.5 Use of Adhesives

In addition to riveted connections adhesives have and are being used in aircraft construction although, generally, they are employed in areas of low stress since their application is still a matter of research. Of these adhesives epoxy resins are the most frequently used since they have the advantages over, say polyester resins, of good adhesive properties, low shrinkage during cure so that residual stresses are reduced, good mechanical properties and thermal stability. The modulus and ultimate strength of epoxy resin are, typically, 5000 and 100 N/mm^2. Epoxy resins are now found extensively as the matrix component in fibrous composites.

Reference

1. Megson, T. H. G., *Structural and Stress Analysis*, 2nd edition. Elsevier, Oxford, 2005.

Problems

P.3.2.1 Examine possible uses of new materials in future aircraft manufacture.

P.3.2.2 Describe the main features of a stressed skin structure. Discuss the structural functions of the various components with particular reference either to the fuselage or to the wing of a medium-sized transport aircraft.

P.3.2.3 The double riveted butt joint shown in Fig. P.3.2-3 connects two plates which are each 2.5 mm thick, the rivets have a diameter of 3 mm. If the failure strength of the rivets in shear is 370 N/mm^2 and the ultimate tensile strength of the plate is 465 N/mm^2 determine the necessary rivet pitch if the joint is to be designed so that failure due to shear in the rivets and failure due to tension in the plate occur simultaneously. Calculate also the joint efficiency.

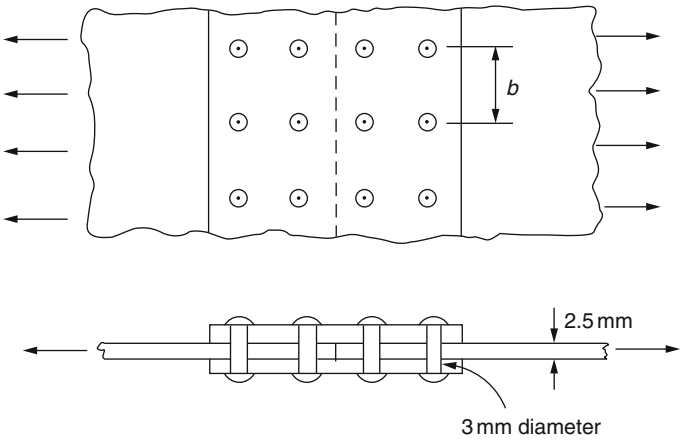


Figure P.3.2-3

Ans. Rivet pitch is 12 mm, joint efficiency is 75 per cent.

P.3.2.4 The rivet group shown in Fig. P.3.2-4 connects two narrow lengths of plate one of which carries a 15 kN load positioned as shown. If the ultimate shear strength of a rivet is 350 N/mm^2 and its failure strength in compression is 600 N/mm^2 determine the minimum allowable values of rivet diameter and plate thickness.

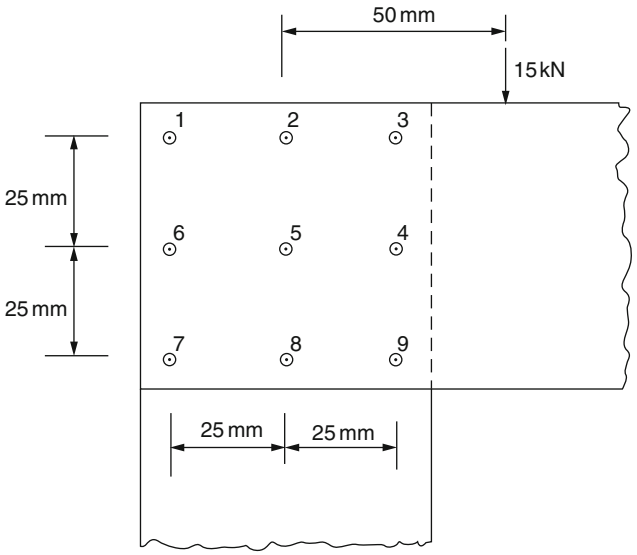


Figure P.3.2-4

Ans. Rivet diameter is 4.0 mm, plate thickness is 1.83 mm.

Aerospace Engineering Desk Reference; ISBN 9781856175753

3.3 Airworthiness

T.H.G. Megson

The airworthiness of an aircraft is concerned with the standards of safety incorporated in all aspects of its construction. These range from structural strength to the provision of certain safeguards in the event of crash landings, and include design requirements relating to aerodynamics, performance and electrical and hydraulic systems. The selection of minimum standards of safety is largely the concern of 'national and international' airworthiness authorities who prepare handbooks of official requirements. The handbooks include operational requirements, minimum safety requirements, recommended practices and design data, etc.

In this chapter we shall concentrate on the structural aspects of airworthiness which depend chiefly on the strength and stiffness of the aircraft. Stiffness problems may be conveniently grouped under the heading *aeroelasticity*. Strength problems arise, as we have seen, from ground and air loads, and their magnitudes depend on the selection of manoeuvring and other conditions applicable to the operational requirements of a particular aircraft.

3.3.1 Factors of Safety-Flight Envelope

The control of weight in aircraft design is of extreme importance. Increases in weight require stronger structures to support them, which in turn lead to further increases in weight and so on. Excesses of structural weight mean lesser amounts of payload, thereby affecting the economic viability of the aircraft. The aircraft designer is therefore constantly seeking to pare his aircraft's weight to the minimum compatible with safety. However, to ensure general minimum standards of strength and safety, airworthiness regulations lay down several factors which the primary structure of the aircraft must satisfy. These are the *limit load*, which is the maximum load that the aircraft is expected to experience in normal operation, the *proof load*, which is the product of the limit load and the *proof factor* (1.0–1.25), and the *ultimate load*, which is the product of the limit load and the *ultimate factor* (usually 1.5). The aircraft's structure must withstand the proof load without detrimental distortion and should not fail until the ultimate load has been achieved. The proof and ultimate factors may be regarded as factors of safety and provide for various contingencies and uncertainties which are discussed in greater detail in Section 3.3.2.

The basic strength and flight performance limits for a particular aircraft are selected by the airworthiness authorities and are contained in the *flight envelope* or V—n diagram shown in Fig. 3.3-1. The curves OA and OF correspond to the stalled condition of the aircraft and are obtained from the well-known aerodynamic relationship

$$\text{Lift} = nW = \frac{1}{2}\rho V^2 S C_{\text{L,max}}$$

Therefore, for speeds below V_A (positive wing incidence) and V_F(negative incidence) the maximum loads which can be applied to the aircraft are governed by $C_{\text{L,max}}$. As the speed increases it is possible to apply the positive and negative limit loads, corresponding to n_1 and n_3, without stalling the aircraft so that AC and FE represent maximum operational load factors for the aircraft. Above the design cruising speed V_C, the cut-off lines CD_1 and D_2E relieve the design cases to be covered since it is not expected that the limit loads will be applied at maximum speed. Values of n_1, n_2 and n_3 are specified by the airworthiness authorities for particular aircraft; typical load factors are shown in Table 3.3-1.

A particular flight envelope is applicable to one altitude only since $C_{\text{L,max}}$ is generally reduced with an increase of altitude, and the speed of sound decreases with altitude thereby reducing the critical Mach number and hence the design diving speed V_D. Flight envelopes are therefore drawn for a range of altitudes from sea level to the operational ceiling of the aircraft.

3.3.2 Load Factor Determination

Several problems require solution before values for the various load factors in the flight envelope can be determined. The limit load, for example, may be produced by a specified manoeuvre or by an encounter with a particularly severe gust (gust cases and the associated gust envelope are discussed in Section 3.4.4). Clearly some knowledge of possible gust conditions is required to determine the limiting case. Furthermore, the fixing of the proof and ultimate factors

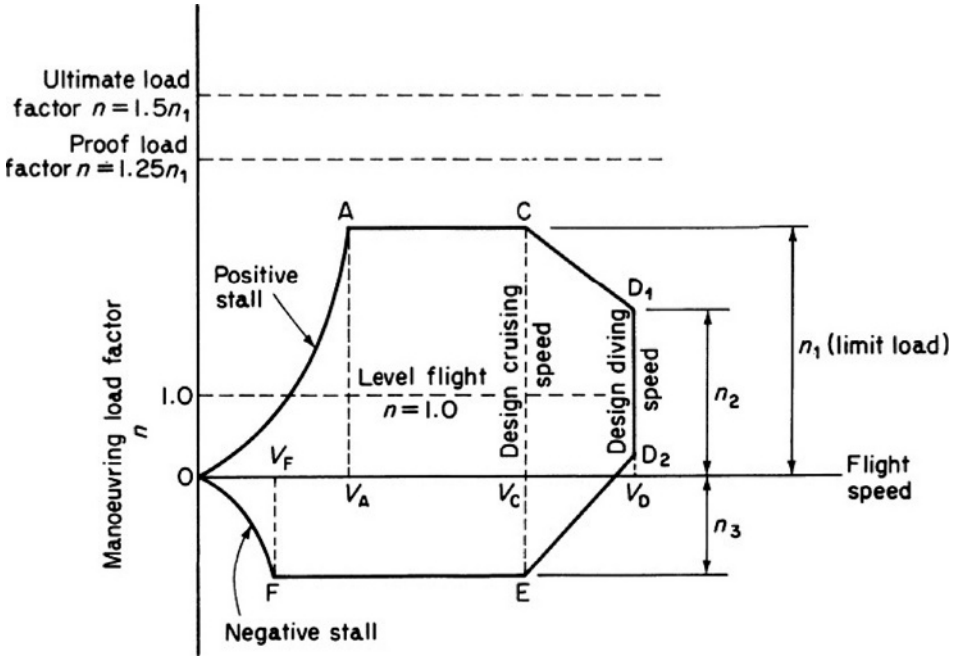


Figure 3.3-1 Flight envelope

Table 3.3-1

	Category		
Load factor n	Normal	Semi-aerobatic	Aerobatic
n_1	$2.1 + 24\,000/(W + 10\,000)$	4.5	6.0
n_2	$0.75n_1$ but $n_2 \not< 2.0$	3.1	4.5
n_3	1.0	1.8	3.0

also depends upon the degree of uncertainty of design, variations in structural strength, structural deterioration, etc. We shall now investigate some of these problems to see their comparative influence on load factor values.

3.3.2.1 Limit Load

An aircraft is subjected to a variety of loads during its operational life, the main classes of which are: manoeuvre loads, gust loads, undercarriage loads, cabin pressure loads, buffeting and induced vibrations. Of these, manoeuvre, undercarriage and cabin pressure loads are determined with reasonable simplicity since manoeuvre loads are controlled design cases, undercarriages are designed for given maximum descent rates and cabin pressures are specified. The remaining loads depend to a large extent on the atmospheric conditions encountered during flight. Estimates of the magnitudes of such loads are only possible therefore if in-flight data on these loads is available. It obviously requires a great number of hours of flying if the experimental data are to include possible extremes of atmospheric conditions. In practice, the amount of data required to establish the probable period of flight time before an aircraft encounters, say, a gust load of a given severity, is a great deal more than that available. It therefore becomes a problem in statistics to extrapolate the available data and calculate the probability of an aircraft being subjected to its proof or ultimate load during its operational life. The aim would be for a zero or negligible rate of occurrence of its ultimate load and an extremely low rate of occurrence of its proof load. Having decided on an ultimate load, then the limit load may be fixed as defined in Section 3.3.1 although the value of the ultimate factor includes, as we have already noted, allowances for uncertainties in design, variation in structural strength and structural deterioration.

3.3.2.2 Uncertainties in Design and Structural Deterioration

Neither of these present serious problems in modern aircraft construction and therefore do not require large factors of safety to minimize their effects. Modern methods of aircraft structural analysis are refined and, in any case, tests to

determine actual failure loads are carried out on representative full scale components to verify design estimates. The problem of structural deterioration due to corrosion and wear may be largely eliminated by close inspection during service and the application of suitable protective treatments.

3.3.2.3 Variation in Structural Strength

To minimize the effect of the variation in structural strength between two apparently identical components, strict controls are employed in the manufacture of materials and in the fabrication of the structure. Material control involves the observance of strict limits in chemical composition and close supervision of manufacturing methods such as machining, heat treatment, rolling, etc. In addition, the inspection of samples by visual, radiographic and other means, and the carrying out of strength tests on specimens, enable below limit batches to be isolated and rejected. Thus, if a sample of a batch of material falls below a specified minimum strength then the batch is rejected. This means of course that an actual structure always comprises materials with properties equal to or better than those assumed for design purposes, an added but unallowed for 'bonus' in considering factors of safety.

Similar precautions are applied to assembled structures with regard to dimension tolerances, quality of assembly, welding, etc. Again, visual and other inspection methods are employed and, in certain cases, strength tests are carried out on sample structures.

3.3.2.4 Fatigue

Although adequate precautions are taken to ensure that an aircraft's structure possesses sufficient strength to withstand the most severe expected gust or manoeuvre load, there still remains the problem of fatigue. Practically, all components of the aircraft's structure are subjected to fluctuating loads which occur a great many times during the life of the aircraft. It has been known for many years that materials fail under fluctuating loads at much lower values of stress than their normal static failure stress. A graph of failure stress against number of repetitions of this stress has the typical form shown in Fig. 3.3-2. For some materials, such as mild steel, the curve (usually known as an *S–N* curve or diagram) is asymptotic to a certain minimum value, which means that the material has an actual infinite-life stress. Curves for other materials, for example aluminium and its alloys, do not always appear to have asymptotic values so that these materials may not possess an infinite-life stress. We shall discuss the implications of this a little later.

Prior to the mid-1940s little attention had been paid to fatigue considerations in the design of aircraft structures. It was felt that sufficient static strength would eliminate the possibility of fatigue failure. However, evidence began to accumulate that several aircraft crashes had been caused by fatigue failure. The seriousness of the situation was highlighted in the early 1950s by catastrophic fatigue failures of two Comet airliners. These were caused by the once-per-flight cabin pressurization cycle which produced circumferential and longitudinal stresses in the fuselage skin. Although these stresses were well below the allowable stresses for single cycle loading, stress concentrations occurred at the corners of the windows and around rivets which raised local stresses considerably above the general stress level. Repeated cycles of pressurization produced fatigue cracks which propagated disastrously, causing an explosion of the fuselage at high altitude.

Several factors contributed to the emergence of fatigue as a major factor in design. For example, aircraft speeds and sizes increased, calling for higher wing and other loadings. Consequently, the effect of turbulence was magnified and

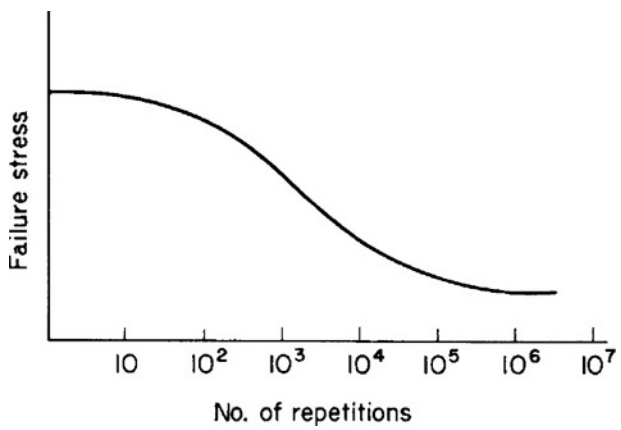


Figure 3.3-2 Typical form of *S–N* diagram.

the magnitudes of the fluctuating loads became larger. In civil aviation, airliners had a greater utilization and a longer operational life. The new 'zinc-rich' alloys, used for their high static strength properties, did not show a proportional improvement in fatigue strength, exhibited high crack propagation rates and were extremely notch sensitive.

Despite the fact that the causes of fatigue were reasonably clear at that time its elimination as a threat to aircraft safety was a different matter. The fatigue problem has two major facets: the prediction of the fatigue strength of a structure and a knowledge of the loads causing fatigue. Information was lacking on both counts. The Royal Aircraft Establishment (RAE) and the aircraft industry therefore embarked on an extensive test programme to determine the behaviour of complete components, joints and other detail parts under fluctuating loads. These included fatigue testing by the RAE of some 50 Meteor 4 tailplanes at a range of temperatures, plus research, also by the RAE, into the fatigue behaviour of joints and connections. Further work was undertaken by some universities and by the industry itself into the effects of stress concentrations.

In conjunction with their fatigue strength testing, the RAE initiated research to develop a suitable instrument for counting and recording gust loads over long periods of time. Such an instrument was developed by J. Taylor in 1950 and was designed so that the response fell off rapidly above 10 Hz. Crossings of g thresholds from 0.2 to 1.8 g at 0.1 g intervals were recorded (note that steady level flight is 1 g flight) during experimental flying at the RAE on three different aircraft over 28 000 km, and the best techniques for extracting information from the data established. Civil airlines cooperated by carrying the instruments on their regular air services for a number of years. Eight different types of aircraft were equipped so that by 1961 records had been obtained for regions including Europe, the Atlantic, Africa, India and the Far East, representing 19 000 hours and 8 million km of flying.

Atmospheric turbulence and the cabin pressurization cycle are only two of the many fluctuating loads which cause fatigue damage in aircraft. On the ground the wing is supported on the undercarriage and experiences tensile stresses in its upper surfaces and compressive stresses in its lower surfaces. In flight these stresses are reversed as aerodynamic lift supports the wing. Also, the impact of landing and ground manoeuvring on imperfect surfaces causes stress fluctuations while, during landing and take-off, flaps are lowered and raised, producing additional load cycles in the flap support structure. Engine pylons are subjected to fatigue loading from thrust variations in take-off and landing and also to inertia loads produced by lateral gusts on the complete aircraft.

A more detailed investigation of fatigue and its associated problems is presented elsewhere whilst a fuller discussion of airworthiness as applied to civil jet aircraft is presented in Ref. [1].

Reference

1. Jenkinson, L.R., Simpkin, P. and Rhodes, D., *Civil Jet Aircraft Design*, Arnold, London, 1999.

Aerospace Engineering Desk Reference; ISBN 9781856175753

3.4 Airframe Loads

T.H.G. Megson

In Chapter 3.2, we discussed in general terms the types of load to which aircraft are subjected during their operational life. We shall now examine in more detail the loads which are produced by various manoeuvres and the manner in which they are calculated.

3.4.1 Aircraft Inertia Loads

The maximum loads on the components of an aircraft's structure generally occur when the aircraft is undergoing some form of acceleration or deceleration, such as in landings, take-offs and manoeuvres within the flight and gust envelopes. Thus, before a structural component can be designed, the inertia loads corresponding to these accelerations and decelerations must be calculated. For these purposes we shall suppose that an aircraft is a rigid body and represent it by a rigid mass, m, as shown in Fig. 3.4-1. We shall also, at this stage, consider motion in the plane of the mass which would correspond to pitching of the aircraft without roll or yaw. We shall also suppose that the centre of gravity (CG) of the mass has coordinates $\bar{x}, \bar{y}$ referred to x and y axes having an arbitrary origin O; the mass is rotating about an axis through O perpendicular to the xy plane with a constant angular velocity ω.

The acceleration of any point, a distance r from O, is $\omega^2 r$ and is directed towards O. Thus, the inertia force acting on the element, δm, is $\omega^2 r \delta m$ in a direction opposite to the acceleration, as shown in Fig. 3.4-1. The components of this inertia force, parallel to the x and y axes, are $\omega^2 r \delta m \cos\theta$ and $\omega^2 r \delta m \sin\theta$, respectively, or, in terms of x and y, $\omega^2 x \delta m$ and $\omega^2 y \delta m$. The resultant inertia forces, F_x and F_y, are then given by

$$F_x = \int \omega^2 x \, \mathrm{d}m = \omega^2 \int x \, \mathrm{d}m$$

$$F_y = \int \omega^2 y \, \mathrm{d}m = \omega^2 \int y \, \mathrm{d}m$$

in which we note that the angular velocity ω is constant and may therefore be taken outside the integral sign. In the above expressions $\int x \, \mathrm{d}m$ and $\int y \, \mathrm{d}m$ are the moments of the mass, m, about the y and x axes, respectively, so that

$$F_x = \omega^2 \bar{x} m \tag{3.4.1}$$

and

$$F_y = \omega^2 \bar{y} m \tag{3.4.2}$$

If the CG lies on the x axis, $\bar{y} = 0$ and $F_y = 0$. Similarly, if the CG lies on the y axis, $F_x = 0$. Clearly, if O coincides with the CG, $\bar{x} = \bar{y} = 0$ and $F_x = F_y = 0$.

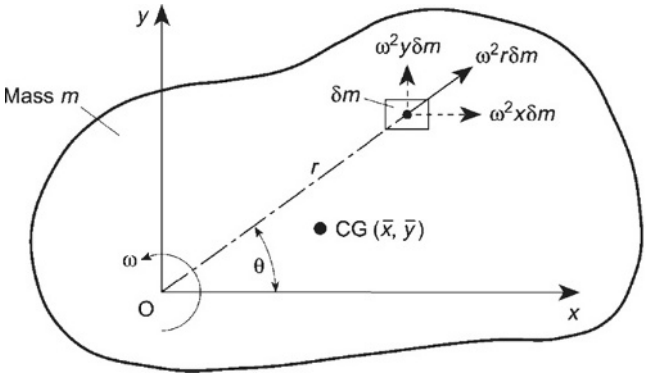


Figure 3.4-1 Inertia forces on a rigid mass having a constant angular velocity.

Suppose now that the rigid body is subjected to an angular acceleration (or deceleration) α in addition to the constant angular velocity, ω, as shown in Fig. 3.4-2. An additional inertia force, $\alpha r\delta m$, acts on the element δm in a direction perpendicular to r and in the opposite sense to the angular acceleration. This inertia force has components $\alpha r\delta m \cos\theta$ and $\alpha r\delta m \sin\theta$, i.e. $\alpha x\delta m$ and $\alpha y\delta m$, in the y and x directions, respectively. Thus, the resultant inertia forces, F_x and F_y, are given by

$$F_x = \int \alpha y \, \mathrm{d}m = \alpha \int y \, \mathrm{d}m$$

and

$$F_y = -\int \alpha x \, \mathrm{d}m = -\alpha \int x \, \mathrm{d}m$$

for α in the direction shown. Then, as before

$$F_x = \alpha \bar{y} m \tag{3.4.3}$$

and

$$F_y = \alpha \bar{x} m \tag{3.4.4}$$

Also, if the CG lies on the x axis, $\bar{y} = 0$ and $F_x = 0$ Similarly, if the CG lies on the y axis, $\bar{x} = 0$ and $F_y = 0$.

The torque about the axis of rotation produced by the inertia force corresponding to the angular acceleration on the element δm is given by

$$\delta T_{\mathrm{O}} = \alpha r^2 \delta m$$

Thus, for the complete mass

$$T_{\mathrm{O}} = \int \alpha r^2 \, \mathrm{d}m = \alpha \int r^2 \mathrm{d}m$$

The integral term in this expression is the moment of inertia, $\mathrm{I_O}$, of the mass about the axis of rotation. Thus

$$T_{\mathrm{O}} = \alpha I_{\mathrm{O}} \tag{3.4.5}$$

Equation (3.4.5) may be rewritten in terms of I_{CG}, the moment of inertia of the mass about an axis perpendicular to the plane of the mass through the CG. Hence, using the parallel axes theorem

$$I_{\mathrm{O}} = m(\bar{r})^2 + I_{\mathrm{CG}}$$

where $\bar{r}$ is the distance between O and the CG. Then

$$I_{\mathrm{O}} = m[(\bar{x})^2 + (\bar{y})^2] + I_{\mathrm{CG}}$$

and

$$T_{\mathrm{O}} = m[(\bar{x})^2 + (\bar{y})^2]\alpha + I_{\mathrm{CG}}\alpha \tag{3.4.6}$$

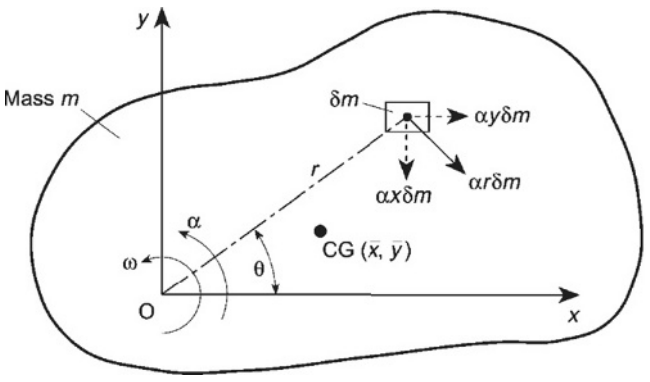


Figure 3.4-2 Inertia forces on a rigid mass subjected to an angular acceleration.

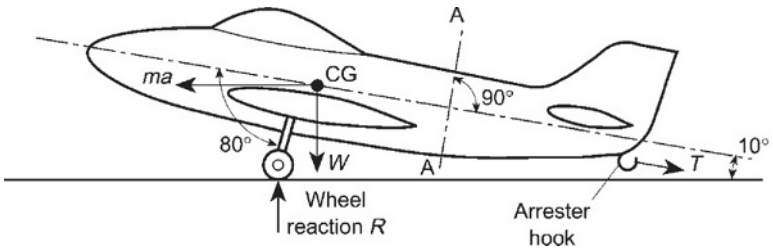


Figure 3.4-3 Forces on the aircraft of Example 3.4.1.

Example 3.4.1

An aircraft having a total weight of 45 kN lands on the deck of an aircraft carrier and is brought to rest by means of a cable engaged by an arrester hook, as shown in Fig. 3.4-3. If the deceleration induced by the cable is $3\,g$ determine the tension, T, in the cable, the load on an undercarriage strut and the shear and axial loads in the fuselage at the section AA; the weight of the aircraft aft of AA is 4.5 kN. Calculate also the length of deck covered by the aircraft before it is brought to rest if the touch-down speed is 25 m/s.

The aircraft is subjected to a horizontal inertia force ma where m is the mass of the aircraft and a its deceleration. Thus, resolving forces horizontally

$$T\cos 10^\circ - ma = 0$$

i.e.

$$T\cos 10^\circ - \frac{45}{g}3g = 0$$

which gives

$$T = 137.1\ \text{kN}$$

Now resolving forces vertically

$$R - W - T\sin 10^\circ = 0$$

i.e.

$$R = 45 + 131.1\sin 10^\circ = 68.8\ \text{kN}$$

Assuming two undercarriage struts, the load in each strut will be $(R/2)/\cos 20^\circ = 36.6$ kN.

Let N and S be the axial and shear loads at the section AA, as shown in Fig. 3.4-4. The inertia load acting at the CG of the fuselage aft of AA is $m_1\,a$, where m_1 is the mass of the fuselage aft of AA. Then

$$m_1 a = \frac{4.5}{g}3\,g = 13.5\ \text{kN}$$

Resolving forces parallel to the axis of the fuselage

$$N - T + m_1 a\cos 10^\circ - 4.5\sin 10^\circ = 0$$

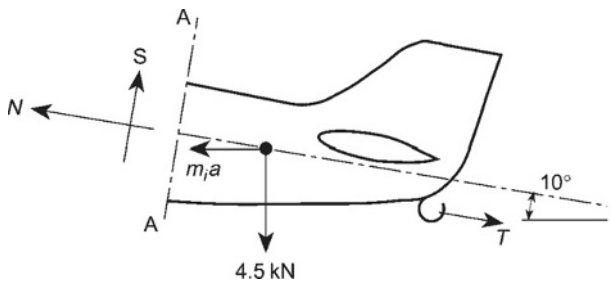


Figure 3.4-4 Shear and axial loads at the section AA of the aircraft of Example 3.4.1.

i.e.

$$N - 137.1 + 13.5 \cos 10^\circ - 4.5 \sin 10^\circ = 0$$

whence

$$N = 124.6 \text{ kN}$$

Now resolving forces perpendicular to the axis of the fuselage

$$S - m_1 a \sin 10^\circ - 4.5 \cos 10^\circ = 0$$

i.e.

$$S - 13.5 \sin 10^\circ - 4.5 \cos 10^\circ = 0$$

so that

$$S = 6.8 \text{ kN}$$

Note that, in addition to the axial load and shear load at the section AA, there will also be a bending moment.

Finally, from elementary dynamics

$$v^2 = v_0^2 + 2as$$

where v_0 is the touchdown speed, v the final speed (=0) and s the length of deck covered. Then

$$v_0^2 = -2as$$

i.e.

$$25^2 = -2(-3 \times 9.81)s$$

which gives

$$s = 10.6 \text{ m}$$

Example 3.4.2

An aircraft having a weight of 250 kN and a tricycle undercarriage lands at a vertical velocity of 3.7 m/s, such that the vertical and horizontal reactions on the main wheels are 1200 kN and 400 kN respectively; at this instant the nose wheel is 1.0 m from the ground, as shown in Fig. 3.4-5. If the moment of inertia of the aircraft about its CG is 5.65×10^8 N s^2 mm determine the inertia forces on the aircraft, the time taken for its vertical velocity to become zero and its angular velocity at this instant.

The horizontal and vertical inertia forces ma_x and ma_y act at the CG, as shown in Fig. 3.4-5, m is the mass of the aircraft, and a_x and a_y its accelerations in the horizontal and vertical directions, respectively. Then, resolving forces horizontally

$$ma_x - 400 = 0$$

whence

$$ma_x = 400 \text{ kN}$$

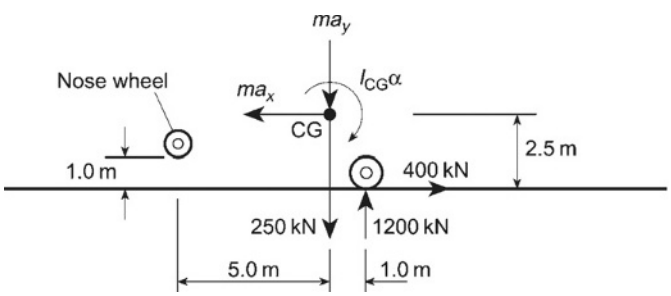


Figure 3.4-5 Geometry of the aircraft of Example 3.4.2.

Now resolving forces vertically

$$ma_y + 250 - 1200 = 0$$

which gives

$$ma_y = 950 \text{ kN}$$

Then

$$a_y = \frac{950}{m} = \frac{950}{250/g} = 3.8g \tag{i}$$

Now taking moments about the CG

$$I_{CG}\alpha - 1200 \times 1.0 - 400 \times 2.5 = 0 \tag{ii}$$

from which

$$I_{CG}\alpha = 2200\, m \text{ kN}$$

Hence

$$\alpha = \frac{I_{CG}\alpha}{I_{CG}} = \frac{2200 \times 10^6}{5.65 \times 10^8} = 3.9 \text{ rad/s}^2 \tag{iii}$$

From Eq. (i), the aircraft has a vertical deceleration of 3.8 g from an initial vertical velocity of 3.7 m/s. Therefore, from elementary dynamics, the time, t, taken for the vertical velocity to become zero, is given by

$$v = v_0 + a_y t \tag{iv}$$

in which $v = 0$ and $v_0 = 3.7$ m/s. Hence

$$0 = 3.7 - 3.8 \times 9.81t$$

whence

$$t = 0.099 \text{ s}$$

In a similar manner to Eq. (iv) the angular velocity of the aircraft after 0.099 s is given by

$$\omega = \omega_0 + \alpha t$$

in which $\omega_0 = 0$ and $\alpha = 3.9$ rad/s^2. Hence

$$\omega = 3.9 \times 0.099$$

i.e.

$$\omega = 0.39 \text{ rad/s}$$

3.4.2 Symmetric Manoeuvre Loads

We shall now consider the calculation of aircraft loads corresponding to the flight conditions specified by flight envelopes. There are, in fact, an infinite number of flight conditions within the boundary of the flight envelope although, structurally, those represented by the boundary are the most severe. Furthermore, it is usually found that the corners A, C, D_1, D_2, E and F (see Fig. 3.3-1) are more critical than points on the boundary between the corners so that, in practice, only the six conditions corresponding to these corner points need be investigated for each flight envelope.

In symmetric manoeuvres we consider the motion of the aircraft initiated by movement of the control surfaces in the plane of symmetry. Examples of such manoeuvres are loops, straight pull-outs and bunts, and the calculations involve the determination of lift, drag and tailplane loads at given flight speeds and altitudes. The effects of atmospheric turbulence and gusts are discussed in Section 3.4.4.

3.4.2.1 Level Flight

Although steady level flight is not a manoeuvre in the strict sense of the word, it is a useful condition to investigate initially since it establishes points of load application and gives some idea of the equilibrium of an aircraft in the longitudinal plane. The loads acting on an aircraft in steady flight are shown in Fig. 3.4-6, with the following notation:

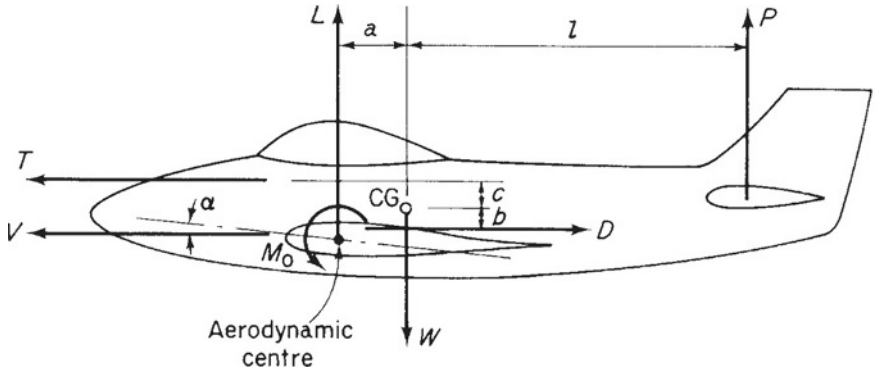


Figure 3.4-6 Aircraft loads in level flight.

L is the lift acting at the aerodynamic centre of the wing.
D is the aircraft drag.
M_0 is the aerodynamic pitching moment of the aircraft *less* its horizontal tail.
P is the horizontal tail load acting at the aerodynamic centre of the tail, usually taken to be at approximately one-third of the tailplane chord.
W is the aircraft weight acting at its CG.
T is the engine thrust, assumed here to act parallel to the direction of flight in order to simplify calculation.

The loads are in static equilibrium since the aircraft is in a steady, unaccelerated, level flight condition. Thus for vertical equilibrium

$$L + P - W = 0 \tag{3.4.7}$$

and for horizontal equilibrium

$$T - D = 0 \tag{3.4.8}$$

and taking moments about the aircraft's CG in the plane of symmetry

$$La - Db - Tc - M_0 - Pl = 0 \tag{3.4.9}$$

For a given aircraft weight, speed and altitude, Eqs (3.4.7)–(3.4.9) may be solved for the unknown lift, drag and tail loads. However, other parameters in these equations, such as M_0, depend upon the wing incidence α which in turn is a function of the required wing lift so that, in practice, a method of successive approximation is found to be the most convenient means of solution.

As a first approximation we assume that the tail load P is small compared with the wing lift L so that, from Eq. (3.4.7), $L \approx W$. From aerodynamic theory with the usual notation

$$L = \frac{1}{2}\rho V^2 S C_L$$

Hence

$$\frac{1}{2}\rho V^2 S C_L \approx W \tag{3.4.10}$$

Equation (3.4.10) gives the approximate lift coefficient C_L and thus (from C_L–α curves established by wind tunnel tests) the wing incidence α. The drag load D follows (knowing V and α) and hence we obtain the required engine thrust T from Eq. (3.4.8). Also M_0, a, b, c and l may be calculated (again since V and α are known) and Eq. (3.4.9) solved for P. As a second approximation this value of P is substituted in Eq. (3.4.7) to obtain a more accurate value for L and the procedure is repeated. Usually three approximations are sufficient to produce reasonably accurate values.

In most cases P, D and T are small compared with the lift and aircraft weight. Therefore, from Eq. (3.4.7) $L \approx W$ and substitution in Eq. (3.4.9) gives, neglecting D and T

$$P \approx W\frac{a}{l} - \frac{M_0}{l} \tag{3.4.11}$$

We see from Eq. (3.4.11) that if a is large then P will most likely be positive. In other words the tail load acts upwards when the CG of the aircraft is far aft. When a is small or negative, i.e., a forward CG, then P will probably be negative and act downwards.

3.4.2.2 General Case of a Symmetric Manoeuvre

In a rapid pull-out from a dive a downward load is applied to the tailplane, causing the aircraft to pitch nose upwards. The downward load is achieved by a backward movement of the control column, thereby applying negative incidence to the elevators, or horizontal tail if the latter is all-moving. If the manoeuvre is carried out rapidly the forward speed of the aircraft remains practically constant so that increases in lift and drag result from the increase in wing incidence only. Since the lift is now greater than that required to balance the aircraft weight, the aircraft experiences an upward acceleration normal to its flight path. This normal acceleration combined with the aircraft's speed in the dive results in the curved flight path shown in Fig. 3.4-7. As the drag load builds up with an increase of incidence the forward speed of the aircraft falls since the thrust is assumed to remain constant during the manoeuvre. It is usual, as we observed in the discussion of the flight envelope, to describe the manoeuvres of an aircraft in terms of a manoeuvring load factor n. For steady level flight $n = 1$, giving 1 g flight, although in fact the acceleration is zero. What is implied in this method of description is that the inertia force on the aircraft in the level flight condition is 1.0 times its weight. It follows that the vertical inertia force on an aircraft carrying out an ng manoeuvre is nW. We may therefore replace the dynamic conditions of the accelerated motion by an equivalent set of static conditions in which the applied loads are in equilibrium with the inertia forces. Thus, in Fig. 3.4-7, n is the manoeuvre load factor while f is a similar factor giving the horizontal inertia force. Note that the actual normal acceleration in this particular case is $(n - 1)g$.

For vertical equilibrium of the aircraft, we have, referring to Fig. 3.4-7 where the aircraft is shown at the lowest point of the pull-out,

$$L + P + T \sin \gamma - nW = 0 \tag{3.4.12}$$

For horizontal equilibrium

$$T \cos \gamma + fW - D = 0 \tag{3.4.13}$$

and for pitching moment equilibrium about the aircraft's CG

$$La - Db - Tc - M_0 - Pl = 0 \tag{3.4.14}$$

Equation (3.4.14) contains no terms representing the effect of pitching acceleration of the aircraft; this is assumed to be negligible at this stage.

Again the method of successive approximation is found to be most convenient for the solution of Eqs (3.4.12)–(3.4.14). There is, however, a difference to the procedure described for the steady level flight case. The engine thrust T is no longer directly related to the drag D as the latter changes during the manoeuvre. Generally, the thrust is regarded as remaining constant and equal to the value appropriate to conditions before the manoeuvre began.

Example 3.4.3

The curves C_D,α and $C_{M,CG}$ for a light aircraft are shown in Fig. 3.4-8(a). The aircraft weight is 8000 N, its wing area 14.5 m^2 and its mean chord 1.35 m. Determine the lift, drag, tail load and forward inertia force for a symmetric manoeuvre corresponding to $n = 4.5$ and a speed of 60 m/s. Assume that engine-off conditions apply and that the air density is 1.223 kg/m^3. Figure 3.4-8(b) shows the relevant aircraft dimensions.

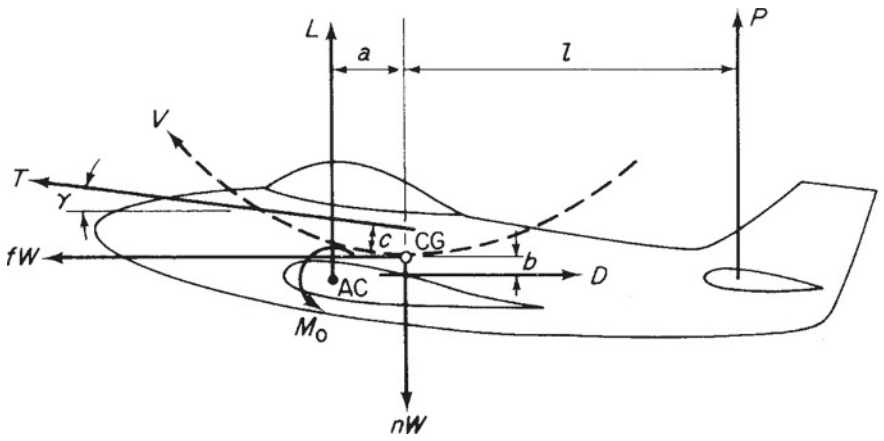


Figure 3.4-7 Aircraft loads in a pull-out from a dive.

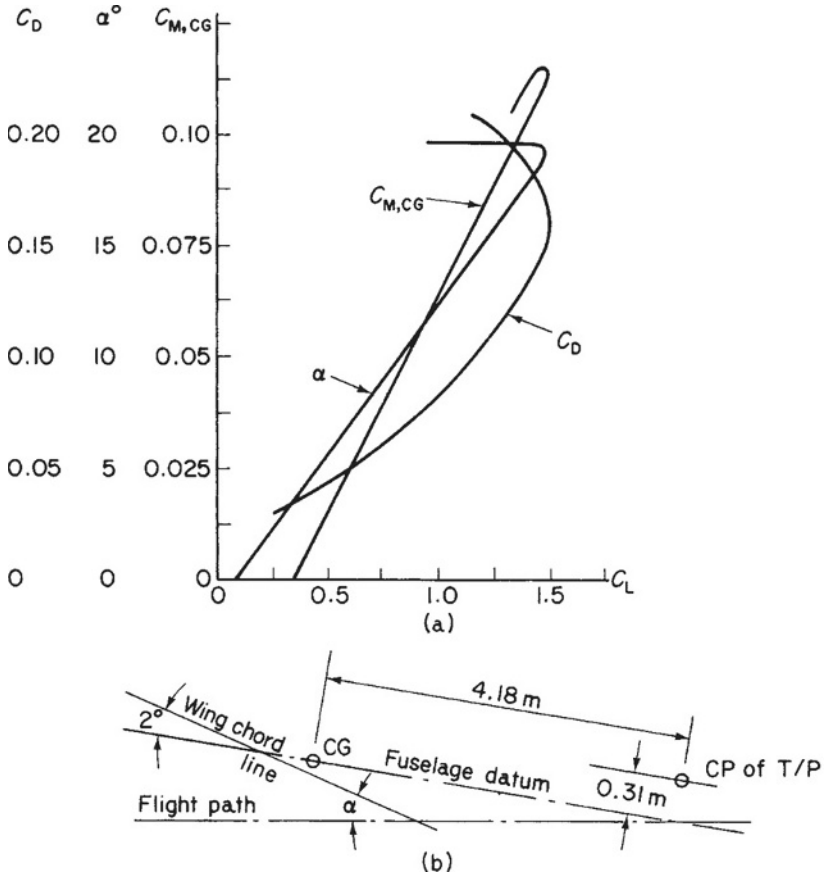


Figure 3.4-8 (a) C_D, α, $C_{M,CG}$–C_L curves for Example 3.4.3; (b) geometry of Example 3.4.3.

As a first approximation we neglect the tail load P. Therefore, from Eq. (3.4.12), since $T = 0$, we have

$$L \approx nW \tag{i}$$

Hence

$$C_L = \frac{L}{\frac{1}{2}\rho V^2 S} \approx \frac{4.5 \times 8000}{\frac{1}{2} \times 1.223 \times 60^2 \times 14.5} = 1.113$$

From Fig. 3.4-8(a), $\alpha = 13.75°$ and $C_{M,CG} = 0.075$. The tail arm l, from Fig. 3.4-8(b), is

$$l = 4.18 \cos(\alpha - 2) + 0.31 \sin(\alpha - 2) \tag{ii}$$

Substituting the above value of α gives $l = 4.123$m. In Eq. (3.4.14) the terms $La - Db - M_0$ are equivalent to the aircraft pitching moment M_{CG} about its CG. Equation (3.4.14) may therefore be written as

$$M_{CG} - Pl = 0$$

or

$$Pl = \frac{1}{2}\rho V^2 S c C_{M,CG} \tag{iii}$$

where c = wing mean chord. Substituting P from Eq. (iii) into Eq. (3.4.12) we have

$$L + \frac{\frac{1}{2}\rho V^2 S c C_{M,CG}}{l} = nW$$

or dividing through by $\frac{1}{2}\rho V^2 S$

$$C_L + \frac{c}{l} C_{M,CG} = \frac{nW}{\frac{1}{2}\rho V^2 S} \tag{iv}$$

We now obtain a more accurate value for C_L from Eq. (iv)

$$C_L = 1.113 - \frac{1.35}{4.123} \times 0.075 = 1.088$$

giving $\alpha = 13.3°$ and $C_{M,CG} = 0.073$.

Substituting this value of α into Eq. (ii) gives a second approximation for l, namely $l = 4.161$ m.

Equation (iv) now gives a third approximation for C_L, i.e. $C_L = 1.099$. Since the three calculated values of C_L are all extremely close, further approximations will not give values of C_L very much different to those above. Therefore, we shall take $C_L = 1.099$. From Fig. 3.4-8(a) $C_D = 0.0875$.

The values of lift, tail load, drag and forward inertia force then follow:

$$\text{Lift } L = \frac{1}{2}\rho V^2 S C_L = \frac{1}{2} \times 1.223 \times 60^2 \times 14.5 \times 1.099 = 35\,000 \text{ N}$$

$$\text{Tail load } P = nW - L = 4.5 \times 8000 - 35\,000 = 1000 \text{ N}$$

$$\text{Drag } D = \frac{1}{2}\rho V^2 S C_D = \frac{1}{2} \times 1.223 \times 60^2 \times 14.5 \times 0.0875 = 2790 \text{ N}$$

$$\text{Forward inertia force } fW = D \text{ (From Eq.(3.4.13))} = 2790 \text{ N}$$

3.4.3 Normal Accelerations Associated with Various Types of Manoeuvre

In Section 3.4.2 we determined aircraft loads corresponding to a given manoeuvre load factor n. Clearly it is necessary to relate this load factor to given types of manoeuvre. Two cases arise: the first involving a steady pull-out from a dive and the second, a correctly banked turn. Although the latter is not a symmetric manoeuvre in the strict sense of the word, it gives rise to normal accelerations in the plane of symmetry and is therefore included.

3.4.3.1 Steady Pull-Out

Let us suppose that the aircraft has just begun its pull-out from a dive so that it is describing a curved flight path but is not yet at its lowest point. The loads acting on the aircraft at this stage of the manoeuvre are shown in Fig. 3.4-9, where R is the radius of curvature of the flight path. In this case the lift vector must equilibrate the normal (to the flight path) component of the aircraft weight and provide the force producing the centripetal acceleration V^2/R of the aircraft towards the centre of curvature of the flight path. Thus

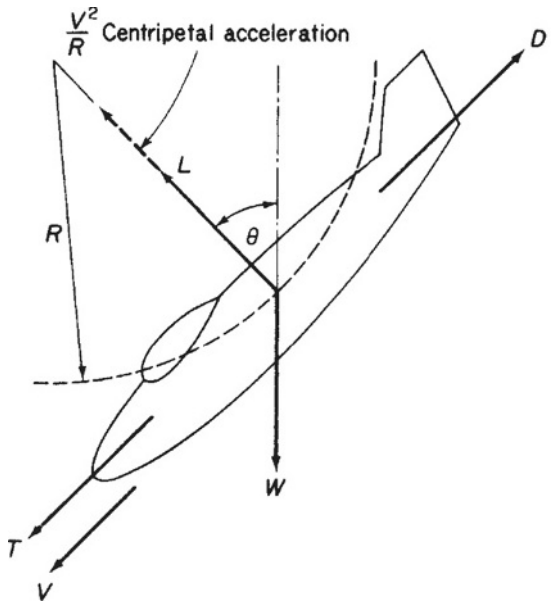


Figure 3.4-9 Aircraft loads and acceleration during a steady pull-out.

$$L = \frac{WV^2}{gR} + W\cos\theta$$

or, since $L = nW$ (see Section 3.4.2)

$$n = \frac{V^2}{gR} + \cos\theta \tag{3.4.15}$$

At the lowest point of the pull-out, $\theta = 0$, and

$$n = \frac{V^2}{gR} + 1 \tag{3.4.16}$$

We see from either Eq. (3.4.15) or Eq. (3.4.16) that the smaller the radius of the flight path, that is the more severe the pull-out, the greater the value of n. It is quite possible therefore for a severe pull-out to overstress the aircraft by subjecting it to loads which lie outside the flight envelope and which may even exceed the proof or ultimate loads. In practice, the control surface movement may be limited by stops incorporated in the control circuit. These stops usually operate only above a certain speed giving the aircraft adequate manoeuvrability at lower speeds. For hydraulically operated controls 'artificial feel' is built in to the system whereby the stick force increases progressively as the speed increases; a necessary precaution in this type of system since the pilot is merely opening and closing valves in the control circuit and therefore receives no direct physical indication of control surface forces.

Alternatively, at low speeds, a severe pull-out or pull-up may stall the aircraft. Again safety precautions are usually incorporated in the form of stall warning devices since, for modern high speed aircraft, a stall can be disastrous, particularly at low altitude.

3.4.3.2 Correctly Banked Turn

In this manoeuvre the aircraft flies in a horizontal turn with no sideslip at constant speed. If the radius of the turn is R and the angle of bank ϕ, then the forces acting on the aircraft are those shown in Fig. 3.4-10. The horizontal component of the lift vector in this case provides the force necessary to produce the centripetal acceleration of the aircraft towards the centre of the turn. Then

$$L\sin\phi = \frac{WV^2}{gR} \tag{3.4.17}$$

and for vertical equilibrium

$$L\cos\phi = W \tag{3.4.18}$$

or

$$L = W\sec\phi \tag{3.4.19}$$

From Eq. (3.4.19) we see that the load factor n in the turn is given by

$$n = \sec\phi \tag{3.4.20}$$

Also, dividing Eq. (3.4.17) by Eq. (3.4.18)

$$\tan\phi = \frac{V^2}{gR} \tag{3.4.21}$$

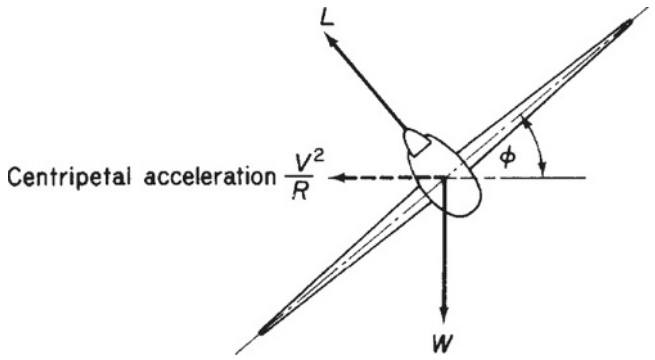


Figure 3.4-10 Correctly banked turn.

Examination of Eq. (3.4.21) reveals that the tighter the turn the greater the angle of bank required to maintain horizontal flight. Furthermore, we see from Eq. (3.4.20) that an increase in bank angle results in an increased load factor. Aerodynamic theory shows that for a limiting value of n the minimum time taken to turn through a given angle at a given value of engine thrust occurs when the lift coefficient C_L is a maximum; that is, with the aircraft on the point of stalling.

3.4.4 Gust Loads

In Section 3.4.2 we considered aircraft loads resulting from prescribed manoeuvres in the longitudinal plane of symmetry. Other types of in-flight load are caused by air turbulence. The movements of the air in turblence are generally known as gusts and produce changes in wing incidence, thereby subjecting the aircraft to sudden or gradual increases or decreases in lift from which normal accelerations result. These may be critical for large, high speed aircraft and may possibly cause higher loads than control initiated manoeuvers.

At the present time two approaches are employed in gust analysis. One method, which has been in use for a considerable number of years, determines the aircraft response and loads due to a single or 'discrete' gust of a given profile. This profile is defined as a distribution of vertical gust velocity over a given finite length or given period of time. Examples of these profiles are shown in Fig. 3.4-11.

Early airworthiness requirements specified an instantaneous application of gust velocity u, resulting in the 'sharp-edged' gust of Fig. 3.4-11(a). Calculations of normal acceleration and aircraft response were based on the assumptions that the aircraft's flight is undisturbed while the aircraft passes from still air into the moving air of the gust and during the time taken for the gust loads to build up; that the aerodynamic forces on the aircraft are determined by the instantaneous incidence of the particular lifting surface and finally that the aircraft's structure is rigid. The second assumption here relating the aerodynamic force on a lifting surface to its instantaneous incidence neglects the fact that in a disturbance such as a gust there is a gradual growth of circulation and hence of lift to a steady state value (Wagner effect). This in general leads to an overestimation of the upward acceleration of an aircraft and therefore of gust loads.

The 'sharp-edged' gust was replaced when it was realized that the gust velocity built up to a maximum over a period of time. Airworthiness requirements were modified on the assumption that the gust velocity increased linearly to a maximum value over a specified gust gradient distance H, hence the 'graded' gust of Fig. 3.4-11(b). In the UK, H is taken as 30.5 m. Since, as far as the aircraft is concerned, the gust velocity builds up to a maximum over a period of time it is no longer allowable to ignore the change of flight path as the aircraft enters the gust. By the time the gust has attained its maximum value the aircraft has developed a vertical component of velocity and, in addition, may be pitching depending on its longitudinal stability characteristics. The effect of the former is to reduce the severity of the gust while the latter may either increase or decrease the loads involved. To evaluate the corresponding gust loads the designer may either calculate the complete motion of the aircraft during the disturbance and hence

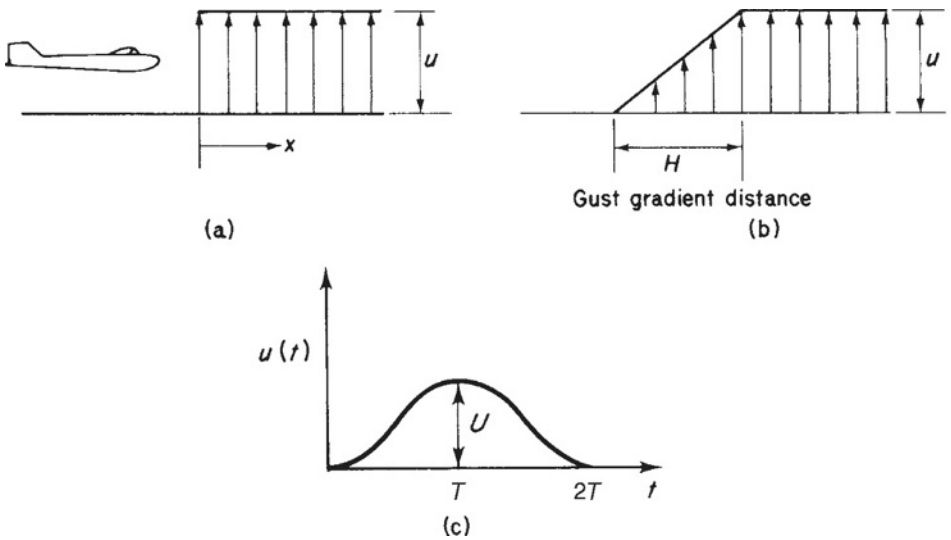


Figure 3.4-11 (a) Sharp-edged gust; (b) graded gust; (c) 1 – cosine gust.

obtain the gust loads, or replace the 'graded' gust by an equivalent 'sharp-edged' gust producing approximately the same effect. We shall discuss the latter procedure in greater detail later.

The calculation of the complete response of the aircraft to a 'graded' gust may be obtained from its response to a 'sharp-edged' or 'step' gust, by treating the former as comprising a large number of small 'steps' and superimposing the responses to each of these. Such a process is known as convolution or Duhamel integration. This treatment is desirable for large or unorthodox aircraft where aeroelastic (structural flexibility) effects on gust loads may be appreciable or unknown. In such cases the assumption of a rigid aircraft may lead to an underestimation of gust loads. The equations of motion are therefore modified to allow for aeroelastic in addition to aerodynamic effects. For small and medium-sized aircraft having orthodox aerodynamic features the equivalent 'sharp-edged' gust procedure is satisfactory.

While the 'graded' or 'ramp' gust is used as a basis for gust load calculations, other shapes of gust profile are in current use. Typical of these is the '1 − cosine' gust of Fig. 3.4-11(c), where the gust velocity u is given by $u(t) = (U/2)[1 - \cos(\pi t/T)]$. Again the aircraft response is determined by superimposing the responses to each of a large number of small steps.

Although the 'discrete' gust approach still finds widespread use in the calculation of gust loads, alternative methods based on *power spectral* analysis are being investigated. The advantage of the power spectral technique lies in its freedom from arbitrary assumptions of gust shapes and sizes. It is assumed that gust velocity is a random variable which may be regarded for analysis as consisting of a large number of sinusoidal components whose amplitudes vary with frequency. The *power spectrum* of such a function is then defined as the distribution of energy over the frequency range. This may then be related to gust velocity. To establish appropriate amplitude and frequency distributions for a particular random gust profile requires a large amount of experimental data. The collection of such data has been previously referred to in Section 3.3.2.

Calculations of the complete response of an aircraft and detailed assessments of the 'discrete' gust and power spectral methods of analysis are outside the scope of this book. More information may be found in Refs [1–4] at the end of the chapter. Our present analysis is confined to the 'discrete' gust approach, in which we consider the 'sharp-edged' gust and the equivalent 'sharp-edged' gust derived from the 'graded' gust.

3.4.4.1 'Sharp-Edged' Gust

The simplifying assumptions introduced in the determination of gust loads resulting from the 'sharp-edged' gust have been discussed in the earlier part of this section. In Fig. 3.4-12 the aircraft is flying at a speed V with wing incidence α_0 in still air. After entering the gust of upward velocity u, the incidence increases by an amount $\tan^{-1} u/V$, or since u is usually small compared with V, u/V. This is accompanied by an increase in aircraft speed from V to $(V^2 + u^2)^{\frac{1}{2}}$, but again this increase is neglected since u is small. The increase in wing lift ΔL is then given by

$$\Delta L = \frac{1}{2}\rho V^2 S \frac{\partial C_L}{\partial \alpha}\frac{u}{V} = \frac{1}{2}\rho V S \frac{\partial C_L}{\partial \alpha} u \tag{3.4.22}$$

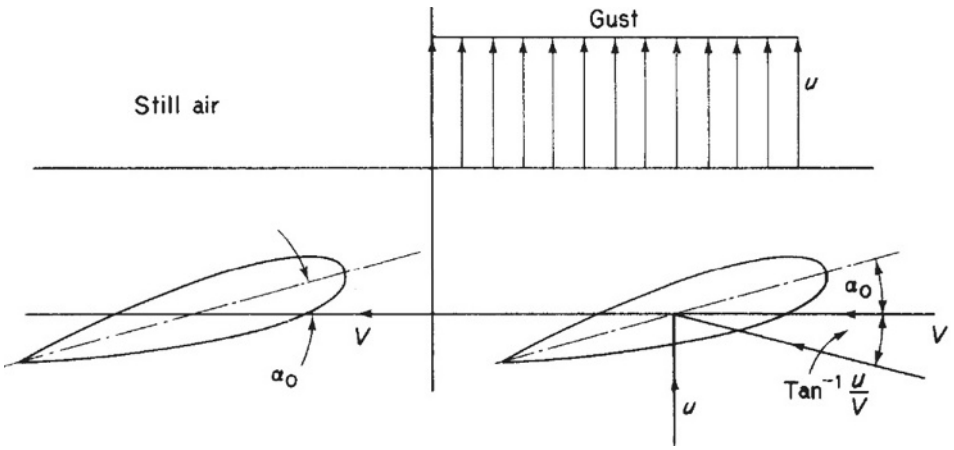


Figure 3.4-12 Increase in wing incidence due to a sharp-edged gust.

where $\partial C_L/\partial\alpha$ is the wing lift–curve slope. Neglecting the change of lift on the tailplane as a first approximation, the gust load factor Δn produced by this change of lift is

$$\Delta n = \frac{\frac{1}{2}\rho V S(\partial C_L/\partial\alpha)u}{W} \tag{3.4.23}$$

where W is the aircraft weight. Expressing Eq. (3.4.23) in terms of the wing loading, $w = W/S$, we have

$$\Delta n = \frac{\frac{1}{2}\rho V(\partial C_L/\partial\alpha)u}{w} \tag{3.4.24}$$

This increment in gust load factor is additional to the steady level flight value $n = 1$. Therefore, as a result of the gust, the total gust load factor is

$$n = 1 + \frac{\frac{1}{2}\rho V(\partial C_L/\partial\alpha)u}{w} \tag{3.4.25}$$

Similarly, for a downgust

$$n = 1 - \frac{\frac{1}{2}\rho V(\partial C_L/\partial\alpha)u}{w} \tag{3.4.26}$$

If flight conditions are expressed in terms of equivalent sea-level conditions then V becomes the equivalent airspeed (EAS), V_E, u becomes u_E and the air density ρ is replaced by the sea-level value ρ_0. Equations (3.4.25) and (3.4.26) are written as

$$n = 1 + \frac{\frac{1}{2}\rho_0 V_E(\partial C_L/\partial\alpha)u_E}{w} \tag{3.4.27}$$

and

$$n = 1 - \frac{\frac{1}{2}\rho_0 V_E(\partial C_L/\partial\alpha)u_E}{w} \tag{3.4.28}$$

We observe from Eqs (3.4.25)–(3.4.28) that the gust load factor is directly proportional to aircraft speed but inversely proportional to wing loading. It follows that high speed aircraft with low or moderate wing loadings are most likely to be affected by gust loads.

The contribution to normal acceleration of the change in tail load produced by the gust may be calculated using the same assumptions as before. However, the change in tailplane incidence is not equal to the change in wing incidence due to downwash effects at the tail. Thus if ΔP is the increase (or decrease) in tailplane load, then

$$\Delta P = \frac{1}{2}\rho_0 V_E^2 S_T \Delta C_{L,T} \tag{3.4.29}$$

where S_T is the tailplane area and $\Delta C_{L,T}$ the increment of tailplane lift coefficient given by

$$\Delta C_{L,T} = \frac{\partial C_{L,T}}{\partial\alpha}\frac{u_E}{V_E} \tag{3.4.30}$$

in which $\partial C_{L,T}/\partial\alpha$ is the rate of change of tailplane lift coefficient with wing incidence. From aerodynamic theory

$$\frac{\partial C_{L,T}}{\partial\alpha} = \frac{\partial C_{L,T}}{\partial\alpha_T}\left(1 - \frac{\partial\varepsilon}{\partial\alpha}\right)$$

where $\partial C_{L,T}/\partial\alpha_T$ is the rate of change of $C_{L,T}$ with tailplane incidence and $\partial\varepsilon/\partial\alpha$ the rate of change of downwash angle with wing incidence. Substituting for $\Delta C_{L,T}$ from Eq. (3.4.30) into Eq. (3.4.29), we have

$$\Delta P = \frac{1}{2}\rho_0 V_E S_T \frac{\partial C_{L,T}}{\partial\alpha} u_E \tag{3.4.31}$$

For positive increments of wing lift and tailplane load

$$\Delta n W = \Delta L + \Delta P$$

or, from Eqs (3.4.27) and (3.4.31)

$$\Delta n = \frac{\frac{1}{2}\rho_0 V_E (\partial C_L/\partial\alpha) u_E}{w}\left(1 + \frac{S_T}{S}\frac{\partial C_{L,T}/\partial\alpha}{\partial C_L/\partial\alpha}\right) \tag{3.4.32}$$

3.4.4.2 The 'Graded' Gust

The 'graded' gust of Fig. 3.4.11(b) may be converted to an equivalent 'sharp-edged' gust by multiplying the maximum velocity in the gust by a *gust alleviation factor*, F. Equation (3.4.27) then becomes

$$n = 1 + \frac{\frac{1}{2}\rho_0 V_E (\partial C_L/\partial\alpha) F u_E}{w} \tag{3.4.33}$$

Similar modifications are carried out on Eqs (3.4.25), (3.4.26), (3.4.28) and (3.4.32). The gust alleviation factor allows for some of the dynamic properties of the aircraft, including unsteady lift, and has been calculated taking into account the heaving motion (i.e. the up and down motion with zero rate of pitch) of the aircraft only.[5]

Horizontal gusts cause lateral loads on the vertical tail or fin. Their magnitudes may be calculated in an identical manner to those above, except that areas and values of lift curve slope are referred to the vertical tail. Also, the gust alleviation factor in the 'graded' gust case becomes F_1 and includes allowances for the aerodynamic yawing moment produced by the gust and the yawing inertia of the aircraft.

3.4.4.3 Gust Envelope

Airworthiness requirements usually specify that gust loads shall be calculated at certain combinations of gust and flight speed. The equations for gust load factor in the above analysis show that n is proportional to aircraft speed for a given gust velocity. Therefore, we may plot a gust envelope similar to the flight envelope of Fig. 3.3-1, as shown in Fig. 3.4-13. The gust speeds $\pm U_1$, $\pm U_2$ and $\pm U_3$ are high, medium and low velocity gusts, respectively. Cut-offs occur at points where the lines corresponding to each gust velocity meet specific aircraft speeds. For example, A and F denote speeds at which a gust of velocity $\pm U_1$ would stall the wing.

The lift coefficient-incidence curve is, as we noted in connection with the flight envelope, affected by compressibility and therefore altitude so that a series of gust envelopes should be drawn for different altitudes. An additional variable in the equations for gust load factor is the wing loading w. Further gust envelopes should therefore be drawn to represent different conditions of aircraft loading.

Typical values of U_1, U_2 and U_3 are 20 m/s, 15.25 m/s and 7.5 m/s. It can be seen from the gust envelope that the maximum gust load factor occurs at the cruising speed V_C. If this value of n exceeds that for the corresponding flight envelope case, that is n_1, then the gust case will be the most critical in the cruise. Let us consider a civil, non-aerobatic aircraft for which n_1 =2.5, $w = 2400$ N/m^2 and $\partial C_L/\partial\alpha = 5.0$/rad. Taking $F = 0.715$, we have, from Eq. (3.4.33),

$$n = 1 + \frac{\frac{1}{2} \times 1.223\, V_C \times 5.0 \times 0.715 \times 15.25}{2400}$$

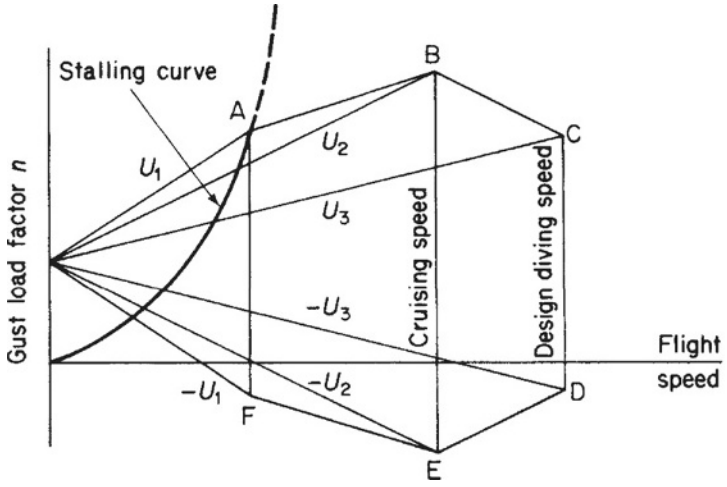


Figure 3.4-13 Typical gust envelope.

giving $n = 1 + 0.0139V_C$, where the cruising speed V_C is expressed as an EAS. For the gust case to be critical

$$1 + 0.0139V_C > 2.5$$

or

$$V_C > 108 \text{ m/s}$$

Thus, for civil aircraft of this type having cruising speeds in excess of 108 m/s, the gust case is the most critical. This would, in fact, apply to most modern civil airliners.

Although the same combination of V and n in the flight and gust envelopes will produce the same total lift on an aircraft, the individual wing and tailplane loads will be different, as shown previously (see the derivation of Eq. (3.4.33)). This situation can be important for aircraft such as the Airbus, which has a large tailplane and a CG forward of the aerodynamic centre. In the flight envelope case the tail load is downwards whereas in the gust case it is upwards; clearly there will be a significant difference in wing load.

The transference of manoeuvre and gust loads into bending, shear and torsional loads on wings, fuselage and tailplanes has been discussed in Section 3.2.1. Further loads arise from aileron application, in undercarriages during landing, on engine mountings and during crash landings. Analysis and discussion of these may be found in Ref. [6].

References

1. Zbrozek, J. K., Atmospheric gusts – present state of the art and further research, *J. Roy. Aero. Soc.*, January 1965.
2. Cox, R. A., A comparative study of aircraft gust analysis procedures, *J. Roy. Aero. Soc.*, October 1970.
3. Bisplinghoff, R. L., Ashley, H. and Halfman, R. L., *Aeroelasticity*, Addison-Wesley Publishing Co. Inc., Cambridge, Mass., 1955.
4. Babister, A.W., *Aircraft Stability and Control*, Pergamon Press, London, 1961.
5. Zbrozek, J. K., *Gust Alleviation Factor*, R. and M. No. 2970, May 1953.
6. *Handbook of Aeronautics No. 1. Structural Principles and Data*, 4th edition, The Royal Aeronautical Society, 1952.

Problems

P. 3.4.1 The aircraft shown in Fig. P. 3.4-1(a) weighs 135 kN and has landed such that at the instant of impact the ground reaction on each main undercarriage wheel is 200 kN and its vertical velocity is 3.5 m/s.

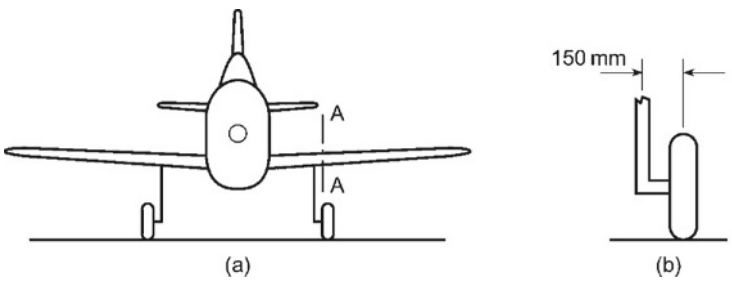


Figure P.3.4-1

If each undercarriage wheel weighs 2.25 kN and is attached to an oleo strut, calculate the axial load and bending moment in the strut; the strut may be assumed to be vertical. Determine also the shortening of the strut when the vertical velocity of the aircraft is zero.

Finally, calculate the shear force and bending moment in the wing at the section AA if the wing, outboard of this section, weighs 6.6 kN and has its CG 3.05 m from AA.

Ans. 193.3 kN, 29.0 kN m (clockwise); 0.32 m; 19.5 kN, 59.6 kN m (anticlock-wise).

P.3.4.2 Determine, for the aircraft of Example 3.4.2, the vertical velocity of the nose wheel when it hits the ground.

Ans. 3.1 m/s.

P.3.4.3 Figure P.3.4-3 shows the flight envelope at sea-level for an aircraft of wing span 27.5 m, average wing chord 3.05 m and total weight 196 000 N. The aerodynamic centre is 0.915 m forward of the CG and the centre of lift for the tail unit is 16.7 m aft of the CG. The pitching moment coefficient is

$$C_{M,0} = -0.0638 \text{ (nose-up positive)}$$

both $C_{M,0}$ and the position of the aerodynamic centre are specified for the complete aircraft less tail unit.

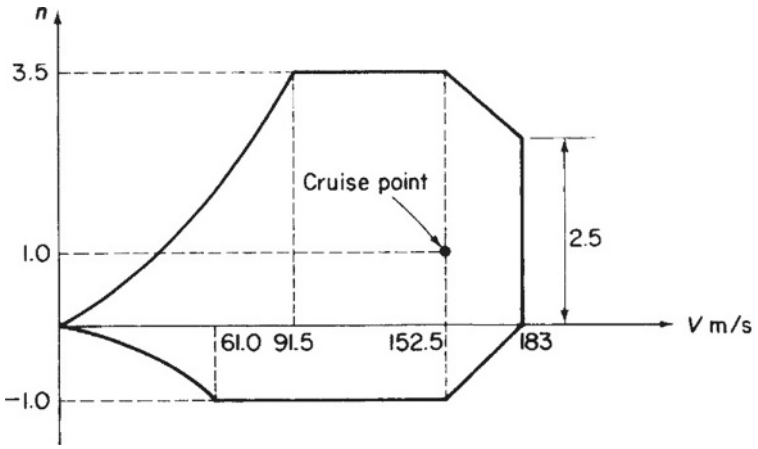


Figure P.3.4-3

For steady cruising flight at sea level the fuselage bending moment at the CG is 600 000 N m. Calculate the maximum value of this bending moment for the given flight envelope. For this purpose it may be assumed that the aerodynamic loadings on the fuselage itself can be neglected, i.e. the only loads on the fuselage structure aft of the CG are those due to the tail lift and the inertia of the fuselage.

Ans. 1549 500 N m at $n = 3.5$, $V = 152.5$ m/s.

P.3.4.4 An aircraft weighing 238 000 N has wings 88.5 m^2 in area for which $C_D = 0.0075 + 0.045C_L^2$ The extra-to-wing drag coefficient based on wing area is 0.0128 and the pitching moment coefficient for all parts excluding the tailplane about an axis through the CG is given by $C_M \cdot c = (0.427C_L - 0.061)$m. The radius from the CG to the line of action of the tail lift may be taken as constant at 12.2 m. The moment of inertia of the aircraft for pitching is 204 000 kg m^2.

During a pull-out from a dive with zero thrust at 215 m/s EAS when the flight path is at 40° to the horizontal with a radius of curvature of 1525 m, the angular velocity of pitch is checked by applying a retardation of 0.25 rad/s^2. Calculate the manoeuvre load factor both at the CG and at the tailplane CP, the forward inertia coefficient and the tail lift.

Ans. $n = 3.78$(CG), $n = 5.19$ at TP, $f = -0.370$, $P = 18\,925$ N.

P.3.4.5 An aircraft flies at sea level in a correctly banked turn of radius 610 m at a speed of 168 m/s. Figure P.3.4-5 shows the relative positions of the CG, aerodynamic centre of the complete aircraft less tailplane and the tailplane centre of pressure for the aircraft at zero lift incidence.

Calculate the tail load necessary for equilibrium in the turn. The necessary data are given in the usual notation as follows:

Weight $W = 133\,500$ N	$dC_L/d\alpha = 4.5$/rad
Wing area $S = 46.5$ m^2	$C_D = 0.01 + 0.05C_L^2$
Wing mean chord $\bar{c} = 3.0$ m	$C_{M,0} = -0.03$

Ans. 73 160 N

P.3.4.6 The aircraft for which the stalling speed V_s in level flight is 46.5 m/s has a maximum allowable manoeuvre load factor n_1 of 4.0. In assessing gyroscopic effects on the engine mounting the following two cases are to be considered:

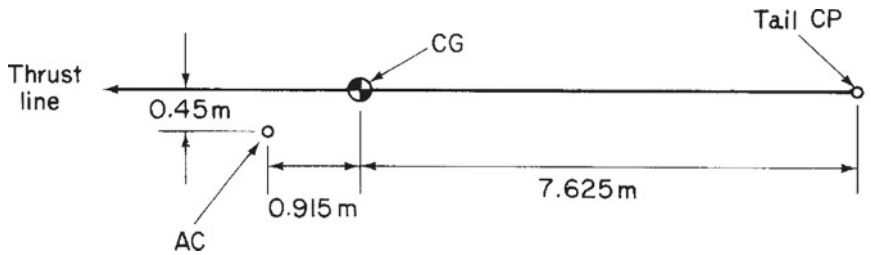


Figure P.3.4-5

(a) Pull-out at maximum permissible rate from a dive in symmetric flight, the angle of the flight path to the horizontal being limited to 60° for this aircraft.
(b) Steady, correctly banked turn at the maximum permissible rate in horizontal flight.

Find the corresponding maximum angular velocities in yaw and pitch.

Ans. (a) Pitch, 0.37 rad/s, (b) Pitch, 0.41 rad/s, Yaw, 0.103 rad/s.

P.3.4.7 A tail-first supersonic airliner, whose essential geometry is shown in Fig. P.3.4-7, flies at 610 m/s true airspeed at an altitude of 18 300 m. Assuming that thrust and drag forces act in the same straight line, calculate the tail lift in steady straight and level flight.

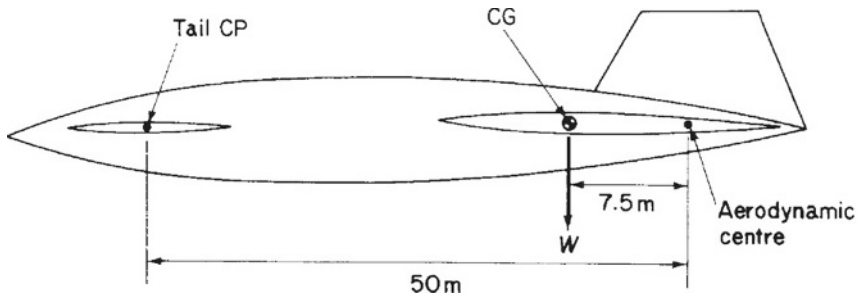


Figure P.3.4-7

If, at the same altitude, the aircraft encounters a sharp-edged vertical up-gust of 18 m/s true airspeed, calculate the changes in the lift and tail load and also the resultant load factor n.
The relevant data in the usual notation are as follows:

Wing: $S = 280\ \text{m}^2$, $\partial C_L/\partial\alpha = 1.5$
Tail: $S_T = 28\ \text{m}^2$, $\partial C_{L,T}/\partial\alpha = 2.0$
Weight $W = 1\,600\,000$ N
$C_{M,0} = -0.01$
Mean chord $\bar{c} = 22.8$ m

At 18 300 m

$\rho = 0.116\ \text{kg/m}^3$

Ans. $P = 267\,852$ N, $\Delta P = 36\,257$ N, $\Delta L = 271\,931$ N, $n = 1.19$

P.3.4.8 An aircraft of all up weight 145 000 N has wings of area 50 m^2 and mean chord 2.5 m. For the whole aircraft $C_D = 0.021 + 0.041C_L^2$ for the wings $dC_L/d\alpha = 4.8$, for the tailplane of area 9.0 m^2, $dC_{L,T}/d\alpha = 2.2$ allowing for the effects of downwash, and the pitching moment coefficient about the aerodynamic centre (of complete aircraft less tailplane) based on wing area is $C_{M,0} = -0.032$. Geometric data are given in Fig. P.3.4-8.

During a steady glide with zero thrust at 250 m/s EAS in which $C_L = 0.08$, the aircraft meets a downgust of equivalent 'sharp-edged' speed 6 m/s. Calculate the tail load, the gust load factor and the forward inertia force, $\rho_0 = 1.223\ \text{kg/m}^3$.

Ans. $P = -28\,902$ N (down), $n = -0.64$, forward inertia force = 40 703 N.

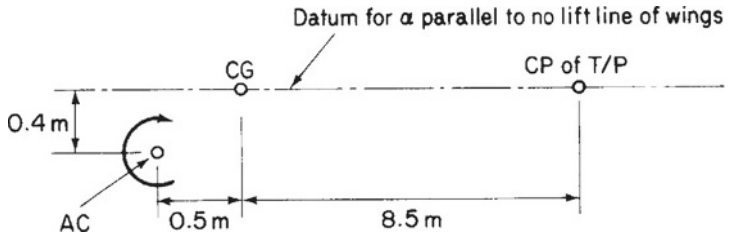


Figure P.3.4-8

SECTION 4

Flight Dynamics

This page intentionally left blank

Aerospace Engineering Desk Reference; ISBN 9781856175753

4.1 Introduction to Flight Dynamics

Michael Cook

4.1.1 Overview

This book is primarily concerned with the provision of good flying and handling qualities in conventional piloted aircraft, although the material is equally applicable to the *uninhabited air vehicle* (UAV). Consequently it is also very much concerned with the stability, control and dynamic characteristics which are fundamental to the determination of those qualities. Since flying and handling qualities are of critical importance to safety and to the piloting task it is essential that their origins are properly understood. Here then, the intention is to set out the basic principles of the subject at an introductory level and to illustrate the application of those principles by means of worked examples.

Following the first flights made by the Wright brothers in December 1903 the pace of aeronautical development quickened and the progress made in the following decade or so was dramatic. However, the stability and control problems that faced the early aviators were sometimes considerable since the flying qualities of their aircraft were often less than satisfactory. Many investigators were studying the problems of stability and control at the time although it is the published works of Bryan (1911) and Lanchester (1908) which are usually accredited with laying the first really secure foundations for the subject. By conducting many experiments with flying models Lanchester was able to observe and successfully describe mathematically some dynamic characteristics of aircraft. The beauty of Lanchester's work was its practicality and theoretical simplicity, thereby lending itself to easy application and interpretation. Bryan, on the other hand, was a mathematician who chose to apply his energies, with the assistance of Mr. Harper, to the problems of the stability and control of aircraft. Bryan developed the general equations of motion of a rigid body with six degrees of freedom to successfully describe the motion of aircraft. His treatment, with very few changes, is still in everyday use. What has changed is the way in which the material is now used, due largely to the advent of the digital computer as an analysis tool. The stability and control of aircraft is a subject which has its origins in aerodynamics and the classical theory of the subject is traditionally expressed in the language of the aerodynamicist. However, most advanced technology aircraft may be described as an *integrated system* comprising airframe, propulsion, flight controls and so on. It is therefore convenient and efficient to utilise powerful computational systems engineering tools to analyse and describe its flight dynamics. Thus, the objective of the present work is to revisit the development of the classical theory and to express it in the language of the systems engineer where it is more appropriate to do so.

Flight dynamics is about the relatively short term motion of aircraft in response to controls or to external disturbances such as atmospheric turbulence. The motion of interest can vary from small excursions about trim to very large amplitude manoeuvring when normal aerodynamic behaviour may well become very non-linear. Since the treatment of the subject in this book is introductory a discussion of large amplitude dynamics is beyond the scope of the present work. The dynamic behaviour of an aircraft is shaped significantly by its stability and control properties which in turn have their roots in the aerodynamics of the airframe. Previously the achievement of aircraft with good stability characteristics usually ensured good flying qualities, all of which depended only on good aerodynamic design. Expanding flight envelopes and the increasing dependence on *automatic flight control systems* (AFCS) for stability augmentation means that good flying qualities are no longer a guaranteed product of good aerodynamic design and good stability characteristics. The reasons for this apparent inconsistency are now reasonably well understood and, put very simply, result from the addition of flight control system dynamics to those of the airframe. Flight control system dynamics are of course a necessary, but not always desirable, by-product of command and stability augmentation.

Modern flight dynamics is concerned not only with the dynamics, stability and control of the basic airframe, but also with the sometimes complex interaction between airframe and flight control system. Since the flight control system comprises motion sensors, a control computer, control actuators and other essential items of control hardware, a study of the subject becomes a multi-disciplinary activity. Therefore, it is essential that the modern flight dynamicist has not only a thorough understanding of the classical stability and control theory of aircraft, but also a working knowledge of control theory and of the use of computers in *flight critical* applications. Thus modern aircraft comprise the airframe together with the flight control equipment and may be treated as a whole *system* using the traditional tools of the aerodynamicist together with the analytical tools of the control engineer.

Thus in a modern approach to the analysis of stability and control it is convenient to treat the airframe as a system component. This leads to the derivation of mathematical models which describe aircraft in terms of aerodynamic *transfer functions*. Described in this way, the stability, control and dynamic characteristics of aircraft are readily interpreted with the aid of very powerful computational systems engineering tools. It follows that the mathematical model of the aircraft is immediately compatible with and provides the foundation for integration with flight control system studies. This is an ideal state of affairs since, today, it is common place to undertake stability and control investigations as a precursor to flight control system development.

Today, the modern flight dynamicist tends to be concerned with the wider issues of flying and handling qualities rather than with the traditional, and more limited issues of stability and control. The former are, of course, largely determined by the latter. The present treatment of the material is shaped by answering the following questions which a newcomer to the subject might be tempted to ask:

(i) *How are the stability and control characteristics of aircraft determined and how do they influence flying qualities?*
The answer to this question involves the establishment of a suitable mathematical framework for the problem, the development of the equations of motion, the solution of the equations of motion, investigation of response to controls and the general interpretation of dynamic behaviour.

(ii) *What are acceptable flying qualities, how are the requirements defined, interpreted and applied, and how do they limit flight characteristics?*
The answer to this question involves a review of contemporary flying qualities requirements and their evaluation and interpretation in the context of stability and control characteristics.

(iii) *When an aircraft has unacceptable flying qualities how may its dynamic characteristics be improved?*
The answer to this question involves an introduction to the rudiments of feedback control as the means for augmenting the stability of the basic airframe.

4.1.2 Flying and Handling Qualities

The flying and handling qualities of an aircraft are those properties which describe the ease and effectiveness with which it responds to pilot commands in the execution of a flight task, or *mission task element* (MTE). In the first instance, therefore, flying and handling qualities are described qualitatively and are formulated in terms of *pilot opinion*; consequently, they tend to be rather subjective. The process involved in the pilot perception of flying and handling qualities may be interpreted in the form of a signal flow diagram as shown in Fig. 4.1-1. The solid lines represent physical, mechanical or electrical signal flow paths, whereas the dashed lines represent sensory feedback information to the pilot. The author's interpretation distinguishes between *flying qualities* and *handling qualities* as indicated. The pilot's perception of flying qualities is considered to comprise a qualitative description of how well the aeroplane carries out the commanded task. On the other hand, the pilot's perception of handling qualities is considered a qualitative description of the adequacy of the short term dynamic response to controls in the execution of the flight task. The two *qualities* are therefore very much interdependent and in practice are probably inseparable. Thus to summarise, the flying qualities may be regarded as being task related, whereas the handling qualities may be regarded as being response related. When the airframe characteristics are augmented by a flight control system the way in which the flight control system may influence the flying and handling qualities is clearly shown in Fig. 4.1-1.

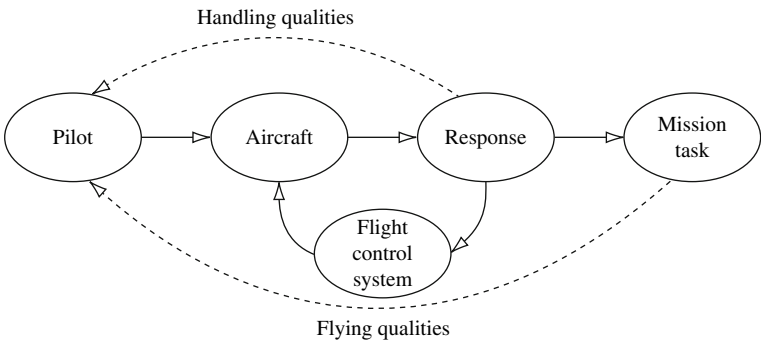


Figure 4.1-1 Flying and handling qualities of conventional aircraft.

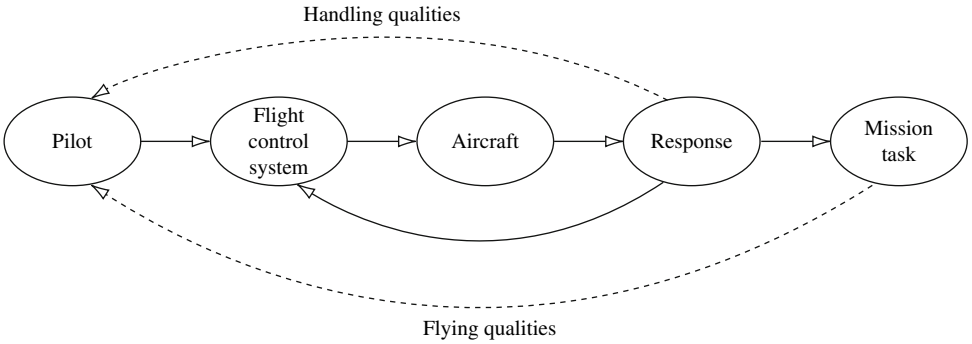


Figure 4.1-2 Flying and handling qualities of FBW aircraft.

An increasing number of advanced modern aeroplanes employ *fly-by-wire* (FBW) primary flight controls and these are usually integrated with the stability augmentation system. In this case, the interpretation of the flying and handling qualities process is modified to that shown in Fig. 4.1-2. Here then, the flight control system becomes an integral part of the primary signal flow path and the influence of its dynamic characteristics on flying and handling qualities is of critical importance. The need for very careful consideration of the influence of the flight control system in this context cannot be over emphasised.

Now the pilot's perception of the flying and handling qualities of an aircraft will be influenced by many factors. For example, the stability, control and dynamic characteristics of the airframe, flight control system dynamics, response to atmospheric disturbances and the less tangible effects of cockpit design. This last factor includes considerations such as control inceptor design, instrument displays and field of view from the cockpit. Not surprisingly the quantification of flying qualities remains difficult. However, there is an overwhelming necessity for some sort of numerical description of flying and handling qualities for use in engineering design and evaluation. It is very well established that the flying and handling qualities of an aircraft are intimately dependent on the stability and control characteristics of the airframe including the flight control system when one is installed. Since stability and control parameters are readily quantified these are usually used as indicators and measures of the likely flying qualities of the aeroplane. Therefore, the prerequisite for almost any study of flying and handling qualities is a descriptive mathematical model of the aeroplane which is capable of providing an adequate quantitative indication of its stability, control and dynamic properties.

4.1.3 General Considerations

In a systematic study of the principles governing the flight dynamics of aircraft it is convenient to break the problem down into manageable descriptive elements. Thus before attempting to answer the questions posed in Section 4.1.1, it is useful to consider and define a suitable framework in which the essential mathematical development may take place.

4.1.3.1 Basic Control–Response Relationships

It is essential to define and establish a description of the basic input–output relationships on which the flying and handling qualities of unaugmented aircraft depend. These relationships are described by the aerodynamic transfer functions which provide the simplest and most fundamental description of airframe dynamics. They describe the control–response relationship as a function of flight condition and may include the influence of atmospheric disturbances when appropriate. These basic relationships are illustrated in Fig. 4.1-3.

Central to this framework is a mathematical model of the aircraft which is usually referred to as *the equations of motion.* The equations of motion provide a complete description of response to controls, subject only to modelling limitations defined at the outset, and is measured in terms of displacement, velocity and acceleration variables. The flight condition describes the conditions under which the observations are made and includes parameters, such as Mach number, altitude, aircraft geometry, mass and trim state. When the airframe is augmented with a flight control system the equations of motion are modified to model this configuration. The response transfer functions, derived from the mathematical solution of the equations of motion, are then no longer the basic aerodynamic transfer functions but are obviously the transfer functions of the augmented aeroplane.

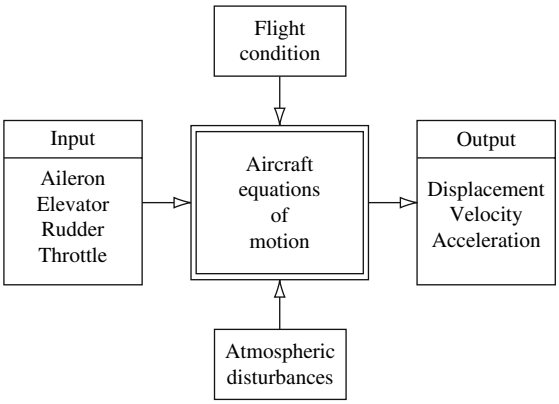


Figure 4.1-3 Basic control–response relationships.

4.1.3.2 Mathematical Models

From the foregoing it is apparent that it is necessary to derive mathematical models to describe the aircraft, its control systems, atmospheric disturbances and so on. The success of any flight dynamics analysis hinges on the suitability of the models for the problem in hand. Often the temptation is to attempt to derive the most accurate model possible. High fidelity models are capable of reproducing aircraft dynamics accurately but are seldom simple. Their main drawback is the lack of *functional visibility*. In very complex aircraft and system models, it may be difficult, or even impossible, to relate response to the simple physical aerodynamic properties of the airframe, or to the properties of the control system components. For the purposes of the investigation of flying and handling qualities it is frequently adequate to use simple approximate models which have the advantage of maximising functional visibility thereby drawing attention to the dominant characteristics. Such models have the potential to enhance the visibility of the physical principles involved thereby facilitating the interpretation of flying and handling qualities enormously. Often, the deterioration in the fidelity of the response resulting from the use of approximate models may be relatively insignificant. For a given problem, it is necessary to develop a model which balances the desire for response fidelity against the requirement to maintain functional visibility. As is so often the case in many fields of engineering, simplicity is a most desirable virtue.

4.1.3.3 Stability and Control

Flying and handling qualities are substantially dependent on, and are usually described in terms of, the stability and control characteristics of an aircraft. It is therefore essential to be able to describe and to quantify stability and control parameters completely. Analysis may then be performed using the stability parameters. Static stability analysis enables the *control displacement* and the *control force* characteristics to be determined for both steady and manoeuvring flight conditions. Dynamic stability analysis enables the temporal *response to controls* and to atmospheric disturbances to be determined for various flight conditions.

4.1.3.4 Stability and Control Augmentation

When an aircraft has deficiencies in flying and handling qualities, it becomes necessary to correct, or *augment*, the aerodynamic characteristics which give rise to those deficiencies. To a limited extent, this could be achieved by modification of the aerodynamic design of the aircraft. In this event it is absolutely essential to understand the relationship between the aerodynamics of the airframe and controls and the stability and control characteristics of that airframe. However, today, many aircraft are designed with their aerodynamics optimised for performance over a very large flight envelope, and a consequence of this is that their flying qualities are often deficient. The intent at the outset is to rectify those deficiencies with a stability augmentation system. Therefore, the alternative to aerodynamic design modification is the introduction of a flight control system. In this case it becomes essential to understand how feedback control techniques may be used to artificially modify the apparent aerodynamic characteristics of the airframe. So once again, but for different reasons, it is absolutely essential to understand the relationship between the aerodynamics of the airframe and its stability and control characteristics. Further, it becomes very important to appreciate the effectiveness of servo systems for autostabilisation whilst acknowledging the attendant advantages, disadvantages and limitations introduced by the system hardware. At this stage of

consideration it is beginning to become obvious why flight dynamics is now a complex multi-disciplinary subject. However, since this work is introductory, the subject of stability augmentation is treated at the most elementary level only.

4.1.4 Aircraft Equations of Motion

The equations of motion of an aeroplane are the foundation on which the entire framework of flight dynamics is built and provide the essential key to a proper understanding of flying and handling qualities. At their simplest, the equations of motion can describe small perturbation motion about trim only. At their most complex they can be completely descriptive embodying static stability, dynamic stability, aeroelastic effects, atmospheric disturbances and control system dynamics simultaneously for a given aeroplane configuration. The equations of motion enable the rather intangible description of flying and handling qualities to be related to quantifiable stability and control parameters, which in turn may be related to identifiable aerodynamic characteristics of the airframe. For initial studies the theory of small perturbations is applied to the equations to ease their analytical solution and to enhance their functional visibility. However, for more advanced applications, which are beyond the scope of the present work, the fully descriptive non-linear form of the equations might be retained. In this case the equations are difficult to solve analytically and recourse would be made to computer simulation techniques to effect a numerical solution.

4.1.5 Aerodynamics

4.1.5.1 Scope

The aerodynamics of an airframe and its controls make a fundamental contribution to the stability and control characteristics of the aircraft. It is usual to incorporate aerodynamic descriptions in the equations of motion in the form of *aerodynamic stability and control derivatives*. Since it is necessary to constrain the motion to well-defined limits in order to obtain the derivatives, the scope of the resulting aircraft model is similarly constrained in its application. It is, however, quite common to find aircraft models constrained in this way being used to predict flying and handling qualities at conditions well beyond the imposed limits. This is not recommended practice! An important aspect of flight dynamics is concerned with the proper definition of aerodynamic derivatives as functions of common aerodynamic parameters. It is also most important that the values of the derivatives are compatible with the scope of the problem to which the aircraft model is to be applied. The processes involved in the estimation or measurement of aerodynamic derivatives provide an essential contribution to a complete understanding of aircraft behaviour.

4.1.5.2 Small Perturbations

The aerodynamic properties of an aircraft vary considerably over the flight envelope and mathematical descriptions of those properties are approximations at best. The limit of the approximation is determined by the ability of mathematics to describe the physical phenomena involved or by the acceptable complexity of the description. The aim is to obtain the simplest approximation consistent with adequate physical representation. In the first instance this aim is best met when the motion of interest is constrained to *small perturbations* about a steady flight condition, which is usually, but not necessarily, trimmed equilibrium. This means that the aerodynamic characteristics can be approximated by *linearising* about the chosen flight condition. Simple approximate mathematical descriptions of aerodynamic stability and control derivatives then follow relatively easily. This is the approach pioneered by Bryan (1911) and it usually works extremely well provided the limitations of the model are recognised from the outset.

4.1.6 Computers

No discussion of flight dynamics would be complete without mention of the very important role played by the computer in all aspects of the subject. It is probably true to say that the development of today's very advanced aircraft would not have been possible without parallel developments in computer technology. In fact there is ample evidence to suggest that the demands of aeronautics have forced the pace of computer development. Computers are used for two main purposes, as a general purpose tool for design and analysis and to provide the "intelligence" in flight control systems.

4.1.6.1 Analytical Computers

In the past all electronic computation whether for analysis, simulation or airborne flight control would have been analogue. Analog computer technology developed rapidly during and immediately after World War II and by the late 1960s had reached its highest level of development following the introduction of the electronic integrated operational amplifier. Its principal role was that of simulation and its main advantages were: its ability to run in real time, continuous electrical signals and its high level of functional visibility. Its main disadvantage was the fact that the electronic hardware required was directly proportional to the functional complexity of the problem to be simulated. This meant that complex aircraft models with complex flight control systems required physically large, and very costly, electronic computer hardware. During the 1960s and 1970s electronic digital computing technology advanced very rapidly and soon displaced the analog computer as the primary tool for design and analysis. However, it took somewhat longer before the digital computer had acquired the capacity and speed necessary to meet the demands of simulation. Today, most of the computational requirements for design, analysis and simulation can be provided by a modest personal computer.

4.1.6.2 Flight Control Computers

In the present context flight control is taken to mean *flight critical stability augmentation,* where a computer malfunction or failure might hazard the continued safe operation of the aircraft. In the case of a FBW computer, a total failure would mean total loss of control of the aircraft, for example. Therefore, hardware integrity is a very serious issue in flight control computer design. The modern aircraft may also have an autopilot computer, air data computer, navigation computer, energy management computer, weapon systems computer and more. Many of these additional computers may be capable of exercising some degree of control over the aircraft, but none will be quite as critical as the stability augmentation computer in the event of a malfunction.

For the last 60 years or more, computers have been used in aircraft for flight control. For much of that time the dedicated analog electronic computer was unchallenged because of its relative simplicity, its easy interface engineering with the mechanical flying controls and its excellent safety record. Toward the end of the 1970s the digital computer had reached the stage of development where its use in flight critical applications became a viable proposition with the promise of vastly expanded control capability. The pursuit of increasingly sophisticated performance goals led to an increase in the complexity of the aerodynamic design of aircraft. This in turn placed greater demands on the flight control system for the maintenance of good flying and handling qualities. The attraction of the digital computer for flight control is its capability for handling very complex control functions easily. The disadvantage is its lack of functional visibility and the consequent difficulty of ensuring safe trouble-free operation. However, the digital flight critical computer is here to stay and is now used in most advanced technology aircraft. Research continues to improve the hardware, software and application. Confidence in digital flight control systems is now such that applications include full FBW civil transport aeroplanes.

These functionally very complex flight control systems have given the modern aeroplane hitherto unobtainable performance benefits. But nothing is free! The consequence of using such systems is the unavoidable introduction of unwanted control system dynamics. These usually manifest themselves as control phase lag and can intrude on the piloting task in an unacceptable way resulting in an aircraft with poor flying and handling qualities. This problem is still a subject of research and is very much beyond the scope of this book.

4.1.6.3 Computer Software Tools

Many computer software tools are now available which are suitable for flight dynamics analysis. Most packages are intended for control system applications, but they are ideal for handling aeronautical system problems and may be installed on a modest personal computer. Software tools used regularly by the author are listed below, but it must be appreciated that the list is by no means exhaustive, nor is it implied that the programs listed are the best or necessarily the most appropriate.

MATLAB is a very powerful control system design and analysis tool which is intended for application to systems configured in state space format. As a result all computation is handled in matrix format. Its screen graphics are good. All of the examples and problems in this book can be solved with the aid of *MATLAB.*

Simulink is a continuous simulation supplementary addition to *MATLAB*, on which it depends for its mathematical modelling. It is also a powerful tool and is easy to apply using a block diagram format for model building. It is not strictly necessary for application to the material in this book although it can be used with advantage for some examples. Its main disadvantage is its limited functional visibility since models are built using interconnecting

blocks, the functions of which are not always immediately obvious to the user. Nevertheless *Simulink* enjoys very widespread use throughout industry and academia.

MATLAB and Simulink, student version release 14 is a combined package available to registered students at low cost.

Program CC version 5 is also a very powerful control system design and analysis tool. It is capable of handling classical control problems in transfer function format as well as modern state space control problems in matrix format. The current version is very similar in use to *MATLAB* to the extent that many procedures are the same. This is not entirely surprising since the source of the underlying mathematical routines is the same for both the languages. An advantage of *Program CC* is that it was written by flight dynamicists for flight dynamicists and as a result its use becomes intuitive once the commands have been learned. Its screen graphics are good and have some flexibility of presentation. A downloadable low cost student version is available which is suitable for solving all examples and problems in this book.

Mathcad version 13 is a very powerful general purpose mathematical problem-solving tool. It is useful for repetitive calculations but it comes into its own for solving difficult non-linear equations. It is also capable of undertaking complex algebraic computations. Its screen graphics are generally very good and are very flexible. In particular, it is a valuable tool for aircraft trim and performance computations where the requirement is to solve simultaneous non-linear algebraic equations. Its use in this role is illustrated in Chapter 4.3. A low cost student version of this software is also available.

20-sim is a modern version of the traditional simulation language and it has been written to capitalise on the functionality of the modern personal computer. Models can be built up from the equations of motion, or from the equivalent matrix equations, or both. Common modules can be assigned icons of the users design and the simulation can then be constructed in a similar way to the block diagram format of *Simulink*. Its versatility is enhanced by its direct compatibility with *MATLAB*. Significant advantages are the excellent functional visibility of the problem, model building flexibility and the infinitely variable control of the model structure. Its screen graphics are excellent and it has the additional facility for direct visualisation of the modelled system running in real time. At the time of writing, the main disadvantage is the lack of a library of aerospace simulation components; however, this will no doubt be addressed as the language matures.

4.1.7 Summary

An attempt has been made in this chapter to give a broad appreciation of what flight dynamics is all about. Clearly, to produce a comprehensive text on the subject would require many volumes, assuming that it were even possible. To reiterate, the present intention is to set out the fundamental principles of the subject only. However, where appropriate, pointers are included in the text to indicate the direction in which the material in question might be developed for more advanced studies.

References

Bryan, G.H. 1911: *Stability in Aviation*. Macmillan and Co, London.
Lanchester, F.W. 1908: *Aerodonetics*. Constable and Co. Ltd, London.
MATLAB and Simulink. The Mathworks Ltd., Matrix House, Cowley Park, Cambridge, CB4 OHH. www.mathworks.co.uk/store.
Mathcad. Adept Scientific, Amor Way, Letchworth, Herts, SG6 1ZA. www.adept-science.co.uk.
Program CC. Systems Technology Inc., 13766 South Hawthorne Boulevard, Hawthorne, CA 90250-7083, USA. www.programcc.com.
20-sim. Controllab Products B.V., Drienerlolaan 5 HO-8266, 7522 NB Enschede, The Netherlands. www.20sim.com.

Aerospace Engineering Desk Reference; ISBN 9781856175753

4.2 System of Axes and Notation

Michael Cook

Before commencing the main task of developing mathematical models of the aircraft it is first necessary to put in place an appropriate and secure foundation on which to build the models. The foundation comprises a mathematical framework in which the equations of motion can be developed in an orderly and consistent way. Since aircraft have six degrees of freedom the description of their motion can be relatively complex. Therefore, motion is usually described by a number of variables which are related to a suitably chosen system of axes. In the UK the scheme of notation and nomenclature in common use is based on that developed by Hopkin (1970) and a simplified summary may be found in the appropriate ESDU (1987) data item. As far as is reasonably possible, the notation and nomenclature used throughout this book correspond with that of Hopkin (1970). By making the appropriate choice of axis systems order and consistency may be introduced to the process of model building. The importance of order and consistency in the definition of the mathematical framework cannot be over-emphasised, since with no understanding of either one, and chaos will surely follow. Only the most basic commonly used axes systems appropriate to aircraft are discussed in the following sections. In addition to the above named references a more expansive treatment may be found in Etkin (1972) or in McRuer *et al.* (1973) for example.

4.2.1 Earth Axes

Since normal atmospheric flight only is considered it is usual to measure aircraft motion with reference to an earth fixed framework. The accepted convention for defining earth axes determines that a reference point o_0 on the surface of the earth is the origin of a right-handed orthogonal system of axes ($o_0x_0y_0z_0$) where, o_0x_0 points to the north, o_0y_0 points to the east and o_0z_0 points vertically "down" along the gravity vector. Conventional earth axes are illustrated in Fig. 4.2-1.

Clearly, the plane ($o_0x_0y_0$) defines the local horizontal plane which is tangential to the surface of the earth. Thus the flight path of an aircraft flying in the atmosphere in the vicinity of the reference point o_0 may be completely described by its coordinates in the axis system. This therefore assumes a *flat earth* where the vertical is "tied" to the gravity vector. This model is quite adequate for localised flight although it is best suited to navigation and performance applications where flight path trajectories are of primary interest.

For investigations involving trans-global navigation the axis system described is inappropriate, a spherical coordinate system being preferred. Similarly, for trans-atmospheric flight involving the launch and re-entry of space vehicles a spherical coordinate system would be more appropriate. However, since in such an application the angular velocity of the earth becomes important it is necessary to define a fixed spatial axis system to which the spherical earth axis system may be referenced.

For flight dynamics applications a simpler definition of earth axes is preferred. Since short-term motion only is of interest it is perfectly adequate to assume flight above a flat earth. The most common consideration is that of motion about *straight and level* flight. Straight and level flight assumes flight in a horizontal plane at a constant altitude and, whatever the subsequent motion of the aircraft might be, the *attitude* is determined with respect to the horizontal. Referring again to Fig. 4.2-1 the horizontal plane is defined by ($o_Ex_Ey_E$) and is parallel to the plane ($o_0x_0y_0$) at the surface of the earth. The only difference is that the o_Ex_E axis points in the arbitrary direction of flight of the aircraft rather than to the north. The o_Ez_E axis points vertically down as before. Therefore, it is only necessary to place the origin o_E in the atmosphere at the most convenient point, which is frequently coincident with the origin of the aircraft body fixed axes. Earth axes ($o_Ex_Ey_Ez_E$) defined in this way are called *datum-path earth axes,* are "tied" to the earth by means of the gravity vector and provide the inertial reference frame for short-term aircraft motion.

4.2.2 Aircraft Body Fixed Axes

4.2.2.1 Generalised Body Axes

It is usual practice to define a right-handed orthogonal axis system fixed in the aircraft and constrained to move with it. Thus when the aircraft is disturbed from its initial flight condition the axes move with the airframe and the motion is quantified in terms of perturbation variables referred to the moving axes. The way in which the axes may be fixed in

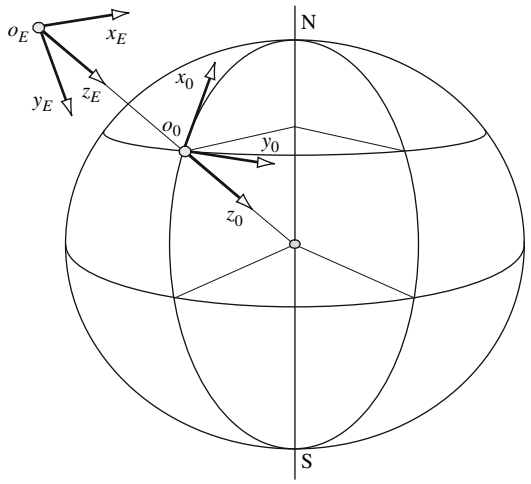


Figure 4.2-1 Conventional earth axes.

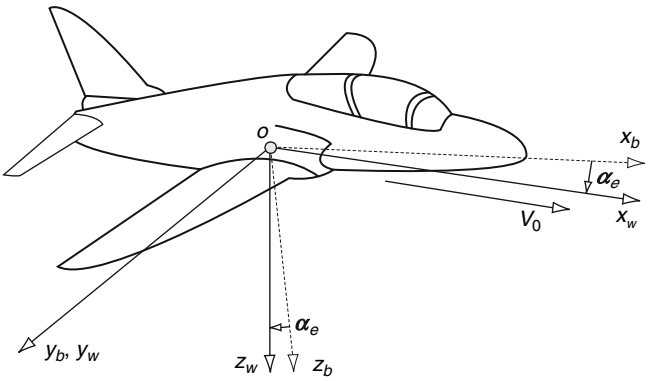


Figure 4.2-2 Moving axes systems.

the airframe is arbitrary although it is preferable to use an accepted standard orientation. The most general axis system is known as a *body axis* system ($ox_by_bz_b$) which is fixed in the aircraft as shown in Fig. 4.2-2. The (ox_bz_b) plane defines the plane of symmetry of the aircraft and it is convenient to arrange the ox_b axis such that it is parallel to the geometrical *horizontal fuselage datum.* Thus in normal flight attitudes the oy_b axis is directed to starboard and the oz_b axis is directed "downwards". The origin o of the axes is fixed at a convenient reference point in the airframe which is usually, but not necessarily, coincident with the centre of gravity (cg).

4.2.2.2 Aerodynamic, Wind or Stability Axes

It is often convenient to define a set of aircraft fixed axes such that the ox axis is parallel to the total velocity vector V_0 as shown in Fig. 4.2-2. Such axes are called *aerodynamic*, *wind* or *stability* axes. In steady symmetric flight wind axes ($ox_wy_wz_w$) are just a particular version of body axes which are rotated about the oy_b axis through the steady body incidence angle α_e until the ox_w axis aligns with the velocity vector. Thus the plane (ox_wz_w) remains the plane of symmetry of the aircraft and the oy_w and the oy_b axes are coincident. Now there is a unique value of body incidence α_e for every flight condition; therefore, the wind axes orientation in the airframe is different for every flight condition. However, for any given flight condition the wind axes orientation is defined and fixed in the aircraft at the outset and is constrained to move with it in subsequent disturbed flight. Typically the body incidence might vary in the range $-10° \leq \alpha_e \leq 20°$ over a normal flight envelope.

4.2.2.3 Perturbation Variables

The motion of the aircraft is described in terms of force, moment, linear and angular velocities and attitude resolved into components with respect to the chosen aircraft fixed axis system. For convenience it is preferable to assume

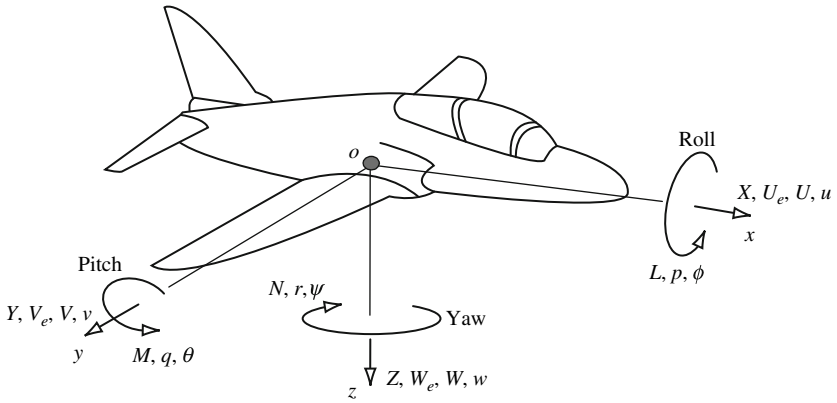


Figure 4.2-3 Motion variables notation.

a generalised *body axis* system in the first instance. Thus initially, the aircraft is assumed to be in steady rectilinear, but not necessarily level, flight when the body incidence is α_e and the steady velocity V_0 resolves into components U_e, V_e and W_e as indicated in Fig. 4.2-3. In steady non-accelerating flight the aircraft is in equilibrium and the forces and moments acting on the airframe are in balance and sum to zero. This initial condition is usually referred to as *trimmed equilibrium*.

Whenever the aircraft is disturbed from equilibrium the force and moment balance is upset and the resulting transient motion is quantified in terms of the perturbation variables. The perturbation variables are shown in Fig. 4.2-3 and summarised in Table 4.2-1.

The positive sense of the variables is determined by the choice of a right-handed axis system. Components of linear quantities, force, velocity, etc., are positive when their direction of action is the same as the direction of the axis to which they relate. The positive sense of the components of rotary quantities, moment, velocity, attitude, etc. is a right-handed rotation and may be determined as follows. Positive *roll* about the *ox* axis is such that the *oy* axis moves towards the *oz* axis, positive *pitch* about the *oy* axis is such that the *oz* axis moves towards the *ox* axis and positive *yaw* about the *oz* axis is such that the *ox* axis moves towards the *oy* axis. Therefore, positive roll is right wing down, positive pitch is nose up and positive yaw is nose to the right as seen by the pilot.

A simple description of the perturbation variables is given in Table 4.2-2. The intention is to provide some insight into the physical meaning of the many variables used in the model. Note that the components of the total linear velocity perturbations (U, V, W) are given by the sum of the steady equilibrium components and the transient perturbation components (u, v, w) thus,

$$\begin{aligned} U &= U_e + u \\ V &= V_e + v \\ W &= W_e + w \end{aligned} \tag{4.2.1}$$

4.2.2.4 Angular Relationships in Symmetric Flight

Since it is assumed that the aircraft is in steady rectilinear, but not necessarily level flight, and that the axes fixed in the aircraft are body axes then it is useful to relate the steady and perturbed angles as shown in Fig. 4.2-4.

With reference to Fig. 4.2-4, the steady velocity vector V_0 defines the flight path and γ_e is the steady flight path angle. As before, α_e is the steady body incidence and θ_e is the steady pitch attitude of the aircraft. The relative angular

Table 4.2-1 Summary of motion variables.

	Trimmed equilibrium			Perturbed		
Aircraft axis	*ox*	*oy*	*oz*	*ox*	*oy*	*oz*
Force	0	0	0	*X*	*Y*	*Z*
Moment	0	0	0	*L*	*M*	*N*
Linear velocity	U_e	V_e	W_e	*U*	*V*	*W*
Angular velocity	0	0	0	*P*	*q*	*r*
Attitude	0	θ_e	0	ϕ	θ	ψ

Table 4.2-2 The perturbation variables.

X	Axial "drag" force	Sum of the components of
Y	Side force	aerodynamic, thrust and
Z	Normal "lift" force	weight forces
L	Rolling moment	Sum of the components of
M	Pitching moment	aerodynamic, thrust and
N	Yawing moment	weight moments
p	Roll rate	Components of angular
q	Pitch rate	velocity
r	Yaw rate	
U	Axial velocity	Total linear velocity
V	Lateral velocity	components of the *cg*
W	Normal velocity	

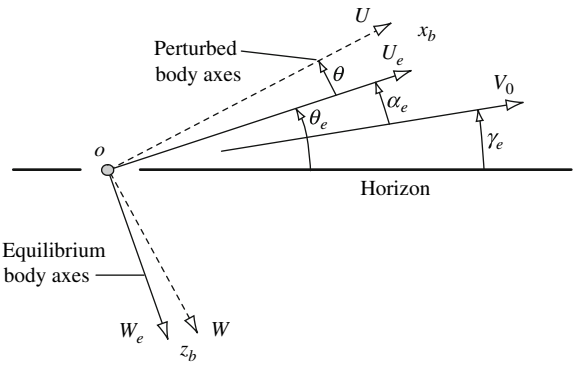


Figure 4.2-4 Generalised body axes in symmetric flight.

change in a perturbation is also shown in Fig. 4.2-4 where it is implied that the axes have moved with the airframe and the motion is viewed at some instant during the disturbance. Thus the steady flight path angle is given by

$$\gamma_e = \theta_e - \alpha_e \tag{4.2.2}$$

In the case when the aircraft fixed axes are wind axes rather than body axes then,

$$\alpha_e = 0 \tag{4.2.3}$$

and in the special case when the axes are wind axes and when the initial condition is level flight,

$$\alpha_e = \theta_e = 0 \tag{4.2.4}$$

It is also useful to note that the perturbation in pitch attitude θ and the perturbation in body incidence α are the same; thus, it is convenient to write,

$$\tan(\alpha_e + \theta) \equiv \tan(\alpha_e + \alpha) = \frac{W}{U} \equiv \frac{W_e + w}{U_e + u} \tag{4.2.5}$$

4.2.2.5 Choice of Axes

Having reviewed the definition of aircraft fixed axis systems an obvious question must be: *when is it appropriate to use wind axes and when is it appropriate to use body axes*? The answer to this question depends on the use to which the equations of motion are to be put. The best choice of axes simply facilitates the analysis of the equations of motion. When starting from first principles it is preferable to use generalised body axes since the resulting equations can cater for most applications. It is then reasonably straightforward to simplify the equations to a wind axis form if the application warrants it. On the other hand, to extend wind axis-based equations to cater for the more general case is not as easy.

When dealing with numerical data for an existing aircraft it is not always obvious which axis system has been used in the derivation of the model. However, by reference to equation (4.2.3) or (4.2.4) and the quoted values of α_e and θ_e it should become obvious which axis system has been used.

When it is necessary to make experimental measurements in an actual aircraft, or in a model, which are to be used subsequently in the equations of motion it is preferable to use a generalised body axis system. Since the measuring equipment is installed in the aircraft its location is precisely known in terms of body axis coordinates which, therefore, determines the best choice of axis system. In a similar way, most aerodynamic measurements and computations are referenced to the free stream velocity vector. For example, in wind tunnel work the obvious reference is the tunnel axis which is coincident with the velocity vector. Thus, for aerodynamic investigations involving the equations of motion a wind axis reference is to be preferred. Traditionally all aerodynamic data for use in the equations of motion are referenced to wind axes.

Thus, to summarise, it is not particularly important which axis system is chosen provided it models the flight condition to be investigated, the end result does not depend on the choice of axis system. However, when compiling data for use in the equations of motion of an aircraft it is quite common for some data to be referred to wind axes and for some data to be referred to body axes. It therefore becomes necessary to have available the mathematical tools for transforming data between different reference axes.

4.2.3 Euler Angles and Aircraft Attitude

The angles defined by the right-handed rotation about the three axes of a right-handed system of axes are called *Euler angles*. The sense of the rotations and the order in which the rotations are considered about the three axes in turn are very important since angles do not obey the commutative law. The attitude of an aircraft is defined as the angular orientation of the airframe fixed axes with respect to earth axes. Attitude angles, therefore, are a particular application of Euler angles. With reference to Fig. 4.2-5 ($ox_0y_0z_0$) are datum or reference axes and ($ox_3y_3z_3$) are aircraft fixed axes, either generalised body axes or wind axes. The attitude of the aircraft, with respect to the datum axes, may be established by considering the rotation about each axis in turn required to bring ($ox_3y_3z_3$) into coincidence with ($ox_0y_0z_0$). Thus, first rotate about $ox_3 ox_3$ through the *roll* angle ϕ to ($ox_2y_2z_2$). Second, rotate about oy_2 through the *pitch* angle θ to ($ox_1y_1z_1$) and third, rotate about oz_1 through the *yaw* angle ψ to ($ox_0y_0z_0$). Clearly, when the attitude of the aircraft is considered with respect to earth axes then ($ox_0y_0z_0$) and ($ox_Ey_Ez_E$) are coincident.

4.2.4 Axes Transformations

It is frequently necessary to transform motion variables and other parameters from one system of axes to another. Clearly, the angular relationships used to describe attitude may be generalised to describe the angular orientation of one set of axes with respect to another. A typical example might be to transform components of linear velocity from aircraft wind axes to body axes. Thus, with reference to Fig. 4.2-5, ($ox_0y_0z_0$) may be used to describe the velocity components in wind axes, ($ox_3y_3z_3$) may be used to describe the components of velocity in body axes and the angles (ϕ, θ, ψ) then describe the generalised angular orientation of one set of axes with respect to the other. It is usual to retain the angular description *of roll, pitch* and *yaw* although the angles do not necessarily describe attitude strictly in accordance with the definition given in Section 4.2.3.

4.2.4.1 Linear Quantities Transformation

Let, for example, (ox_3, oy_3, oz_3) represent components of a linear quantity in the axis system ($ox_3y_3z_3$) and let (ox_0, oy_0, oz_0) represent components of the same linear quantity transformed into the axis system ($ox_0y_0z_0$). The linear

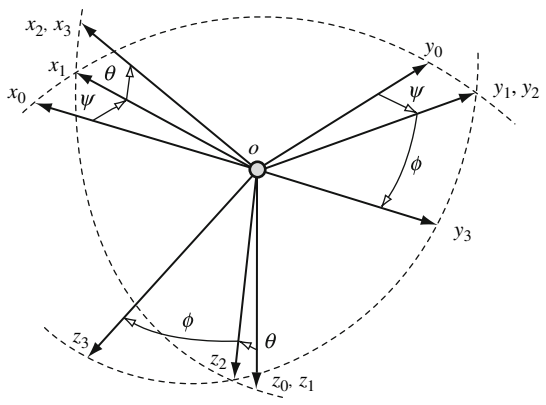


Figure 4.2-5 The Euler angles.

quantities of interest would be, for example, acceleration, velocity, displacement, etc. Resolving through each rotation in turn and in the correct order then, with reference to Fig. 4.2-5, it may be shown that:

(i) after *rolling* about ox_3 through the angle ϕ,

$$\begin{aligned} ox_3 &= ox_2 \\ oy_3 &= oy_2 \cos\phi + oz_2 \sin\phi \\ oz_3 &= -oy_2 \sin\phi + oz_2 \cos\phi \end{aligned} \tag{4.2.6}$$

Alternatively, writing equation (4.2.6) in the more convenient matrix form,

$$\begin{bmatrix} ox_3 \\ oy_3 \\ oz_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} ox_2 \\ oy_2 \\ oz_2 \end{bmatrix} \tag{4.2.7}$$

(ii) similarly, after *pitching* about oy_2 through the angle θ,

$$\begin{bmatrix} ox_2 \\ oy_2 \\ oz_2 \end{bmatrix} = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} ox_1 \\ oy_1 \\ oz_1 \end{bmatrix} \tag{4.2.8}$$

(iii) and after *yawing* about oz_1 through the angle ψ,

$$\begin{bmatrix} ox_1 \\ oy_1 \\ oz_1 \end{bmatrix} = \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ox_0 \\ oy_0 \\ oz_0 \end{bmatrix} \tag{4.2.9}$$

By repeated substitution equations (4.2.7), (4.2.8) and (4.2.9) may be combined to give the required transformation relationship

$$\begin{bmatrix} ox_3 \\ oy_3 \\ oz_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\psi & \sin\psi & 0 \\ -\sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} ox_0 \\ oy_0 \\ oz_0 \end{bmatrix} \tag{4.2.10}$$

or

$$\begin{bmatrix} ox_3 \\ oy_3 \\ oz_3 \end{bmatrix} = \mathbf{D} \begin{bmatrix} ox_0 \\ oy_0 \\ oz_0 \end{bmatrix} \tag{4.2.11}$$

where the *direction cosine matrix* $\mathbf{D}$ is given by

$$\mathbf{D} = \begin{bmatrix} \cos\theta\cos\psi & \cos\theta\sin\psi & -\sin\theta \\ \sin\phi\sin\theta\cos\psi & \sin\phi\sin\theta\sin\psi & \sin\phi\cos\theta \\ -\cos\phi\sin\psi & +\cos\phi\cos\psi & \\ \cos\phi\sin\theta\cos\psi & \cos\phi\sin\theta\sin\psi & \cos\phi\cos\theta \\ +\sin\phi\sin\psi & -\sin\phi\cos\psi & \end{bmatrix} \tag{4.2.12}$$

As shown, equation (4.2.11) transforms linear quantities from $(ox_0y_0z_0)$ to $(ox_3y_3z_3)$. By inverting the direction cosine matrix $\mathbf{D}$ the transformation from $(ox_3y_3z_3)$ to $(ox_0y_0z_0)$ is obtained as given by equation (4.2.13):

$$\begin{bmatrix} ox_0 \\ oy_0 \\ oz_0 \end{bmatrix} = \mathbf{D}^{-1} \begin{bmatrix} ox_3 \\ oy_3 \\ oz_3 \end{bmatrix} \tag{4.2.13}$$

Example 4.2.1

To illustrate the use of equation (4.2.11) consider the very simple example in which it is required to resolve the velocity of the aircraft through both the incidence angle and the sideslip angle into aircraft axes. The situation prevailing is assumed to be steady and is shown in Fig. 4.2-6.

The axes $(oxyz)$ are generalised aircraft body axes with velocity components U_e, V_e and W_e respectively. The free stream velocity vector is V_0 and the angles of incidence and sideslip are α_e and β_e respectively. With reference to

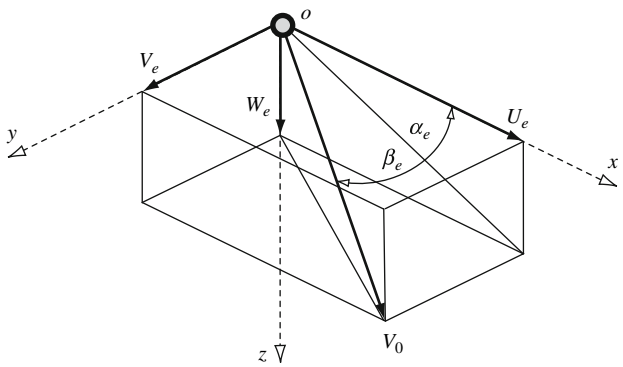


Figure 4.2-6 Resolution of velocity through incidence and sideslip angles.

equation (4.2.11), axes $(oxyz)$ correspond with axes $(ox_3y_3z_3)$ and V_0 corresponds with ox_0 of axes $(ox_0y_0z_0)$; therefore, the following vector substitutions may be made:

$$(ox_0, oy_0, oz_0) = (V_0, 0, 0) \quad \text{and} \quad (ox_3, oy_3, oz_3) = (U_e, V_e, W_e)$$

and the angular correspondence means that the following substitution may be made:

$$(\phi, \theta, \psi) = (0, \alpha_e, -\beta_e)$$

Note that a positive sideslip angle is equivalent to a negative yaw angle. Thus making the substitutions in equation (4.2.9),

$$\begin{bmatrix} U_e \\ V_e \\ W_e \end{bmatrix} = \begin{bmatrix} \cos\alpha_e \cos\beta_e & -\cos\alpha_e \sin\beta_e & -\sin\alpha_e \\ \sin\beta_e & \cos\beta_e & 0 \\ \sin\alpha_e \cos\beta_e & -\sin\alpha_e \sin\beta_e & \cos\alpha_e \end{bmatrix} \begin{bmatrix} V_0 \\ 0 \\ 0 \end{bmatrix} \tag{4.2.14}$$

Or, equivalently,

$$\begin{aligned} U_e &= V_0 \cos\alpha_e \cos\beta_e \\ V_e &= V_0 \sin\beta_e \\ W_e &= V_0 \sin\alpha_e \cos\beta_e \end{aligned} \tag{4.2.15}$$

Example 4.2.2

Another very useful application of the direction cosine matrix is to calculate height perturbations in terms of aircraft motion. Equation (4.2.13) may be used to relate the velocity components in aircraft axes to the corresponding components in earth axes as follows:

$$\begin{bmatrix} U_E \\ V_E \\ W_E \end{bmatrix} = \mathbf{D}^{-1} \begin{bmatrix} U \\ V \\ W \end{bmatrix}$$

$$= \begin{bmatrix} \cos\psi\cos\theta & \cos\psi\sin\theta\sin\phi - \sin\psi\cos\phi & \cos\psi\sin\theta\cos\phi + \sin\psi\sin\phi \\ \sin\psi\cos\theta & \sin\psi\sin\theta\sin\phi + \cos\psi\cos\phi & \sin\psi\sin\theta\cos\phi - \cos\psi\sin\phi \\ -\sin\theta & \cos\theta\sin\phi & \cos\theta\cos\phi \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix} \tag{4.2.16}$$

where U_E, V_E and W_E are the perturbed total velocity components referred to earth axes. Now, since height is measured positive in the "upwards" direction, the rate of change of height due to the perturbation in aircraft motion is given by

$$\dot{h} = -W_E$$

Hence, from equation (4.2.16),

$$\dot{h} = U \sin\theta - V \cos\theta \sin\phi - W\cos\theta \cos\phi \tag{4.2.17}$$

4.2.4.2 Angular Velocities Transformation

Probably the most useful angular quantities transformation relates the angular velocities p, q, r of the aircraft fixed axes to the resolved components of angular velocity, the attitude rates $\dot{\phi}, \dot{\theta}, \dot{\psi}$ with respect to datum axes. The easiest way to deal with the algebra of this transformation whilst retaining a good grasp of the physical implications is to superimpose the angular rate vectors on to the axes shown in Fig. 4.2-5, and the result of this is shown in Fig. 4.2-7.

The angular *body rates* p, q, r are shown in the aircraft axes ($ox_3y_3z_3$) then, considering each rotation in turn necessary to bring the aircraft axes into coincidence with the datum axes ($ox_0y_0z_0$). First, *roll* about ox_3ox_3 through the angle ϕ with angular velocity $\dot{\phi}$. Second, *pitch* about oy_2 through the angle θ with angular velocity $\dot{\theta}$. And third, *yaw* about oz_1 through the angle ψ with angular velocity $\dot{\psi}$. Again, it is most useful to refer the attitude rates to earth axes in which case the datum axes ($ox_0y_0z_0$) are coincident with earth axes ($o_Ex_Ey_Ez_E$). The *attitude rate* vectors are clearly shown in Fig. 4.2-7. The relationship between the aircraft body rates and the attitude rates, referred to datum axes, is readily established as follows:

(i) *Roll rate* p is equal to the sum of the components of $\dot{\phi}, \dot{\theta}, \dot{\psi}$ resolved along ox_3,

$$p = \dot{\phi} - \dot{\psi}\sin\theta \tag{4.2.18}$$

(ii) *Pitch rate* q is equal to the sum of the components of $\dot{\phi}, \dot{\theta}, \dot{\psi}$ resolved along oy_3,

$$q = \dot{\theta}\cos\phi + \dot{\psi}\sin\phi\cos\theta \tag{4.2.19}$$

(iii) *Yaw rate* r is equal to the sum of the components of $\dot{\phi}, \dot{\theta}, \dot{\psi}$ resolved along oz_3,

$$r = \dot{\psi}\cos\phi\cos\theta - \dot{\theta}\sin\phi \tag{4.2.20}$$

Equations (4.2.18), (4.2.19) and (4.2.20) may be combined into the convenient matrix notation

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} 1 & 0 & -\sin\theta \\ 0 & \cos\phi & \sin\phi\cos\theta \\ 0 & -\sin\phi & \cos\phi\cos\theta \end{bmatrix} \begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} \tag{4.2.21}$$

and the inverse of equation (4.2.21) is

$$\begin{bmatrix} \dot{\phi} \\ \dot{\theta} \\ \dot{\psi} \end{bmatrix} = \begin{bmatrix} 1 & \sin\phi\tan\theta & \cos\phi\tan\theta \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi\sec\theta & \cos\phi\sec\theta \end{bmatrix} \begin{bmatrix} p \\ q \\ r \end{bmatrix} \tag{4.2.22}$$

When the aircraft perturbations are small, such that (ϕ, θ, ψ) may be treated as small angles, equations (4.2.21) and (4.2.22) may be approximated by

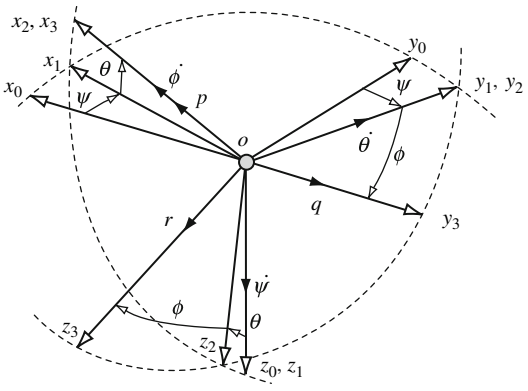


Figure 4.2-7 Angular rates transformation.

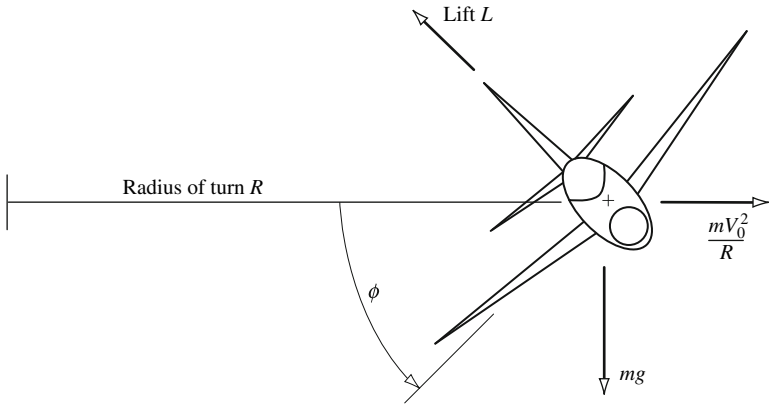


Figure 4.2-8 Aircraft in a steady banked turn.

$$\begin{aligned} p &= \dot{\phi} \\ q &= \dot{\theta} \\ r &= \dot{\psi} \end{aligned} \tag{4.2.23}$$

Example 4.2.3

To illustrate the use of the angular velocities transformation, consider the situation when an aircraft is flying in a steady level coordinated turn at a speed of 250 m/s at a bank angle of 60°. It is required to calculate the turn rate $\dot{\psi}$, the yaw rate r and the pitch rate q. The forces acting on the aircraft are shown in Fig. 4.2-8.

By resolving the forces acting on the aircraft vertically and horizontally and eliminating the lift L between the two resulting equations it is easily shown that the radius of turn is given by

$$R = \frac{V_0^2}{g \tan \phi} \tag{4.2.24}$$

The time to complete one turn is given by

$$t = \frac{2\pi R}{V_0} = \frac{2\pi V_0}{g \tan \phi} \tag{4.2.25}$$

therefore the rate of turn is given by

$$\dot{\psi} = \frac{2\pi}{t} = \frac{g \tan \phi}{V_0} \tag{4.2.26}$$

Thus, $\dot{\psi} = 0.068$ rad/s. For the conditions applying to the turn, $\dot{\phi} = \dot{\theta} = \theta = 0$ and thus equation (4.2.21) may now be used to find the values of r and q:

$$\begin{bmatrix} p \\ q \\ r \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos 60^\circ & \sin 60^\circ \\ 0 & -\sin 60^\circ & \cos 60^\circ \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \dot{\psi} \end{bmatrix}$$

Therefore, $p = 0$, $q = 0.059$ rad/s and $r = 0.034$ rad/s. Note that p, q and r are the angular velocities that would be measured by rate gyros fixed in the aircraft with their sensitive axes aligned with the ox, oy and oz aircraft axes respectively.

4.2.5 Aircraft Reference Geometry

The description of the geometric layout of an aircraft is an essential part of the mathematical modelling process. Forthe purposes of flight dynamics analysis it is convenient that the geometry of the aircraft can be adequately described by a small number of dimensional reference parameters which are defined and illustrated in Fig. 4.2-9.

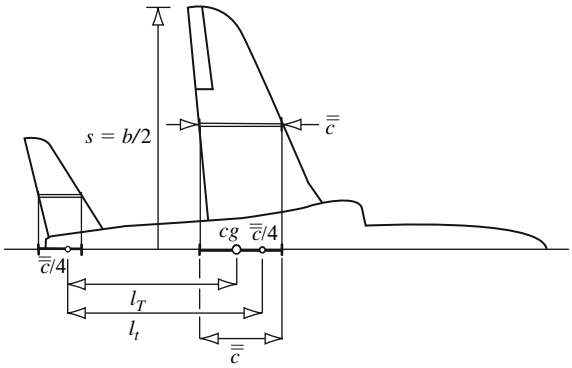


Figure 4.2-9 Longitudinal reference geometry.

4.2.5.1 Wing Area

The reference area is usually the gross plan area of the wing, including that part within the fuselage, and is denoted S:

$$S = b\bar{c} \tag{4.2.27}$$

where b is the wing span and $\bar{c}$ is the standard mean chord of the wing.

4.2.5.2 Mean Aerodynamic Chord

The *mean aerodynamic chord* of the wing (*mac*) is denoted $\bar{\bar{c}}$ and is defined as

$$\bar{\bar{c}} = \frac{\int_{-s}^{s} c_y^2 \mathrm{d}y}{\int_{-s}^{s} c_y \mathrm{d}y} \tag{4.2.28}$$

The reference *mac* is located on the centre line of the aircraft by projecting $\bar{\bar{c}}$ from its spanwise location as shown in Fig. 4.2-9. Thus for a swept wing the leading edge of the *mac* lies aft of the leading edge of the root chord of the wing. The *mac* represents the location of the root chord of a rectangular wing which has the same aerodynamic influence on the aircraft as the actual wing. Traditionally *mac* is used in stability and control studies since a number of important aerodynamic reference centres are located on it.

4.2.5.3 Standard Mean Chord

The *standard mean chord* of the wing (*smc*) is effectively the same as the geometric mean chord and is denoted $\bar{c}$. For a wing of symmetric planform it is defined as

$$\bar{c} = \frac{\int_{-s}^{s} c_y \mathrm{d}y}{\int_{-s}^{s} \mathrm{d}y} \tag{4.2.29}$$

where $s = b/2$ is the semi-span and c_y is the local chord at spanwise coordinate y. For a straight tapered wing equation (4.2.29) simplifies to

$$\bar{c} = \frac{S}{b} \tag{4.2.30}$$

The reference *smc* is located on the centre line of the aircraft by projecting $\bar{c}$ from its spanwise location in the same way that the *mac* is located. Thus for a swept wing the leading edge of the *smc* also lies aft of the leading edge of the root chord of the wing. The *smc* is the mean chord preferred by aircraft designers since it relates very simply to the geometry of the aircraft. For most aircraft the *smc* and *mac* are sufficiently similar in length and location that they are practically interchangeable. It is quite common to find references that quote a mean chord without specifying which. This is not good practice although the error incurred by assuming the wrong chord is rarely serious. However, the reference chord used in any application should always be clearly defined at the outset.

4.2.5.4 Aspect Ratio

The aspect ratio of the aircraft wing is a measure of its spanwise slenderness and is denoted A and is defined as follows:

$$A = \frac{b^2}{S} = \frac{b}{\bar{c}} \tag{4.2.31}$$

4.2.5.5 Centre of Gravity Location

The centre of gravity, *cg*, of an aircraft is usually located on the reference chord as indicated in Fig. 4.2-9. Its position is quoted as a fraction of $\bar{\bar{c}}$ (or $\bar{c}$), denoted h, and is measured from the leading edge of the reference chord as shown. The *cg* position varies as a function of aircraft loading, the typical variation being in the range 10–40% of $\bar{\bar{c}}$. Or, equivalently, $0.1 \leq h \leq 0.4$.

4.2.5.6 Tail Moment Arm and Tail Volume Ratio

The *mac* of the horizontal tailplane, or foreplane, is defined and located in the airframe in the same way as the *mac* of the wing as indicated in Fig. 4.2-9. The wing and tailplane aerodynamic forces and moments are assumed to act at their respective aerodynamic centres which, to a good approximation, lie at the quarter chord points of the *mac* of the wing and tailplane respectively. The tail moment arm l_T is defined as the longitudinal distance between the centre of gravity and the aerodynamic centre of the tailplane as shown in Fig. 4.2-9. The *tail volume ratio* $\overline{V}_T$ is an important geometric parameter and is defined as

$$\overline{V}_T = \frac{S_T l_T}{S\bar{\bar{c}}} \tag{4.2.32}$$

where S_T is the gross area of the tailplane and *mac* $\bar{\bar{c}}$ is the longitudinal reference length. Typically, the tail volume ratio has a value in the range $0.5 \leq \overline{V}_T \leq 1.3$ and is a measure of the aerodynamic effectiveness of the tailplane as a stabilising device.

Sometimes, especially in stability and control studies, it is convenient to measure the longitudinal tail moment about the aerodynamic centre of the *mac* of the wing. In this case the tail moment arm is denoted l_t, as shown in Fig. 4.2-9, and a slightly modified tail volume ratio is defined.

4.2.5.7 Fin Moment Arm and Fin Volume Ratio

The *mac* of the fin is defined and located in the airframe in the same way as the *mac* of the wing as indicated in Fig. 4.2-10. As for the tailplane, the fin moment arm l_F is defined as the longitudinal distance between the centre of gravity and the aerodynamic centre of the fin as shown in Fig. 4.2-10. The *fin volume ratio* $\overline{V}_F$ is also an important geometric parameter and is defined as

$$\overline{V}_F = \frac{S_F l_F}{Sb} \tag{4.2.33}$$

where S_F is the gross area of the fin and the wing span b is the lateral–directional reference length. Again, the fin volume ratio is a measure of the aerodynamic effectiveness of the fin as a directional stabilising device.

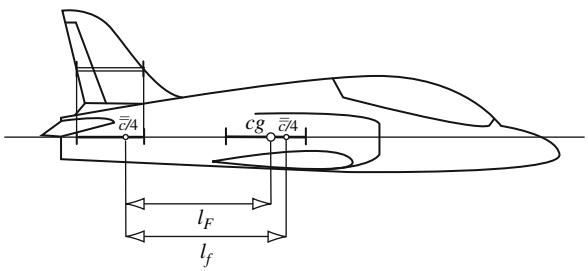


Figure 4.2-10 Fin moment arm.

As stated above it is sometimes convenient to measure the longitudinal moment of the aerodynamic forces acting at the fin about the aerodynamic centre of the *mac* of the wing. In this case the fin moment arm is denoted l_f as shown in Fig. 4.2-10.

4.2.6 Controls Notation

4.2.6.1 Aerodynamic Controls

Sometimes it appears that some confusion exists with respect to the correct notation applying to aerodynamic controls, especially when unconventional control surfaces are used. Hopkin (1970) defines a notation which is intended to be generally applicable, but since a very large number of combinations of control motivators is possible the notation relating to control inceptors may become ill defined and hence application dependent. However, for the conventional aircraft there is a universally accepted notation, which accords with Hopkin (1970), and it is simple to apply. Generally, a positive *control action* by the pilot gives rise to a positive aircraft response, whereas a positive *control surface displacement* gives rise to a negative aircraft response. Thus:

(i) **In roll**: positive right push force on the stick ⇒ positive stick displacement ⇒ right aileron up and left aileron down (negative mean) ⇒ right wing down roll response (positive).
(ii) **In pitch**: positive pull force on the stick ⇒ positive aft stick displacement ⇒ elevator trailing edge up (negative) ⇒ nose up pitch response (positive).
(iii) **In yaw**: positive push force on the right rudder pedal ⇒ positive rudder bar displacement ⇒ rudder trailing edge displaced to the right (negative) ⇒ nose to the right yaw response (positive).

Roll and pitch control stick displacements are denoted δ_ξ and δ_η respectively and rudder pedal displacement is denoted δ_ζ. Aileron, elevator and rudder surface displacements are denoted ξ, η and ζ respectively as indicated in Fig. 4.2-11. It should be noted that since ailerons act differentially the displacement ξ is usually taken as the mean value of the separate displacements of each aileron.

4.2.6.2 Engine Control

Engine thrust τ is controlled by throttle lever displacement ε. Positive throttle lever displacement is usually in the forward push sense and results in a positive increase in thrust. For a turbojet engine the relationship between thrust and throttle lever angle is approximated by a simple first order lag transfer function:

$$\frac{\tau(s)}{\varepsilon(s)} = \frac{k_\tau}{(1 + sT_\tau)} \tag{4.2.34}$$

where k_τ is a suitable gain constant and T_τ is the lag time constant which is typically of the order of 2–3 s.

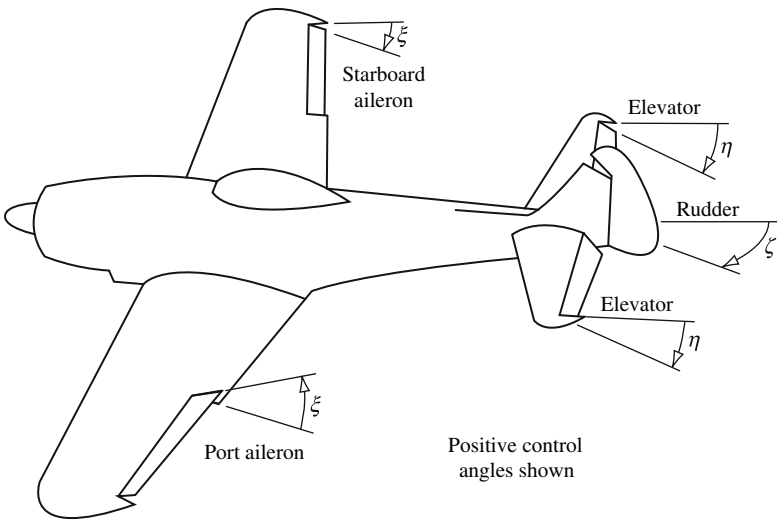


Figure 4.2-11 Aerodynamic controls notation.

4.2.7 Aerodynamic Reference Centres

With reference to Fig. 4.2-12, the *centre of pressure, cp*, of an aerofoil, wing or complete aircraft is the point at which the resultant aerodynamic force F acts. It is usual to resolve the force into the *lift* component perpendicular to the velocity vector and the *drag* component parallel to the velocity vector, denoted L and D respectively in the usual way. The *cp* is located on the *mac* and thereby determines an important aerodynamic reference centre.

Now simple theory establishes that the resultant aerodynamic force F generated by an aerofoil comprises two components, that due to camber F_c and that due to angle of attack F_α, both of which resolve into lift and drag forces as indicated. The aerodynamic force due to camber is constant and acts at the midpoint of the aerofoil chord and for a symmetric aerofoil section this force is zero. The aerodynamic force due to angle of attack acts at the quarter chord point and varies directly with angle of attack at angles below the stall. This also explains why the zero lift angle of attack of a cambered aerofoil is usually a small negative value since, at this condition, the lift due to camber is equal and opposite to the lift due to angle of attack. Thus at low speeds, when the angle of attack is generally large, most of the aerodynamic force is due to the angle of attack-dependent contribution and the *cp* is nearer to the quarter chord point. On the other hand, at high speeds, when the angle of attack is generally small, a larger contribution to the aerodynamic force is due to the camber-dependent component and the *cp* is nearer to the midpoint of the chord. Thus, in the limit the *cp* of an aerofoil generally lies between the quarter chord and mid-chord points. More generally, the interpretation for an aircraft recognises that the *cp* moves as a function of angle of attack, Mach number and configuration. For example, at low angles of attack and high Mach numbers the *cp* tends to move aft and *vice versa*. Consequently the *cp* is of limited use as an aerodynamic reference point in stability and control studies. It should be noted that the *cp* of the complete aircraft in trimmed equilibrium flight corresponds with the *controls fixed neutral point* $h_n\bar{\bar{c}}$ which is discussed in Chapter 4.3.

If, instead of the *cp*, another *fixed* point on the *mac* is chosen as an aerodynamic reference point then, at this point, the total aerodynamic force remains the same but is accompanied by a pitching moment about the point. Clearly, the most convenient reference point on the *mac* is the quarter chord point since the pitching moment is the moment of the aerodynamic force due to camber and remains constant with variation in angle of attack. This point is called the *aerodynamic centre*, denoted *ac*, and at low Mach numbers lies at, or very close to, the quarter chord point, $\bar{\bar{c}}/4$. It is for this reason that the *ac*, or equivalently, the quarter chord point of the reference chord is preferred as a reference point. The corresponding equivalent aerofoil model is shown in Fig. 4.2-12. Since the *ac* remains essentially fixed in position during small perturbations about a given flight condition, and since the pitching moment is nominally constant about the *ac*, it is used as a reference point in stability and control studies. It is important to appreciate that as the flight condition Mach number is increased so the *ac* moves aft and in supersonic flow conditions it is located at, or very near to, $\bar{\bar{c}}/2$.

The definition of aerodynamic centre given above applies most strictly to the location of the *ac* on the chord of an aerofoil. However, it also applies reasonably well to its location on the *mac* of a wing and is also used extensively for locating the *ac* on the *mac* of a wing–body combination without too much loss of validity. It should be appreciated that the complex aerodynamics of a wing and body combination might result in an *ac* location which is not at the quarter chord point although, typically, it would not be too far removed from that point.

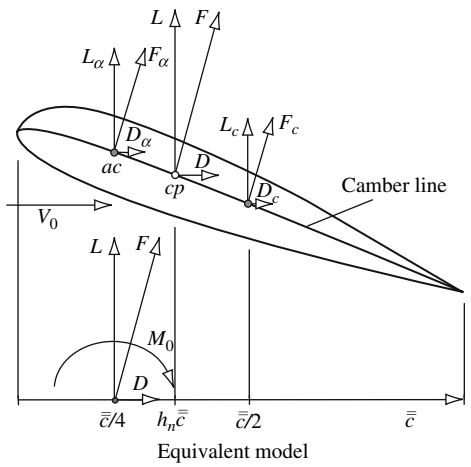


Figure 4.2-12 Aerodynamic reference centres.

References

ESDU 1987: *Introduction to Aerodynamic Derivatives, Equations of Motion and Stability.* Engineering Sciences Data Unit, Data Item No. 86021. Aerodynamics Series, Vol. 9a, Stability of Aircraft. Engineering Sciences Data, ESDU International Ltd., 27 Corsham Street, London. www.esdu.com.

Etkin, B. 1972: *Dynamics of Atmospheric Flight*. New York: John Wiley and Sons, Inc.

Hopkin, H.R. 1970: *A Scheme of Notation and Nomenclature for Aircraft Dynamics and Associated Aerodynamics*. Aeronautical Research Council, Reports and Memoranda No. 3562. Her Majesty's Stationery Office, London.

McRuer, D. Ashkenas, I. and Graham, D. 1973: *Aircraft Dynamics and Automatic Control*. Princeton, NJ: Princeton University Press.

Problems

1. A tailless aircraft of 9072 kg mass has a delta wing with aspect ratio 1 and area 37 m^2. Show that the aerodynamic mean chord

$$\bar{\bar{c}} = \frac{\int_0^{\frac{b}{2}} c^2 \mathrm{d}y}{\int_0^{\frac{b}{2}} c \mathrm{d}y}$$

of a delta wing is two-thirds of its root chord and that for this wing it is 8.11 m.

(CU 1983)

2. With the aid of a diagram describe the axes systems used in aircraft stability and control analysis. State the conditions when the use of each axis system might be preferred.

(CU 1982)

3. Show that in a longitudinal symmetric small perturbation the components of aircraft weight resolved into the ox and oz axes are given by

$$X_g = -mg\theta \cos\theta_e - mg \sin\theta_e$$
$$Z_g = mg \cos\theta_e - mg\theta \sin\theta_e$$

where θ is the perturbation in pitch attitude and θ_e is the equilibrium pitch attitude.

(CU 1982)

4. With the aid of a diagram showing a generalised set of aircraft body axes, define the parameter notation used in the mathematical modelling of aircraft motion.

(CU 1982)

5. In the context of aircraft motion, what are the Euler angles? If the standard right-handed aircraft axis set is rotated through pitch θ and yaw ψ angles only, show that the initial vector quantity (x_0, y_0, z_0) is related to the transformed vector quantity (x, y, z) as follows:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} \cos\theta\cos\psi & \cos\theta\sin\psi & -\sin\theta \\ -\sin\psi & \cos\psi & 0 \\ \sin\theta\cos\psi & \sin\theta\sin\psi & \cos\theta \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}$$

(CU 1982)

6. Define the span, gross area, aspect ratio and mean aerodynamic chord of an aircraft wing.

(CU 2001)

7. Distinguish between the centre of pressure and the aerodynamic centre of an aerofoil. Explain why the pitching moment about the quarter chord point of an aerofoil is nominally constant in subsonic flight.

(CU 2001)

Aerospace Engineering Desk Reference; ISBN 9781856175753

4.3 Static Equilibrium and Trim

Michael Cook

4.3.1 Trim Equilibrium

4.3.1.1 Preliminary Considerations

In normal flight it is usual for the pilot to adjust the controls of an aircraft such that on releasing the controls it continues to fly at the chosen flight condition. By this means the pilot is relieved of the tedium of constantly maintaining the control inputs and the associated control forces which may be tiring. The aircraft is then said to be *trimmed,* and the trim state defines the initial condition about which the dynamics of interest may be studied. Thus all aircraft are equipped with the means for pre-setting or adjusting the *datum* or *trim* setting of the primary control surfaces. The ailerons, elevator and rudder are all fitted with *trim tabs* which, in all except the smallest aircraft, may be adjusted from the cockpit in flight. However, all aircraft are fitted with a continuously adjustable elevator trim tab. It is an essential requirement that an aircraft must be stable if it is to remain in equilibrium following trimming. In particular, the static stability characteristics about all three axes largely determine the *trimmability* of an aircraft. Thus static stability is concerned with the control actions required to establish equilibrium and with the characteristics required to ensure that the aircraft remains in equilibrium. Dynamic stability is also important of course, and largely determines the characteristics of the transient motion, following a disturbance, about a trimmed flight condition.

The object of trimming is to bring the forces and moments acting on the aircraft into a state of equilibrium. That is the condition when the axial, normal and side forces, and the roll, pitch and yaw moments are all zero. The force balance is often expressed approximately as the requirement for the lift to equal the weight and the thrust to equal the drag. Provided that the aircraft is stable it will then stay in equilibrium until it is disturbed by pilot control inputs or by external influences such as turbulence. The transient motion following such a disturbance is characterised by the dynamic stability characteristics and the stable aircraft will eventually settle into its equilibrium state once more. The maintenance of trimmed equilibrium requires the correct simultaneous adjustment of the main flight variables in all six degrees of freedom and is dependent on airspeed or Mach number, flight path angle, airframe configuration, weight and centre of gravity (*cg*) position. As these parameters change during the course of a typical flight so trim adjustments are made as necessary. Fortunately, the task of trimming an aircraft is not as challenging as it might at first seem. The symmetry of a typical airframe confers symmetric aerodynamic properties on the airframe that usually reduces the task to that of longitudinal trim only. Lateral–directional trim adjustments are only likely to be required when the aerodynamic symmetry is lost, due to loss of an engine in a multi-engined aircraft, for example.

Lateral–directional stability is designed-in to most aircraft and ensures that in roll the aircraft remains at *wings level* and that in yaw it tends to *weathercock* into the wind when the ailerons and rudder are at their zero or datum positions. Thus, under normal circumstances the aircraft will naturally seek lateral–directional equilibrium without interference by the pilot. This applies even when significant changes are made to airspeed, configuration, weight and *cg* position, for example, since the symmetry of the airframe is retained throughout. However, such variations in flight condition can lead to dramatic changes in longitudinal trim.

Longitudinal trim involves the simultaneous adjustment of elevator angle and thrust to give the required airspeed and flight path angle for a given airframe configuration. Equilibrium is achievable only if the aircraft is longitudinally stable and the control actions to trim depend on the *degree* of longitudinal static stability. Since the longitudinal flight condition is continuously variable it is very important that trimmed equilibrium is possible at all conditions. For this reason considerable emphasis is given to ensuring adequate longitudinal static stability and trim control. Because of their importance *static stability* and *trim* are often interpreted to mean *longitudinal static stability* and *trim.*

The commonly used theory of longitudinal static stability was developed by Gates and Lyon (1944), and derives from a full, static and dynamic, stability analysis of the equations of motion of an aircraft. An excellent and accessible summary of the findings of Gates and Lyon is given in Duncan (1959) and also in Babister (1961). In the interests of understanding and physical interpretation the theory is often reduced to a linearised form retaining only the principal aerodynamic and configuration parameters. It is in this simplest form that the theory is reviewed here since it is only required as the basis on which to build the small perturbation dynamics model. It is important to appreciate that although the longitudinal static stability model is described only in terms of the aerodynamic properties of the

airframe, the control and trim properties as seen by the pilot must conform to the same physical interpretation even when they are augmented by a flight control system. It is also important to note that static and dynamic stability are, in reality, inseparable. However, the separate treatment of static stability is a useful means for introducing the concept of stability insofar as it determines the control and trim characteristics of the aircraft.

4.3.1.2 Conditions for Stability

The static stability of an aircraft is commonly interpreted to describe its tendency to converge on the initial equilibrium condition following a small disturbance from trim. Dynamic stability, on the other hand, describes the transient motion involved in the process of recovering equilibrium following the disturbance. Fig. 4.3-1 includes two illustrations showing the effects of static stability and static instability in an otherwise dynamically stable aircraft. Following an initial disturbance displacement, for example in pitch, at time $t = 0$ the subsequent response time history is shown and is clearly dependent on the stability of the aircraft. It should be noted that the damping of the dynamic oscillatory component of the responses shown was deliberately chosen to be low in order to best illustrate the static and dynamic stability characteristics.

In establishing trim equilibrium the pilot adjusts the elevator angle and thrust to obtain a lift force sufficient to support the weight and thrust sufficient to balance the drag at the desired speed and flight path angle. Since the airframe is symmetric the equilibrium side force is of course zero. Provided that the speed is above the minimum drag speed then the force balance will remain stable with speed. Therefore, the static stability of the aircraft reduces to a consideration of the effects of angular disturbances about the three axes. Following such a disturbance the aerodynamic forces and moments will no longer be in equilibrium, and in a statically stable aircraft the resultant moments will cause the aircraft to converge on its initial condition. The condition for an aircraft to be statically stable is therefore easily deduced.

Consider a positive pitch, or incidence, disturbance from equilibrium. This is in the nose up sense and results in an increase in incidence α and hence in lift coefficient C_L. In a stable aircraft the resulting pitching moment must be restoring, that is, in the negative or nose down sense. And of course the converse must be true following a nose down disturbance. Thus the condition for longitudinal static stability may be determined by plotting pitching moment M, or pitching moment coefficient C_m, for variation in incidence α about the trim value α_e as shown in Fig. 4.3-2. The nose up disturbance increases α and takes the aircraft to the out-of-trim point p where the pitching moment coefficient becomes negative and is therefore restoring. Clearly, a nose down disturbance leads to the same conclusion. As indicated, the aircraft is stable when the slope of this plot is negative. Thus, the condition for stable trim at incidence α_e may be expressed:

$$C_m = 0 \tag{4.3.1}$$

and

$$\frac{dC_m}{d\alpha} < 0 \tag{4.3.2}$$

The above observation is only strictly valid when it is assumed that the aerodynamic force and moment coefficients are functions of incidence only. This is usually an acceptable approximation for subsonic aircraft and, indeed, the plot

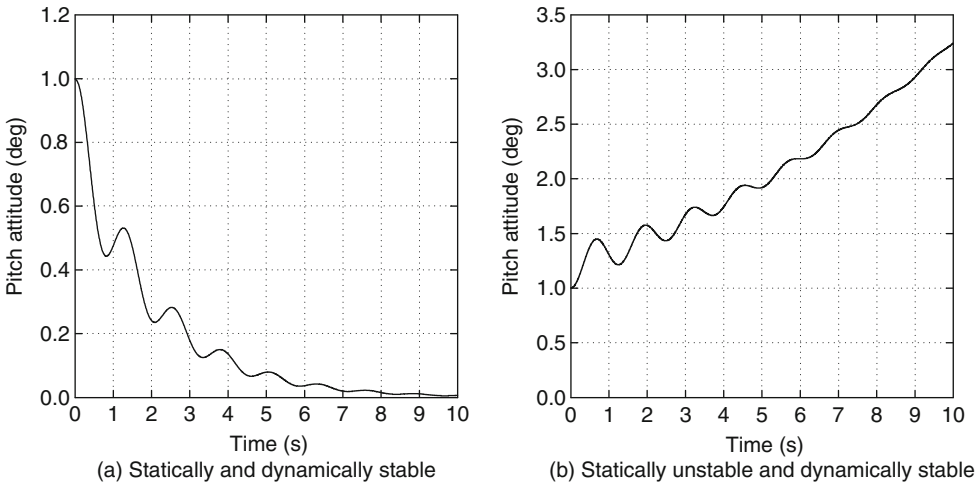


Figure 4.3-1 Stability.

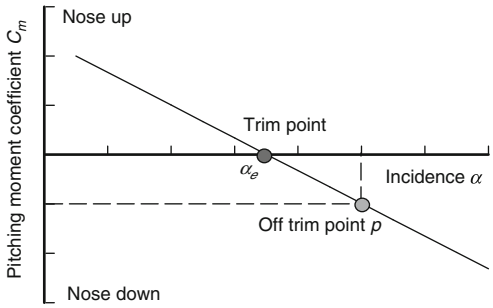


Figure 4.3-2 Pitching moment variation with incidence for a stable aircraft.

of pitching moment coefficient against incidence may well be very nearly linear as shown in Fig. 4.3-2. However, this argument becomes increasingly inappropriate with increasing Mach number. As compressibility effects become significant so the aerodynamic force and moment coefficients become functions of both incidence and Mach number. When this occurs equation (4.3.2) may not always guarantee that stable trim can be obtained. The rather more complex analysis by Gates and Lyon (1944) takes speed effects into account and defines a general requirement for longitudinal static stability as

$$\frac{dC_m}{dC_L} < 0 \tag{4.3.3}$$

For subsonic aircraft equations (4.3.2) and (4.3.3) are completely interchangeable since α and C_L are linearly, or very nearly linearly, related by the lift curve slope a.

In a similar way the conditions for lateral–directional static stability may be deduced as

$$\frac{dC_l}{d\phi} < 0 \tag{4.3.4}$$

and

$$\frac{dC_n}{d\beta} > 0 \tag{4.3.5}$$

where C_l and C_n are rolling moment and yawing moment coefficients respectively and ϕ and β are roll angle and sideslip angle respectively.

4.3.1.3 Degree of Stability

It was shown above that the condition for an aircraft to possess static stability about all three axes at a given trim condition is that the gradients of the $C_m - \alpha$ and $C_l - \phi$ plots must be negative, whilst the gradient of the $C_n - \beta$ plot must be positive. Now, obviously, a very large range of values of the gradients is possible and the magnitude of the gradient determines the *degree of stability* possessed by the aircraft. Variation in the degree of longitudinal static stability is illustrated in Fig. 4.3-3. The degree of stability is described in terms of *stability margin* which quantifies how much stability the aircraft has over and above zero or neutral stability. Thus, for example, the longitudinal static stability margin is directly related to the gradient of the C_m–α plot.

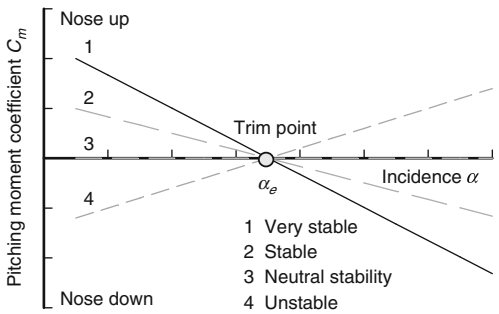


Figure 4.3-3 The degree of longitudinal static stability.

With reference to Fig. 4.3-3 and for a given disturbance in α it is clear that the corresponding restoring pitching moment C_m is greatest for a very stable aircraft. The magnitude of the restoring moment decreases as the degree of stability, or stability margin, is reduced and becomes zero at neutral stability. Clearly, when the aircraft is unstable the moment is of the opposite sign and is therefore divergent. Thus the higher the degree of stability the greater is the restoring moment following a disturbance. This means that a very stable aircraft will be very resistant to upset. This in turn means that greater control actions will be needed to encourage the aircraft to change its trim state or to manoeuvre. It follows then, that the stability margins determine the magnitude of the control actions required to trim the aircraft. It is easy to appreciate that a consequence of this is that too much stability can be as hazardous as too little stability since the available control power is limited.

As mentioned before, the lateral–directional static stability of the aircraft is usually fixed by design and usually remains more or less constant throughout the flight envelope. The lateral–directional stability margins therefore remain substantially constant for all flight conditions. This situation may well break down when large amplitude manoeuvring is considered. Under such circumstances normally linear aerodynamic behaviour may well become very non-linear and cause dramatic changes to observed lateral–directional stability and control characteristics. Although of considerable interest to the flight dynamicist, non-linear behaviour is beyond the scope of this book and constant lateral–directional static stability is assumed throughout.

4.3.1.4 Variation in Stability

Changes in the aerodynamic operating conditions of an aircraft which result in pitching moment changes inevitably lead to variation in longitudinal static stability. Such variation in stability is normally manifest as a non-linear version of the $C_m - C_L$ characteristic shown in Fig. 4.3-2. For the subsonic classical aircraft such changes are usually small and may result in some non-linearity of the pitching moment characteristic with change in trim. In general the variation in the degree of stability is acceptably small. For the modern supersonic high performance aircraft the situation is not so well defined. Large flight envelopes and significant variation in flight condition can lead to dramatic changes in static stability. For example, it is possible for such an aircraft to be stable at some conditions and unstable at others. It is easy to see how such variations might arise in a physical sense, but it is much more difficult to describe the variations in mathematical terms. A brief review of some of the more obvious sources of variation in stability follows.

4.3.1.4.1 Power effects

Probably the most significant variation in longitudinal static stability arises from the effects of power. Direct effects result from the point of application and line of action of the thrust forces with respect to the *cg*. Clearly, as illustrated in Fig. 4.3-4, a high thrust line results in a nose down pitching moment and *vice versa*. In normal trimmed flight the thrust moment is additive to the aerodynamic moment and the total pitching moment would be trimmed to zero by adjustment of the elevator. However, any aerodynamic perturbation about trim which results in a thrust perturbation is potentially capable of giving rise to a non-linear stability characteristic. The precise nature of the variation in stability is dependent on the operating characteristics of the installed power unit which may not be easy to identify.

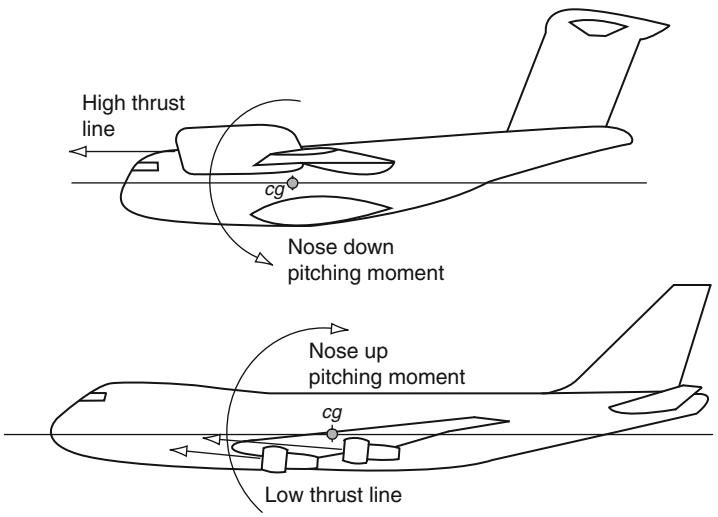


Figure 4.3-4 Typical thrust line effects on pitching moment.

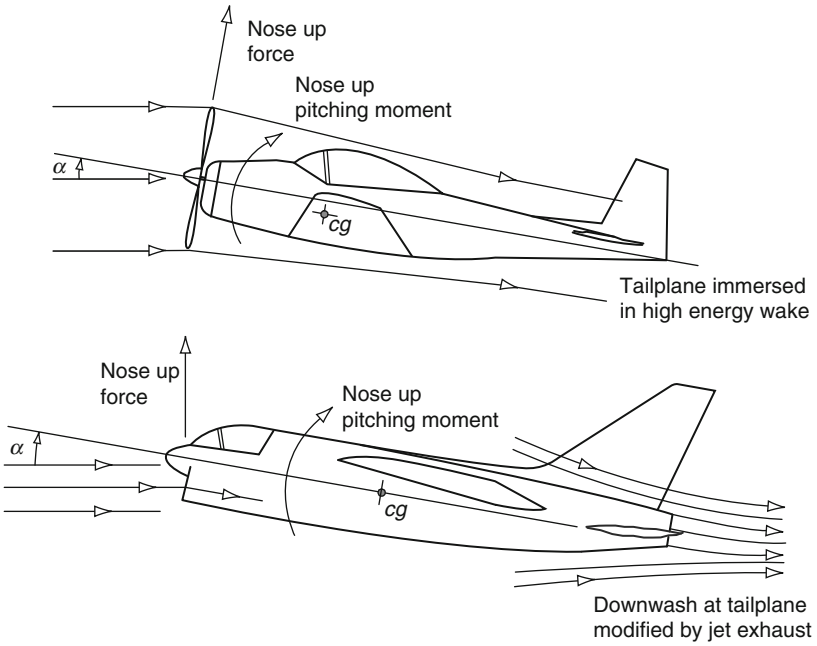


Figure 4.3-5 Typical induced flow effects on pitching moment.

Indirect power effects are caused by the induced flow associated with a propeller and its wake or the intake and exhaust of a gas turbine engine. Some of the more obvious induced flow effects are illustrated in Fig. 4.3-5. The process of turning the incident flow through the body incidence angle into the propeller disc or into the engine intake creates a normal force at the propeller or engine intake as shown. In general this effect gives rise to a nose up pitching moment. The magnitude of the normal force is dependent on the body incidence angle and on the increase in flow energy at the propeller disc or engine intake. The force will therefore vary considerably with trim condition. The force is also sensitive to aerodynamic perturbations about trim; it is therefore easy to appreciate its contribution to pitching moment non-linearity.

The wake behind a propeller is a region of high energy flow which modifies the aerodynamic operating conditions over parts of the wing and tailplane. The greatest effect on pitching moment arises from the tailplane. The effectiveness of the tailplane is enhanced simply because of the increased flow velocity and the reduction in down-wash angle. These two effects together increase the nose down pitching moment available and hence increase the degree of stability of the aircraft. The induced flow effects associated with the propeller-driven aircraft can have a significant influence on its longitudinal static stability. These effects also change with aerodynamic conditions especially at high angles of attack. It is therefore quite common to see some non-linearity in the pitching moment trim plot for such an aircraft at high values of lift coefficient. It should also be noted that the propeller wake rotates about the longitudinal axis. Although less significant, the rotating flow has some influence on the lateral–directional static stability of the aircraft.

The exhaust from a jet engine, being a region of very high velocity and reduced pressure, creates an inflow field as indicated in Fig. 4.3-5. Clearly the influence on pitching moment will depend on the relative location of the aerodynamic surfaces of the aircraft and the engine exhausts. When the tailplane is immersed in this induced flow field then there is a change in the downwash angle. Thus the effect is to increase the static stability when the downwash angle is reduced and *vice versa*. In general this effect is not very significant, except perhaps for the aircraft with engines mounted in pods on the rear fuselage and in which the tailplane is very close to the exhaust wake.

4.3.1.4.2 Other effects

Although power effects generally make the most significant contribution to variation in longitudinal static stability other potentially important contributory sources also exist. For example, wing sweep back and aircraft geometry which result in significant variation in downwash at the tailplane generally tend to reduce the available stability, an effect which is clearly dependent on the aerodynamic trim condition. The fuselage alone is usually unstable and the condition worsens with increasing Mach number. On the other hand, at high subsonic and supersonic Mach numbers the aerodynamic centres of the wing and tailplane move aft. This has the effect of increasing the available nose down pitching moment which is a stabilising characteristic. And finally, since all airframes have some degree of flexibility

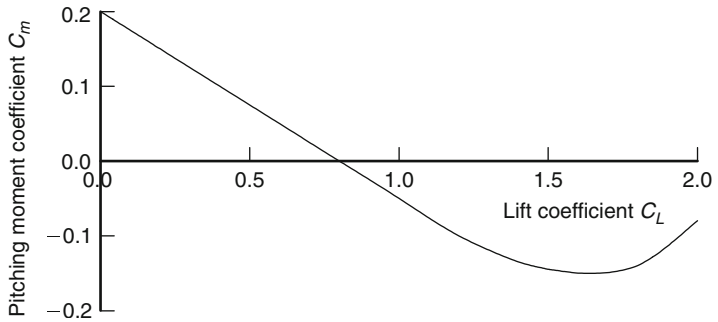


Figure 4.3-6 Stability reversal at high lift coefficient.

the structure distorts under the influence of aerodynamic loads. Today aeroelastic distortion of the structure is carefully controlled by design and is not usually significant in influencing static stability. However, in the very large civil transport aircraft the relative geometric disposition of the wing and tailplane changes with loading conditions; some contribution to the variation in pitching moment is therefore inevitable but the contribution is usually small.

Taking all of these effects together, the prospect of ever being able to quantitatively define the longitudinal static stability of an aircraft may seem daunting. Fortunately these effects are well understood and can be minimised by design. The result for most aircraft is a pitching moment trim characteristic with some non-linear tendency at higher values of trim lift coefficient. In extreme cases the stability of the aircraft can actually reverse at high values of lift coefficient to result in an unstable pitch up characteristic. A typical pitching moment trim plot for an aircraft with a pitch up characteristic is shown in Fig. 4.3-6.

Example 4.3.1

To illustrate the variation in the pitching moment characteristic for a typical subsonic aircraft, the relevant data obtained from wind tunnel experiments on a l/6th scale model of the Handley Page HP-137 are shown plotted in Fig. 4.3-7. The data were extracted from a report by Storey (1966). They were obtained at a tunnel speed of 200 ft/s and the Reynolds number was $\mathrm{Re} = 1.2 \times 10^6$ based on mean aerodynamic chord $\bar{\bar{c}}$. The HP-137 is in fact the well known Jetstream aircraft; however, it is not known if the data shown are representative of the actual aircraft flying today.

The plots show the characteristic for the aircraft without tail and for the aircraft with tail at various combinations of setting angle η_T and elevator angle η. Clearly, all of the plots are reasonably linear at all values of lift coefficient up to

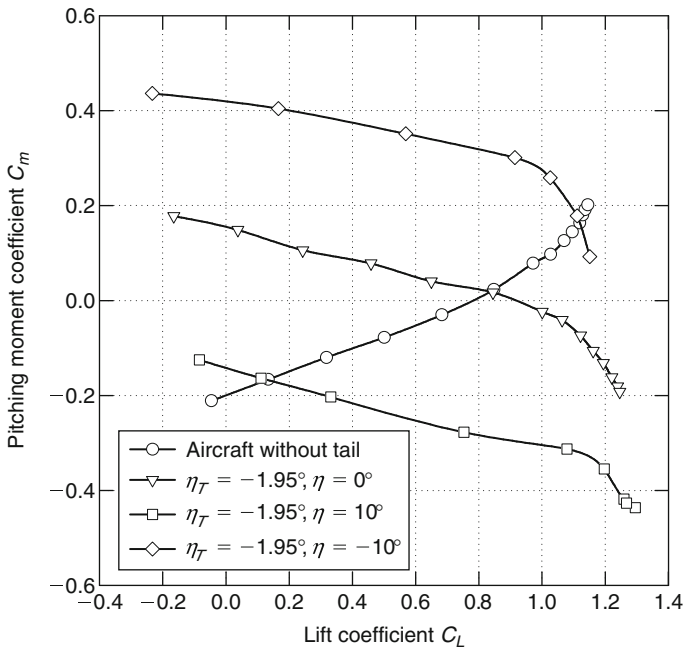


Figure 4.3-7 C_m–α plots for a 1/6th scale model of the Handley Page Jetstream.

the stall. Without a tailplane the aircraft is unstable since the slope of the plot is positive. With tailplane the slope, and hence the degree of stability, is more or less constant. Assuming that the trim ($C_m = 0$) range of lift coefficient is approximately $-0.2 \leq C_L \leq 1.0$ then, by interpolation, it can be seen that this can be obtained with an elevator angle range of approximately $-0.6° \leq \eta \leq 0°$. Clearly this is well within the control capability of the tailplane and elevator configuration shown in this example.

This kind of experimental analysis would be used to confirm the geometric design of the tailplane and elevator. In particular, it is essential to establish that the aircraft has an adequate stability margin across the trim envelope, that the elevator angle required to trim the aircraft is within its aerodynamic capability and that a sufficient margin of elevator control range remains for manoeuvring.

4.3.2 The Pitching Moment Equation

Having established the importance of pitching moment in the determination of longitudinal static stability, further analysis of stability requires the development of the pitching moment equation. A fully representative general pitching moment equation is difficult to develop since it is very dependent on the geometry of the aircraft. However, it is possible to develop a simple approximation to the pitching moment equation, which is sufficiently representative for most preliminary studies and which provides considerable insight into the basic requirements for static stability and trim.

4.3.2.1 Simple Development of the Pitching Moment Equation

For the development of the simplest possible pitching moment equation it is usual to define a model showing only the normal forces and pitching moments acting on the aircraft. It is assumed that at steady level flight the thrust and drag are in equilibrium and act at the *cg* and further, for small disturbances in incidence, changes in this equilibrium are insignificant. This assumption therefore implies that small disturbances in incidence cause significant changes in lift forces and pitching moments only. The model defined in these terms is shown in Fig. 4.3-8.

For the purposes of modelling pitching behaviour the model comprises two parts, the wing and fuselage combination and the tailplane. It is then assumed that the wing and fuselage behave aerodynamically like a wing alone. Clearly, this is not true since the fuselage may make significant aerodynamic contributions and, in any event, its presence will interfere with the aerodynamic properties of the wing to a greater or lesser extent. However, for conventional subsonic aircraft with a reasonably high aspect ratio wing this is a very satisfactory approximation. The tailplane is treated as a separate component since it provides the principal aerodynamic mechanism for controlling longitudinal static stability and trim. The following analysis establishes the fundamental importance of the tailplane parameters in the provision of longitudinal static stability.

Referring to Fig. 4.3-8 it is seen that the wing–fuselage lift L_w and residual pitching moment M_0 act at the aerodynamic centre *ac* of the combination which is assumed to be coincident with the aerodynamic centre of the wing alone. In a similar way the lift L_T and pitching moment M_T of the tailplane are assumed to act at its aerodynamic centre. The longitudinal geometry of the model is entirely related to the mean aerodynamic chord *mac* ($\bar{\bar{c}}$) as shown in Fig. 4.3-8. An expression for the total pitching moment M about the *cg* may therefore be written:

$$M = M_0 + L_w(h - h_0)\bar{\bar{c}} - L_T l_T + M_T \tag{4.3.6}$$

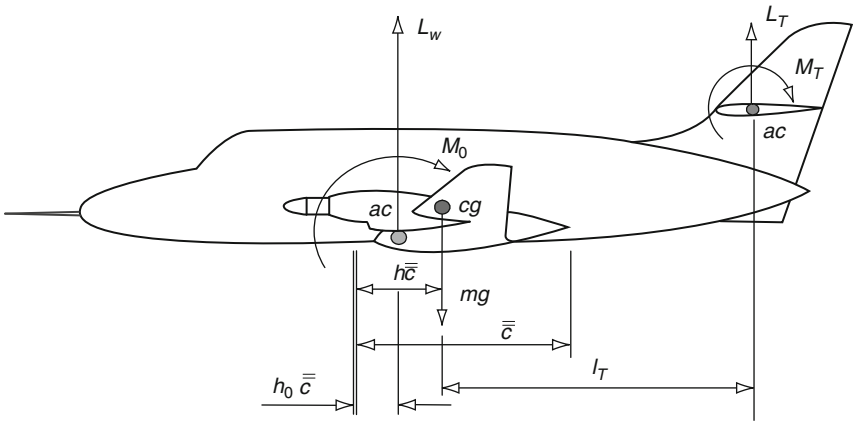


Figure 4.3-8 Simple pitching moment model.

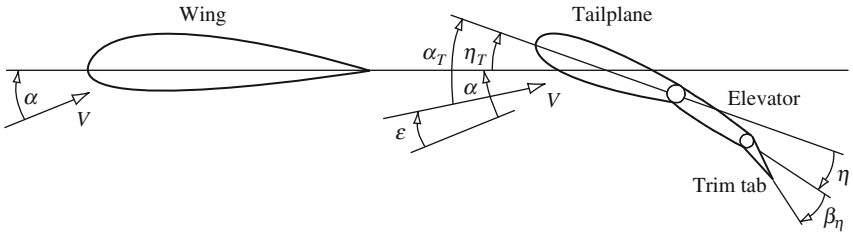


Figure 4.3-9 Wing–tailplane flow geometry.

If, as is usual, it is assumed that the tailplane aerofoil section is symmetric then M_T becomes zero. Thus, in the more convenient coefficient form equation (4.3.6) may be written:

$$C_m = C_{m_0} + C_{L_w}(h - h_0) - C_{L_T}\overline{V}_T \tag{4.3.7}$$

To facilitate further analysis of pitching moment it is necessary to express the tailplane lift coefficient C_{L_T} in terms of more accessible tailplane parameters. Tailplane lift coefficient may be expressed:

$$C_{L_T} = a_0 + a_1\alpha_T + a_2\eta + a_3\beta_\eta \tag{4.3.8}$$

where a_0, a_1, a_2 and a_3 are constant aerodynamic coefficients, α_T is the local incidence, η is the elevator angle and β_η is the elevator trim tab angle. Note that since a symmetric tailplane aerofoil section is assumed a_o is also zero. The local tailplane incidence is influenced by the *tailplane setting angle* η_T and the local flow distortion due to the effect of the downwash field behind the wing. The flow geometry is shown in Fig. 4.3-9.

Clearly the angle of attack of the tailplane is given by

$$\alpha_T = \alpha - \varepsilon + \eta_T \tag{4.3.9}$$

where ε is the downwash angle at the tailplane. Since, to a good approximation, for small disturbances the downwash angle is a function of wing–body incidence α only:

$$\alpha - \varepsilon = \alpha\left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) = \frac{C_{L_w}}{a}\left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) \tag{4.3.10}$$

whence

$$\alpha_T = \frac{C_{L_w}}{a}\left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) + \eta_T \tag{4.3.11}$$

Now substituting the expression for α_T given by equation (4.3.11) into equation (4.3.8), substituting the resulting expression for C_{L_T} into equation (4.3.7) and noting that a_0 is zero then, the pitching moment equation in its simplest and most general form is obtained:

$$C_m = C_{m_0} + C_{L_w}(h - h_0) - \overline{V}_T\left(C_{L_w}\frac{a_1}{a}\left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) + a_2\eta + a_3\beta_\eta + a_1\eta_T\right) \tag{4.3.12}$$

A simple computational algorithm for estimating the rate of change of downwash with angle of attack $\mathrm{d}\varepsilon/\mathrm{d}\alpha$ is given in Stribling (1984) and its use is illustrated in one of the *Mathcad* trim programs.

4.3.2.2 Elevator Angle to Trim

It has already been shown, in equation (4.3.1), that the condition for trim is that the total pitching moment can be adjusted to zero, that is, $C_m = 0$. Applying this condition to equation (4.3.12) the elevator angle required to trim the aircraft is given by

$$\eta = \frac{1}{\overline{V}_T a_2}(C_{m_0} + C_{L_w}(h - h_0)) - \frac{C_{L_w}}{a}\left(\frac{a_1}{a_2}\right)\left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) - \frac{a_3}{a_2}\beta_\eta - \frac{a_1}{a_2}\eta_T \tag{4.3.13}$$

When the elevator tab is set at its neutral position, $\beta_\eta = 0$ and for a given *cg* position h the elevator angle to trim varies only with lift coefficient. For any other tab setting a different elevator angle is required to trim. Therefore, to an extent, elevator and elevator tab provide interchangeable means for achieving longitudinal trim.

4.3.2.3 Test for Longitudinal Static Stability

The basic requirement for an aircraft to be statically stable at a given trim condition is stated in equation (4.3.2). By differentiating equation (4.3.12) with respect to C_L, or equivalently C_{L_w}, and noting that η_T and, by definition, C_{m_0} are constants then the condition for the aircraft to be stable is given by

$$\frac{dC_m}{dC_{L_w}} < 0$$

where

$$\frac{dC_m}{dC_{L_w}} = (h - h_0) - \overline{V}_T\left(\frac{a_1}{a}\left(1 - \frac{d\varepsilon}{d\alpha}\right) + a_2\frac{d\eta}{dC_{L_w}} + a_3\frac{d\beta_\eta}{dC_{L_w}}\right) \tag{4.3.14}$$

Thus at a given *cg* position h, the longitudinal static stability of the aircraft and the aerodynamic control characteristics, that is, *elevator angle to trim*, $d\eta/dC_{L_w}$, and *elevator tab angle to trim*, $d\beta_\eta/dC_{L_w}$, are interdependent. Further analysis is usually carried out by separating the effects of elevator angle and tab angle in equation (4.3.14). *Controls fixed stability* is concerned with the interdependence of elevator angle to trim and stability, whereas *controls free stability* is concerned with the interdependence of elevator tab angle to trim and stability.

4.3.3 Longitudinal Static Stability

4.3.3.1 Controls Fixed Stability

The condition described as *controls fixed* is taken to mean the condition when the elevator and elevator tab are held at constant settings corresponding to the prevailing trim condition. In practice this means that the pilot is flying the aircraft with his hands on the controls and is holding the controls at the *fixed* setting required to trim. This, of course, assumes that the aircraft is stable and remains in trim.

Since the controls are fixed:

$$\frac{d\eta}{dC_{L_w}} = \frac{d\beta\eta}{dC_{L_w}} = 0 \tag{4.3.15}$$

and equation (4.3.14) may be written:

$$\frac{dC_m}{dC_{L_w}} = (h - h_0) - \overline{V}_T\frac{a_1}{a}\left(1 - \frac{d\varepsilon}{d\alpha}\right) \tag{4.3.16}$$

Or, writing,

$$K_n = -\frac{dC_m}{dC_{L_w}} = h_n - h \tag{4.3.17}$$

where K_n is the *controls fixed stability margin*, the slope of the C_m–C_L plot. The location of the *controls fixed neutral point* h_n on the mean aerodynamic chord $\bar{c}$ is therefore given by

$$h_n = h_0 + \overline{V}_T\frac{a_1}{a}\left(1 - \frac{d\varepsilon}{d\alpha}\right) \tag{4.3.18}$$

For a statically stable aircraft the stability margin K_n is positive, and the greater its value the greater the degree of stability possessed by the aircraft. With reference to equation (4.3.17) it is clear that the aircraft will be stable when the *cg* position h is ahead of the controls fixed neutral point h_n. The acceptable margins of stability therefore determine the permitted range of *cg* position in a given aircraft. The aft limit often corresponds with the controls fixed neutral point, whereas the forward limit is determined by the maximum permissible stability margin. Remember, Section 4.3.1.3, that too much stability can be as hazardous as too little stability.

The meaning of controls fixed stability is easily interpreted by considering the pilot actions required to trim an aircraft in a controls fixed sense. It is assumed at the outset that the aircraft is in fact stable and hence can be trimmed to an equilibrium flight condition. When the aircraft is in a trimmed initial equilibrium state the pitching moment is zero and equation (4.3.12) may be written:

$$0 = C_{m_0} + C_{L_w}(h - h_0) - \overline{V}_T\left(C_{L_w}\frac{a_1}{a}\left(1 - \frac{d\varepsilon}{d\alpha}\right) + a_2\eta + a_3\beta_\eta + a_1\eta_T\right) \tag{4.3.19}$$

It is assumed that the pilot is holding the controls at the required elevator angle, the power is set to give steady level flight and the elevator tab is set at its datum, $\beta_\eta = 0$. Now, to retrim the aircraft at a new flight condition in a controls fixed sense it is necessary for the pilot to move the controls to the new elevator setting and then to hold the controls at that setting. For example, to retrim at a higher speed in a more nose down attitude, the pilot would move the control column forward until his new condition was established and would then simply hold the column at that position. This would of course leave the aircraft in a descending condition unless the power were increased sufficient to maintain level flight at the higher speed. However, power variations are not allowed for in the simple model reviewed here.

Thus to trim a stable aircraft at any condition in its speed envelope simply requires the selection of the correct elevator angle, all other parameters remaining constant. Therefore, the variable in controls fixed stability analysis is elevator angle to trim. Differentiating equation (4.3.19) with respect to C_{L_w} and making the same assumptions as before but allowing elevator angle η to vary with trim, then after some rearrangement it may be shown that

$$\frac{\mathrm{d}\eta}{\mathrm{d}C_{L_w}} = \frac{-1}{\overline{V}_T a_2}(h_n - h) = \frac{-1}{\overline{V}_T a_2} K_n \tag{4.3.20}$$

Thus, since $\overline{V}_T$ and a_2 are constants, the *elevator angle to trim* characteristic $\mathrm{d}\eta/\mathrm{d}C_{L_w}$ is proportional to the controls fixed stability margin K_n. Measurements of elevator angle to trim for a range of flight conditions, subject to the assumptions described, provide a practical means for determining controls fixed stability characteristics from flight experiments. However, in such experiments it is not generally possible to completely eliminate the effects of power on the results.

Example 4.3.2

The practical evaluation of controls fixed static stability centres on the application of equations (4.3.13), (4.3.19) and (4.3.20) to a stable aircraft. It is relatively straightforward to obtain measurements of the elevator angle η required to trim an aircraft at a chosen value of lift coefficient C_L. Provided that the power and elevator trim tab angle β_η are maintained at a constant setting throughout the measurement process then the above mentioned equations apply directly. A flight test exercise conducted in a Handley Page Jetstream by the author, under these conditions, provided the trim data plotted in Fig. 4.3-10 for three different *cg* positions. At any given value of lift coefficient C_L the corresponding value of elevator angle to trim η is given by the solution of equation (4.3.13), or alternatively equation (4.3.19). The plots are clearly non-linear and the non-linearity in this aircraft is almost entirely due to the effects of power.

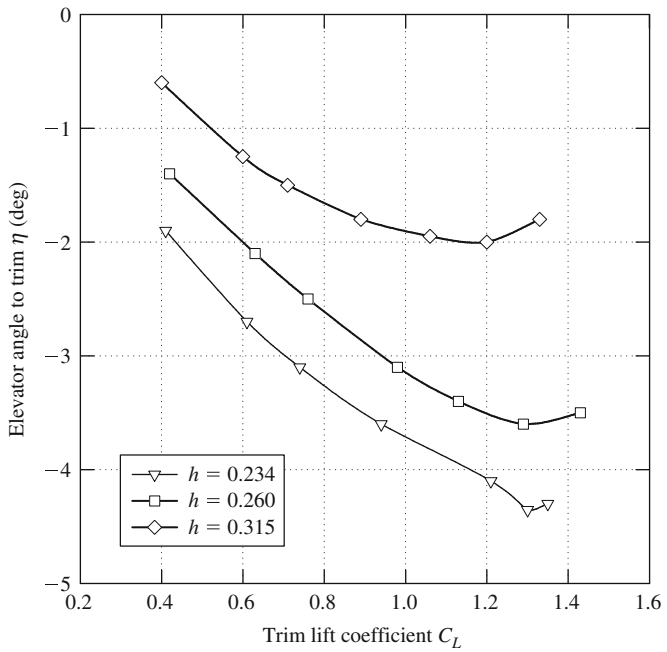


Figure 4.3-10 Plot of elevator angle to trim.

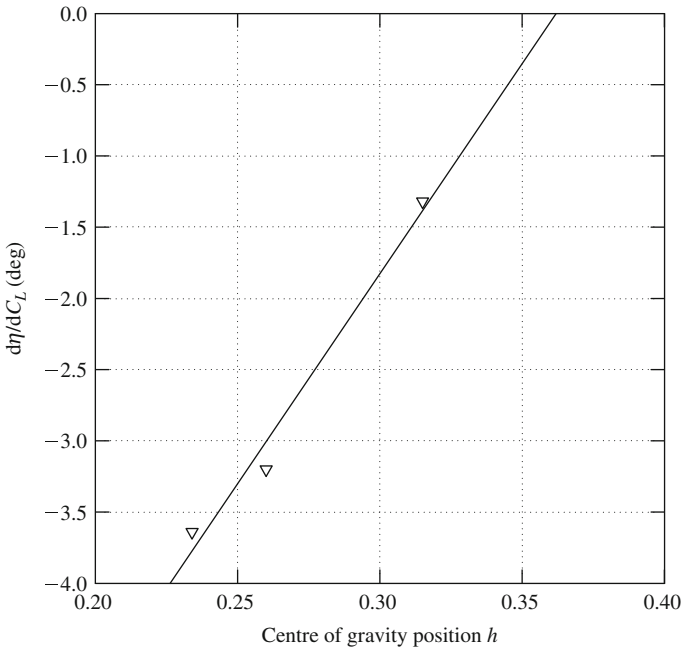


Figure 4.3-11 Determination of controls fixed neutral point.

Since the gradients of the plots shown in Fig. 4.3-10 are all negative the aircraft is statically stable in accordance with equation (4.3.20). However, for any given *cg* position the gradient varies with lift coefficient indicating a small variation in stability margin. In a detailed analysis the stability margin would be evaluated at each value of trimmed lift coefficient to quantify the variation in stability. In the present example the quality of the data was not good enough to allow such a complete analysis. To establish the location of the controls fixed neutral point h_n equation (4.3.20) must be solved at each value of trim lift coefficient. This is most easily done graphically as shown in Fig. 4.3-11.

Equation (4.3.20) is solved by plotting $\mathrm{d}\eta/\mathrm{d}C_L$ against *cg* position h as shown. In this example, the mean gradient for each *cg* position is plotted rather than the value at each trim point. Since equation (4.3.20) represents a linear plot a straight line may be fitted to the three data points as shown. Extrapolation to the neutral stability point at which $\mathrm{d}\eta/\mathrm{d}C_L = 0$ corresponds with a *cg* position of approximately $h = 0.37$. Clearly, three data points through which to draw a line is barely adequate for this kind of evaluation. A controls fixed neutral point h_n at 37% of *mac* correlates well with the known properties of the aircraft. The most aft *cg* position permitted is in fact at 37% of *mac*. Having established the location of the controls fixed neutral point the controls fixed stability margin K_n for each *cg* position follows from the application of equation (4.3.20).

In a more searching stability evaluation rather more data points would be required and data of much better quality would be essential. Although limited, the present example does illustrate the typical controls fixed longitudinal static stability characteristics of a well behaved classical aircraft.

4.3.3.2 Controls Free Stability

The condition described as *controls free* is taken to mean the condition when the elevator is free to float at an angle corresponding to the prevailing trim condition. In practice this means that the pilot can fly the aircraft with his hands off the controls whilst the aircraft remains in its trimmed flight condition. Again, it is assumed that the aircraft is stable, otherwise it will diverge when the controls are released. Now this situation can only be obtained if the controls can be adjusted such that the elevator will float at the correct angle for the desired *hands-off* trim condition. This is arranged by adjusting the elevator trim tab until the required trim is obtained. Thus controls free stability is concerned with the trim tab and its control characteristics.

When the controls are free, the elevator hinge moment H is zero and the elevator floats at an indeterminate angle η. It is therefore necessary to eliminate elevator angle from the pitching moment equation (4.3.12) in order to facilitate the analysis of controls free stability. Elevator hinge moment coefficient is given by the expression

$$C_H = b_1\alpha_T + b_2\eta + b_3\beta_\eta \tag{4.3.21}$$

where b_1, b_2 and b_3 are constants determined by the design of the elevator and trim tab control circuitry. Substituting for local tailplane incidence α_T as given by equation (4.3.11), then equation (4.3.21) may be rearranged to determine the angle at which the elevator floats. Thus,

$$\eta = \frac{1}{b_2}C_H - \frac{C_{L_w}}{a}\frac{b_1}{b_2}\left(1 - \frac{d\varepsilon}{d\alpha}\right) - \frac{b_3}{b_2}\beta_\eta - \frac{b_1}{b_2}\eta_T \tag{4.3.22}$$

To eliminate elevator angle from the pitching moment equation, substitute equation (4.3.22) into equation (4.3.12) to obtain

$$C_m = C_{m_0} + C_{L_w}(h - h_0) - \overline{V}_T\left(\begin{array}{l} C_{L_w}\dfrac{a_1}{a}\left(1 - \dfrac{d\varepsilon}{d\alpha}\right)\left(1 - \dfrac{a_2 b_1}{a_1 b_2}\right) + a_3\beta_\eta\left(1 - \dfrac{a_2 b_3}{a_3 b_2}\right) \\ +a_1\eta_T\left(1 - \dfrac{a_2 b_1}{a_1 b_2}\right) + \dfrac{a_2}{b_2}C_H \end{array}\right) \tag{4.3.23}$$

Now in the controls free condition $C_H = 0$ and noting that η_T, C_{m_0} and, since the tab is set at the trim value, β_η are constants then, differentiating equation (4.3.23) with respect to C_{L_w}:

$$\frac{dC_m}{dC_{L_w}} = (h - h_0) - \overline{V}_T\frac{a_1}{a}\left(1 - \frac{d\varepsilon}{d\alpha}\right)\left(1 - \frac{a_2 b_1}{a_1 b_2}\right) \tag{4.3.24}$$

Or, writing,

$$K_n' = -\frac{dC_m}{dC_{L_w}} = h_n' - h \tag{4.3.25}$$

where K_n' is the *controls free stability margin*, the slope of the C_m–C_L plot with the controls free. The location of the *controls free neutral point* h_n' on the mean aerodynamic chord $\bar{c}$ is given by

$$\begin{aligned} h_n' &= h_0 + \overline{V}_T\frac{a_1}{a}\left(1 - \frac{d\varepsilon}{d\alpha}\right)\left(1 - \frac{a_2 b_1}{a_1 b_2}\right) \\ &= h_n - \overline{V}_T\frac{a_2 b_1}{a b_2}\left(1 - \frac{d\varepsilon}{d\alpha}\right) \end{aligned} \tag{4.3.26}$$

Thus, as before, for a statically stable aircraft the controls free stability margin K_n' is positive and the greater its value the greater the degree of stability possessed by the aircraft. With reference to equation (4.3.25) it is clear that for controls free stability the *cg* position h must be ahead of the controls free neutral point h_n'. Equation (4.3.26) shows the relationship between the controls fixed and the controls free neutral points. The numerical values of the elevator and tab constants are such that usually $h_n' > h_n$, which means that it is common for the controls free neutral point to lie aft of the controls fixed neutral point. Thus an aircraft that is stable controls fixed will also usually be stable controls free and it follows that the controls free stability margin K_n' will be greater than the controls fixed stability margin K_n.

The meaning of controls free stability is readily interpreted by considering the pilot actions required to trim the aircraft in a controls free sense. It is assumed that the aircraft is stable and is initially in a hands-off trim condition. In this condition the pitching moment is zero and hence equation (4.3.23) may be written:

$$0 = C_{m_0} + C_{L_w}(h - h_0) - \overline{V}_T\left(\begin{array}{l} C_{L_w}\dfrac{a_1}{a}\left(1 - \dfrac{d\varepsilon}{d\alpha}\right)\left(1 - \dfrac{a_2 b_1}{a_1 b_2}\right) \\ +a_3\beta_\eta\left(1 - \dfrac{a_2 b_3}{a_3 b_2}\right) + a_1\eta_T\left(1 - \dfrac{a_2 b_1}{a_1 b_2}\right) \end{array}\right) \tag{4.3.27}$$

Now, to retrim the aircraft, it is necessary for the pilot to grasp the control column and move it to the position corresponding with the elevator angle required for the new trim condition. However, if he now releases the control it will simply move back to its original trim position since an out-of-trim elevator hinge moment, and hence stick force, will exist at the new position. To rectify the problem he must use the trim tab. Having moved the control to the position corresponding with the new trim condition he will be holding a force on the control. By adjusting the trim tab he can null the force and following which, he can release the control and it will stay in the new hands-off position as required. Thus trim tab adjustment is equivalent to control force adjustment, which in turn is directly related to elevator hinge moment adjustment in a mechanical flying control system. To reiterate the previous illustration, consider the situation when the pilot wishes to retrim the aircraft at a higher speed in a more nose down attitude. As before, he will *push* the control column forward until he obtains the desired condition which leaves him holding an out-of-trim force and descending. Elevator tab adjustment will enable him to reduce the control force to zero whereupon he can release the control to enjoy his new hands-off trim condition. Since he will be descending it would

normally be necessary to increase power in order to regain level flight. However, as already stated thrust variations are not allowed for in this model; if they were, the analysis would be considerably more complex.

Thus to trim a stable aircraft at any hands-off flight condition in its speed envelope simply requires the correct selection of elevator tab angle. The variable in controls free stability analysis is therefore elevator tab angle to trim. Differentiating equation (4.3.27) with respect to C_{L_w} and making the same assumptions as previously but allowing elevator tab angle β_η to vary with trim, then after some rearrangement it may be shown that

$$\frac{\mathrm{d}\beta_\eta}{\mathrm{d}C_{L_w}} = \frac{-(h'_n - h)}{a_3 \overline{V}_T \left(1 - \dfrac{a_2 b_3}{a_3 b_2}\right)} = \frac{-K'_n}{a_3 \overline{V}_T \left(1 - \dfrac{a_2 b_3}{a_3 b_2}\right)} \tag{4.3.28}$$

Since it is usual for

$$-a_3 \overline{V}_T \left(1 - \frac{a_2 b_3}{a_3 b_2}\right) > 0 \tag{4.3.29}$$

then the *elevator tab angle to trim* characteristic $\mathrm{d}\beta_\eta/\mathrm{d}C_{L_w}$ is positive and is proportional to the controls free stability margin K'_n. Measurement of the tab angle to trim a range of flight conditions, subject to the assumptions described, provides a practical means for determining controls free stability characteristics from flight experiments. However, since tab angle, elevator hinge moment and control force are all equivalent, it is often more meaningful to investigate control force to trim directly since this is the parameter of direct concern to the pilot.

To determine the equivalence between elevator tab angle to trim and control force to trim, consider the aircraft in a stable hands-off trim state with the tab set at its correct trim value. If the pilot moves the controls in this condition the elevator hinge moment, and hence control force, will vary. Equation (4.3.23) is applicable and may be written:

$$0 = C_{m_0} + C_{L_w}(h - h_0) - \overline{V}_T \left(\begin{array}{l} C_{L_w} \dfrac{a_1}{a} \left(1 - \dfrac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) \left(1 - \dfrac{a_2 b_1}{a_1 b_2}\right) + a_3 \beta_\eta \left(1 - \dfrac{a_2 b_3}{a_3 b_2}\right) \\ + a_1 \eta_T \left(1 - \dfrac{a_2 b_1}{a_1 b_2}\right) + \dfrac{a_2}{b_2} C_H \end{array} \right) \tag{4.3.30}$$

where β_η is set at to its datum trim position and is assumed constant and hinge moment coefficient C_H is allowed to vary with trim condition. Differentiate equation (4.3.30) with respect to C_{L_w} subject to these constraints and rearrange to obtain

$$\frac{\mathrm{d}C_H}{\mathrm{d}C_{L_w}} = \frac{-1}{\overline{V}_T \frac{a_2}{b_2}}(h'_n - h) = \frac{-1}{\overline{V}_T \frac{a_2}{b_2}} K'_n \tag{4.3.31}$$

Comparison of equation (4.3.31) with equation (4.3.28) demonstrates the equivalence of tab angle to trim and hinge moment to trim. Further, if the elevator control force is denoted F_η and g_η denotes the mechanical gearing between the control column and elevator then,

$$F_\eta = g_\eta H = \frac{1}{2}\rho V^2 S_\eta \bar{\bar{c}}_\eta g_\eta C_H \tag{4.3.32}$$

where S_η is the elevator area aft of the hinge line and $\bar{\bar{c}}_\eta$ is the mean aerodynamic chord of the elevator aft of the hinge line. This therefore demonstrates the relationship between control force and hinge moment although equation (4.3.32) shows the relationship also depends on the square of the speed.

Example 4.3.3

The practical evaluation of controls free static stability is undertaken in much the same way as the evaluation of controls fixed stability discussed in Example 4.3.2. In this case the evaluation of controls free static stability centres on the application of equations (4.3.30)–(4.3.32) to a stable aircraft. It is relatively straightforward to obtain measurements of the elevator stick force F_η, and hence hinge moment coefficient C_H, required to trim an aircraft at a chosen value of lift coefficient C_L. Provided that the power and elevator trim tab angle β_η are maintained at a constant setting throughout the measurement process then the above mentioned equations apply directly. As before, a flight test exercise conducted in a Handley Page Jetstream under these conditions provided the trim data plotted in Fig. 4.3-12 for three different *cg* positions. At any given value of lift coefficient C_L the corresponding value of elevator hinge moment to trim C_H is given by the solution of equation (4.3.30). Again, the plots are non-linear due primarily to the effects of power. However, since force measurements are involved the influence of friction in the mechanical control runs is significant and inconsistent. The result of this is data with rather too many spurious points. In order to provide a meaningful example the obviously spurious data points have been "adjusted" to correlate with the known characteristics of the aircraft.

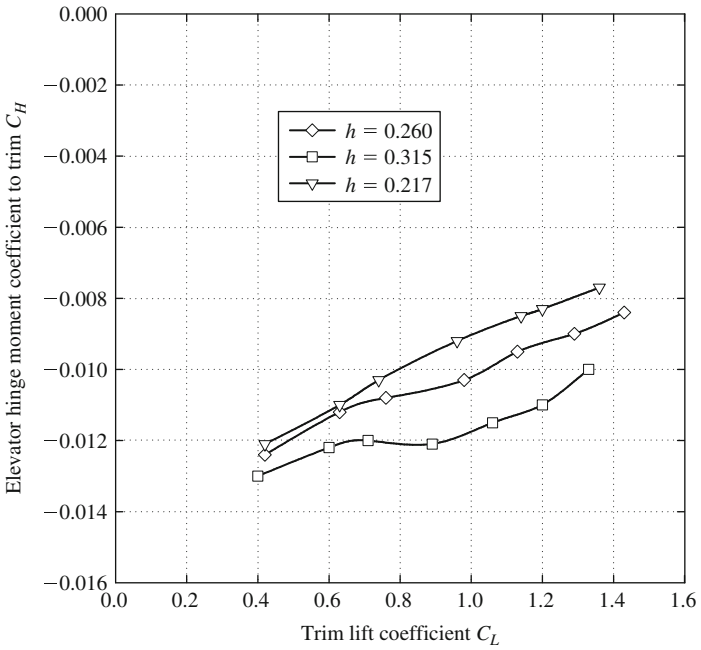


Figure 4.3-12 Plot of hinge moment coefficient to trim.

Since the gradients of the plots shown in Fig. 4.3-12 are all positive the aircraft is statically stable in accordance with equation (4.3.31). However, for any given *cg* position the gradient varies with lift coefficient indicating rather inconsistent variations in stability margin. However, in this case, the variations are more likely to be the result of poor quality data rather than orderly changes in the aerodynamic properties of the aircraft. Again, in a detailed analysis the stability margin would be evaluated at each value of trimmed lift coefficient in order to quantify the variation in stability.

In the present example the quality of the data was clearly not good enough to allow such a complete analysis. To establish the location of the controls free neutral point h'_n equation (4.3.31) must be solved at each value of trim lift coefficient. This is most easily done graphically as shown in Fig. 4.3-13.

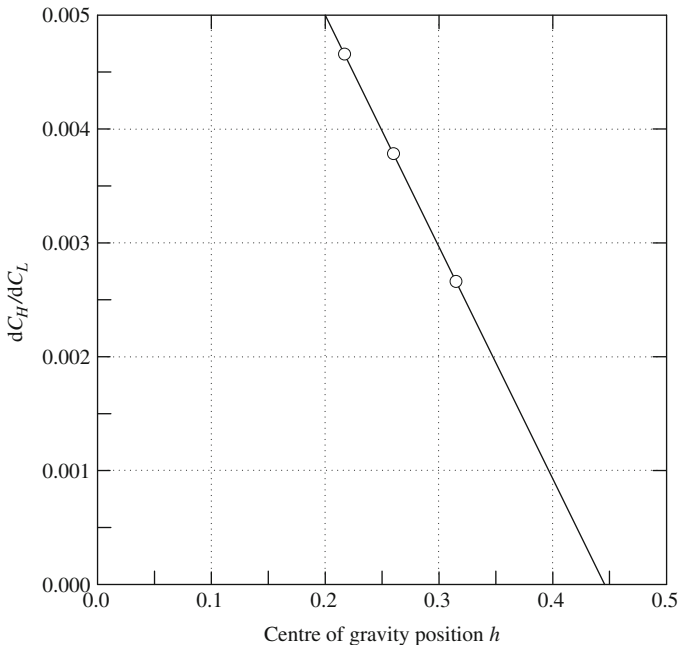


Figure 4.3-13 Determination of controls free neutral point.

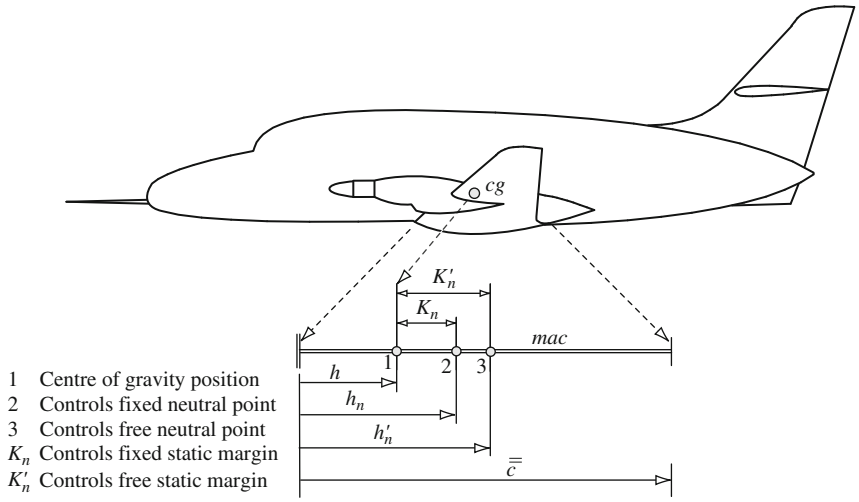


Figure 4.3-14 Longitudinal stability margins.

Equation (4.3.31) is solved by plotting dC_H/dC_L against *cg* position h as shown. In this example, the mean gradient for each *cg* position is plotted rather than the value at each trim point. Since equation (4.3.31) represents a linear plot a straight line may be fitted to the three data points as shown. Extrapolation to the neutral stability point at which $dC_H/dC_L = 0$ corresponds with a *cg* position of approximately $h = 0.44$. A controls free neutral point h'_n at 44% of *mac* correlates reasonably well with the known properties of the aircraft. Having established the location of the controls free neutral point the controls free stability margin K'_n for each *cg* position follows from the application of equation (4.3.25).

4.3.3.3 Summary of Longitudinal Static Stability

A physical interpretation of the meaning of longitudinal static stability may be brought together in the summary shown in Fig. 4.3-14.

The important parameters are neutral point positions and their relationship to the *cg* position which, in turn, determines the stability margins of the aircraft. The stability margins determine literally how much stability the aircraft has in hand, in the controls fixed and free senses, over and above neutral stability. The margins therefore indicate how safe the aircraft is. However, equally importantly, the stability margins provide a measure of the control actions required to trim the aircraft. In particular, the controls fixed stability margin is a measure of the control displacement required to trim and the controls free stability margin is a measure of the control force required to trim. From a flying and handling qualities point of view it is the interpretation of stability in terms of control characteristics which is by far the most important consideration. In practice, the assessment of longitudinal static stability is frequently concerned only with the measurement of control characteristics as illustrated by Examples 4.3.2 and 4.3.3.

4.3.4 Lateral Static Stability

Lateral static stability is concerned with the ability of the aircraft to maintain wings level equilibrium in the roll sense. Wing dihedral is the most visible parameter which confers lateral static stability on an aircraft although there are many other contributions, some of which are destabilising. Since all aircraft are required to fly with their wings level in the steady trim state, lateral static stability is designed in from the outset. Dihedral is the easiest parameter to adjust in the design process in order to "tune" the degree of stability to an acceptable level. Remember that too much lateral static stability will result in an aircraft that is reluctant to manoeuvre laterally, so it is important to obtain the correct degree of stability.

The effect of dihedral as a means for providing lateral static stability is easily appreciated by considering the situation depicted in Fig. 4.3-15. Following a small lateral disturbance in roll ϕ the aircraft will commence to slide "downhill" sideways with a sideslip velocity v. Consider the resulting change in the aerodynamic conditions on the leading wing which has dihedral angle Γ. Since the wing has dihedral, the sideslip velocity has a small component v' resolved perpendicular to the plane of the wing panel where

$$v' = v \sin \Gamma \tag{4.3.33}$$

The velocity component v' combines with the axial velocity component U_e to increase the angle of attack of the leading wing by α'. Since $v' << U_e$ the change in angle of attack α' is small and the total disturbed axial velocity

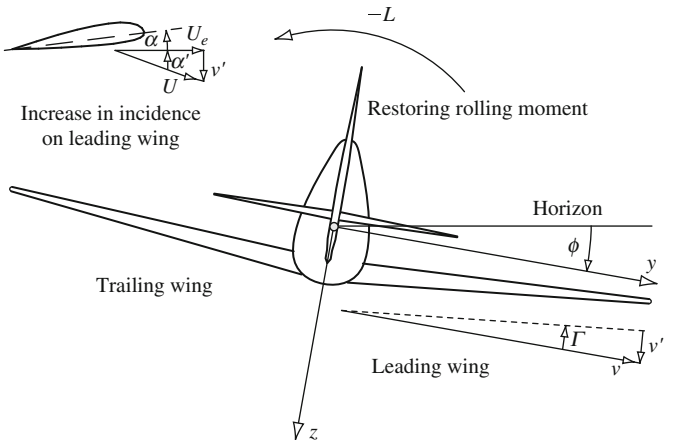


Figure 4.3-15 Dihedral effect.

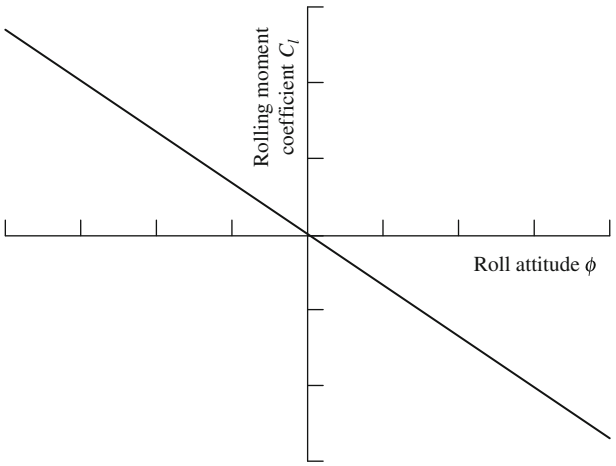


Figure 4.3-16 C_l–ϕ plot for a stable aircraft.

component $U \cong U_e$. The increase in angle of attack on the leading wing gives rise to an increase in lift which in turn gives rise to a restoring rolling moment $-L$. The corresponding aerodynamic change on the wing trailing into the sideslip results in a small decrease in lift which also produces a restoring rolling moment. The net effect therefore is to create a negative rolling moment which causes the aircraft to recover its zero sideslip wings level equilibrium. Thus, the condition for an aircraft to be laterally stable is that the rolling moment resulting from a positive disturbance in roll attitude must be negative, or in mathematical terms:

$$\frac{dC_l}{d\phi} < 0 \tag{4.3.34}$$

where C_l is the rolling moment coefficient. This is shown graphically in Fig. 4.3-16 and may be interpreted in a similar way to the pitching moment plot shown in Fig. 4.3-2.

The sequence of events following a sideslip disturbance are shown for a laterally stable, neutrally stable and unstable aircraft on Fig. 4.3-17. However, it must be remembered that once disturbed the subsequent motion will be determined by the lateral dynamic stability characteristics as well.

4.3.5 Directional Static Stability

Directional static stability is concerned with the ability of the aircraft to yaw or *weathercock* into wind in order to maintain directional equilibrium. Since all aircraft are required to fly with zero sideslip in the yaw sense, positive directional stability is designed in from the outset. The fin is the most visible contributor to directional static stability although, as in the case of lateral stability, there are many other contributions, some of which are destabilising.

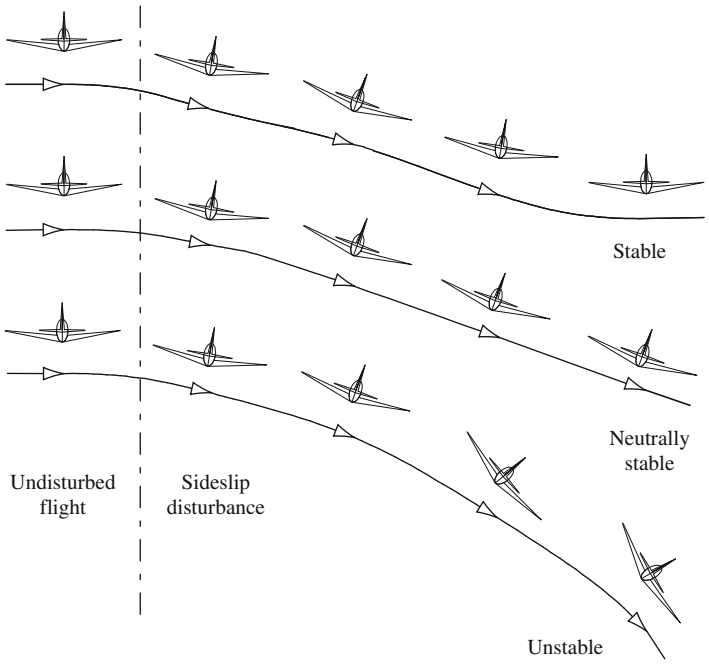


Figure 4.3-17 The effect of dihedral on lateral stability.

Again, it is useful to remember that too much directional static stability will result in an aircraft that is reluctant to manoeuvre directionally, so it is important to obtain the correct degree of stability.

Consider an aircraft that is subject to a positive sideslip disturbance as shown in Fig. 4.3-18. The combination of sideslip velocity v and axial velocity component U results in a positive sideslip angle β. Note that a positive sideslip angle equates to a negative yaw angle since the nose of the aircraft has swung to the left of the resultant total velocity vector V. Now, as shown in Fig. 4.3-18, in the disturbance the fin is at a non-zero angle of attack equivalent to the sideslip angle β. The fin therefore generates lift L_F which acts in the sense shown thereby creating a positive yawing moment N. The yawing moment is stabilising since it causes the aircraft to yaw to the right until the sideslip angle is reduced to zero. Thus, the condition for an aircraft to be directionally stable is readily established and is

$$\frac{dC_n}{d\psi} < 0 \quad \text{or, equivalently,} \quad \frac{dC_n}{d\beta} > 0 \tag{4.3.35}$$

where C_n is the yawing moment coefficient.

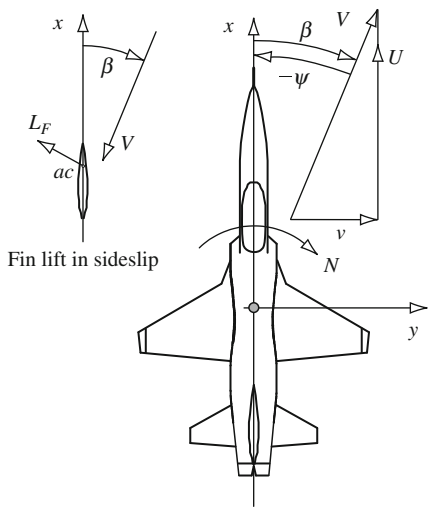


Figure 4.3-18 Directional weathercock effect.

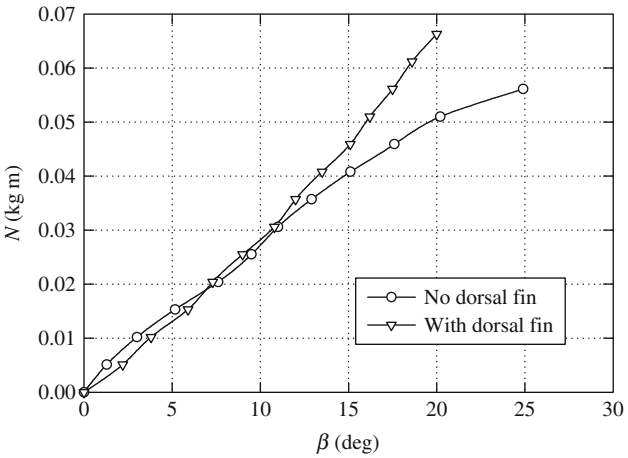


Figure 4.3-19 Plot of yawing moment against sideslip for a stable aircraft.

A typical plot of yawing moment against sideslip angle for a directionally stable aircraft is shown in Fig. 4.3-19. The plots show the results of a wind tunnel test on a simple conventional aircraft model. For small disturbances in yaw the plot is reasonably linear since it is dominated by the lifting properties of the fin. However, as the fin approaches the stall its lifting properties deteriorate and other influences begin to dominate resulting ultimately in loss of directional stability. The main destabilising contribution comes from the fuselage which at small yaw angles is masked by the powerful fin effect. The addition of a dorsal fin significantly delays the onset of fin stall thereby enabling directional static stability to be maintained at higher yaw disturbance angles as indicated in Fig. 4.3-19.

Fin effectiveness also deteriorates with increasing body incidence angle since the base of the fin becomes increasingly immersed in the fuselage wake thereby reducing the effective working area of the fin. This problem has become particularly evident in a number of modern combat aircraft. Typically, such aircraft have two engines mounted side by side in the rear fuselage. This results in a broad flat fuselage ahead of the fin which creates a substantial wake to dramatically reduce fin effectiveness at moderate to high angles of incidence. For this reason many aircraft of this type have noticeably large fins and in some cases the aircraft have two fins attached to the outer edges of the upper fuselage.

4.3.6 Calculation of Aircraft Trim Condition

As described in Section 4.3.1, the condition for an aircraft to remain in steady trimmed flight requires that the forces and moments acting on the aircraft sum to zero and that it is stable. Thus, in order to calculate the trim condition of an aircraft it is convenient to assume straight or symmetric flight and to apply the principles described earlier in this chapter. For a given aircraft mass, *cg* position, altitude and airspeed, symmetric trim is described by the aerodynamic operating condition, namely angle of attack, thrust, pitch attitude, elevator angle and flight path angle. Other operating condition parameters can then be derived as required.

The forces and moments acting on an aeroplane in the general case of steady symmetric climbing flight are shown in Fig. 4.3-20 where the symbols have their usual meanings. Since the aircraft is symmetric, the lateral–directional forces and moments are assumed to remain in equilibrium throughout, and the problem reduces to the establishment

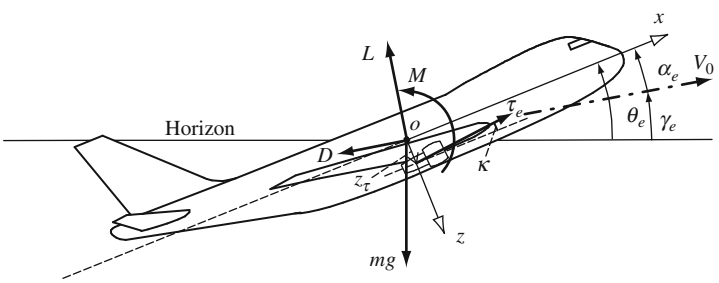


Figure 4.3-20 Symmetric forces and moments acting on a trimmed aircraft.

of longitudinal equilibrium only. Thus, the reference axes are aircraft body axes which define the plane of symmetry oxz, with the origin o located at the aircraft cg as shown.

4.3.6.1 Defining the Trim Condition

The total axial force X is given by resolving the total lift L, total drag D, weight mg and thrust τ_e into the ox axis and these components must sum to zero in trim. Whence

$$X = L\sin\alpha_e + \tau_e\cos\kappa - D\cos\alpha_e - mg\sin(\alpha_e + \gamma_e) = 0 \tag{4.3.36}$$

where α_e is the equilibrium body incidence, γ_e is the steady flight path angle and k is the inclination of the thrust line to the ox body axis (positive nose up). Similarly, the total normal force Z is given by resolving the forces into the oz axis and these also must sum to zero in trim. Whence

$$Z = mg\cos(\alpha_e + \gamma_e) - L\cos\alpha_e - D\sin\alpha_e - \tau_e\sin\kappa = 0 \tag{4.3.37}$$

The development of the aerodynamic pitching moment about the cg is described in Section 4.3.2 and is given by equation (4.3.6). However, since the total pitching moment is required, equation (4.3.6) must be modified to include the thrust, and any other significant, moment contributions. As before, the total drag moment is assumed insignificant since the normal displacement between the cg and aerodynamic centre is typically small for most aircraft configurations. Also, the tailplane zero lift pitching moment M_T is assumed small since the aerofoil section is usually symmetrical and the tailplane drag moment is very small since the tailplane setting would be designed to trim at small local incidence angle. Thus, the total pitching moment about the cg is given by the sum of the wing–body, tailplane and thrust moments, and these moments must sum to zero in trim. Whence

$$M = M_0 + L_w(h - h_0)\bar{c} - L_T l_T + \tau_e z_\tau = 0 \tag{4.3.38}$$

where L_w is the wing–body lift and L_T is the tailplane lift. The other symbols are evident from Fig. 4.3-20. It is convenient to write equations (4.3.36)–(4.3.38) in coefficient form

$$\frac{mg}{\frac{1}{2}\rho V_0^2 S}\sin(\alpha_e + \gamma_e) = C_\tau\cos\kappa + C_L\sin\alpha_e - C_D\cos\alpha_e \tag{4.3.39}$$

$$\frac{mg}{\frac{1}{2}\rho V_0^2 S}\cos(\alpha_e + \gamma_e) = C_L\cos\alpha_e + C_D\sin\alpha_e + C_\tau\sin\kappa \tag{4.3.40}$$

$$0 = C_{m_0} + (h - h_0)C_{L_w} - \overline{V}_T C_{L_T} + \frac{z_\tau}{\bar{c}}C_\tau = 0 \tag{4.3.41}$$

where the thrust coefficient is given by

$$C_\tau = \frac{\tau_e}{\frac{1}{2}\rho V_0^2 S} \tag{4.3.42}$$

the total lift coefficient is given by

$$C_L = C_{L_w} + \frac{S_T}{S}C_{L_T} \tag{4.3.43}$$

and the total drag coefficient is given by

$$C_D = C_{D_0} + \frac{1}{\pi A e}C_L^2 \equiv C_{D_0} + KC_L^2 \tag{4.3.44}$$

The wing–body lift coefficient, which is assumed to comprise wing aerodynamic properties only, is given by

$$C_{L_w} = a(\alpha_w - \alpha_{w0}) \equiv a(\alpha_e + \alpha_{wr} - \alpha_{w0}) \tag{4.3.45}$$

where α_{wr} is the wing rigging angle as shown in Fig. 4.3-21 and α_{wo} is the zero lift angle of attack of the wing.

Simultaneous solution of equations (4.3.39)–(4.3.45) for a given flight condition determines the values of the aerodynamic coefficients and the body incidence defining the aircraft trim state.

4.3.6.2 Elevator Angle to Trim

Once the trim condition is determined, the important elevator angle to trim can be derived along with other useful trim variables. However, the basic aerodynamic relationships described earlier represent the simplest possible

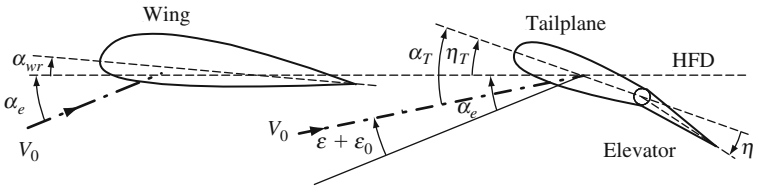


Figure 4.3-21 Practical wing–tailplane aerodynamic geometry.

definitions in the interests of functional visibility. For a practical aircraft application it is necessary to take additional contributions into account when assembling the defining equations. For example, the wing–tail aerodynamic relationship will be modified by the constraints of a practical layout as illustrated in Fig. 4.3-21. The illustration is of course a modified version of that shown in Fig. 4.3-9 to include the wing rigging angle α_{wr} and a zero lift downwash term ε_0.

The aircraft fixed reference for the angle definitions is the *horizontal fuselage datum* (HFD) which is usually a convenient *base line* or *centre line* for the aircraft geometric layout. It is convenient to define the aircraft *ox* body axis parallel to the HFD, with its origin located at the *cg*, and this is shown in Fig. 4.3-22.

With reference to Fig. 4.3-21 it is seen that wing angle of attack is given by

$$\alpha_w = \alpha_e + \alpha_{wr} \tag{4.3.46}$$

and tailplane angle of attack is given by

$$\alpha_T = \eta_T + \alpha_e - \varepsilon - \varepsilon_0 = \eta_T + \alpha_w - \alpha_{wr} - \varepsilon - \varepsilon_0 \tag{4.3.47}$$

With reference to equation (4.3.10):

$$\alpha_w - \varepsilon = \alpha_w \left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) \tag{4.3.48}$$

and equation (4.3.47) may be written:

$$\alpha_T = \eta_T + \alpha_w \left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) - \alpha_{wr} - \varepsilon_0 \tag{4.3.49}$$

It is assumed that the elevator trim tab angle is zero and that aircraft trim is determined by the elevator angle to trim η_e. As before, it is assumed that $a_0 = 0$ since the tailplane aerofoil section is typically symmetrical. The tailplane lift coefficient given by equation (4.3.8) may therefore be re-stated with the appropriate substitution of equation (4.3.49):

$$C_{L_T} = a_1 \alpha_T + a_2 \eta_e = a_1 \left(\eta_T + \alpha_w \left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) - \alpha_{wr} - \varepsilon_0\right) + a_2 \eta_e \tag{4.3.50}$$

Thus, the elevator angle to trim follows by rearrangement of equation (4.3.50):

$$\eta_e = \frac{C_{L_T}}{a_2} - \frac{a_1}{a_2}\left(\eta_T + \alpha_w \left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) - \alpha_{wr} - \varepsilon_0\right) \tag{4.3.51}$$

Note that equation (4.3.51) is equivalent to equation (4.3.13).

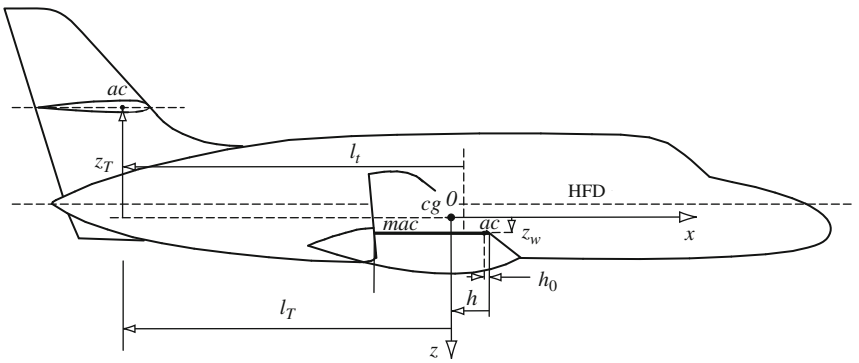


Figure 4.3-22 Practical aircraft longitudinal geometry.

4.3.6.3 Controls Fixed Static Stability

The location of the controls fixed neutral point on the mean aerodynamic chord and the controls fixed static margin are very important parameters in any aircraft trim assessment, since they both influence the aerodynamic, thrust and control requirements for achieving trim. In practice, the achievement of a satisfactory range of elevator angles to trim over the flight envelope is determined by the static margin, and this in turn places constraints on the permitted range of *cg* positions. The neutral point usually determines the most aft *cg* limit in a stable aircraft. Fortunately, the simple expressions given by equations (4.3.17) and (4.3.18) are sufficient for most practical assessment and they are repeated here for convenience. The neutral point location h_n is given by

$$h_n = h_0 + \overline{V}_T \frac{a_1}{a}\left(1 - \frac{\mathrm{d}\varepsilon}{\mathrm{d}\alpha}\right) \tag{4.3.52}$$

and the static margin K_n is given by

$$K_n = h_n - h \tag{4.3.53}$$

Estimation of the wing–body aerodynamic centre location h_0 on the mean aerodynamic chord requires careful consideration. For a subsonic wing, typically $h_0 = 0.25$ and for the purpose of illustrating the simple theory in Section 4.3.3 this value is often assumed, incorrectly, to apply to a wing–body combination. However, the presence of the fuselage usually causes a forward shift of the combined wing–body aerodynamic centre to a value more like $h_0 = 0.1$, or less. Clearly, this has an impact on the requirements for trim and it is important to obtain the best estimate of its location. This can be done by wind tunnel tests on a wing–body combination, or more conveniently by reference to empirical data sources. Estimation of h_0 is described in ESDU 92024, Volume 4b in the ESDU Aerodynamics Series (2006).

Estimation of the rate of change of downwash angle at the tail with wing angle of attack is another parameter that requires careful estimation for the same reasons. Typical values are in the region of $\mathrm{d}\varepsilon/\mathrm{d}\alpha \approx 0.5$, but the geometric location of the tailplane with respect to the wing strongly influences the actual value. Again, a value can be estimated by wind tunnel test of a suitable model. Alternatively, $\mathrm{d}\varepsilon/\mathrm{d}\alpha$ can be estimated with the aid of ESDU 80020, Volume 9a in the ESDU Aerodynamics Series (2006). A simple computer program for estimating $\mathrm{d}\varepsilon/\mathrm{d}\alpha$ may be found in Stribling (1984), and the use of the program is illustrated in the next section.

4.3.6.4 "AeroTrim": a *Mathcad* Trim Program

A computer program called "AeroTrim" has been written by the author in the *Mathcad* language to implement the trim calculations described above. Since *Mathcad* permits the development of programs in the format of a mathematical document, the listing is easy to read and is self-explanatory. Because of its computational visibility *Mathcad* is an ideal tool for programs of this type, although it could be written in a number of alternative languages. AeroTrim is a simple generic trim calculator and is limited to subsonic flight at altitudes up to 36,000 ft. However, it is very easy for the user to modify the program to suit particular requirements and it should be regarded as a foundation for such further development. Indeed, the author has produced versions of the program to deal with transonic flight conditions, aircraft performance and versions substantially extended to include aerodynamic derivative estimation.

The program includes numerical data for the Cranfield University Jetstream 31 flying laboratory aircraft. To use the program for other aircraft applications it is necessary only to delete and replace the numerical data where prompted to do so. Although based on simple mathematical models, the program produces plausible estimates for the known trim characteristics of the Jetstream, but the small differences from observed practice are thought to be due mainly to propeller effects which are notoriously difficult to model adequately.

With the program loaded into *Mathcad*, operation is as simple as clicking on the calculate button. Thus the impact on trim of changing one or more of the numerical input values can be evaluated instantaneously. Points to note include:

Section 1 The user inputs flight condition data for which a trim evaluation is required.

Section 2 Calculates atmospheric temperature, air density and density ratio for the chosen altitude based on the ISA model. Currently limited to the troposphere, but easily modified to include the stratosphere.

Section 3 The user defines the velocity range over which the trim conditions are required, but bearing in mind that the computations are only valid for subsonic flight conditions. The counter sets the number of velocity steps through the range, currently set at 10. The range expression sets the starting velocity, currently set at 100 kt, and the increment, currently set at 15 kt.

Section 4 The user inserts values for the aircraft geometry constants taking care to observe the body axis system used. All of this information would be readily available in a dimensioned three-view drawing of the aircraft.

Section 5 The user inputs values for the principal wing–body aerodynamic parameters for the aircraft. Unknowns obviously have to be estimated by whatever means are available.

Section 6 Repeats Section 5 for the tailplane aerodynamic parameters.

Section 7 Calculates some basic wing–body–tail parameters.

Section 8 Estimates $d\varepsilon/d\alpha$ for the given aircraft geometry using a simple algorithm described by Stribling (1984). Since the model does not include fuselage interference effects or thrust effects it may underestimate the parameter by a small amount. However, results obtained with the algorithm would seem to be plausible and appropriate.

Section 9 Estimates the induced drag factor K in the drag polar $C_D = C_{D_0} + KC_L^2$ using an empirical method described in Shevell (1989), which is based on industrial flight test experience. The very limited data for the fuselage drag factor s_d and the empirical constant k_D were plotted and curves were fitted to give expressions suitable for inclusion in the computation. Results obtained for the Jetstream compare very favourably with the known drag properties of the aircraft.

Section 10 Calculates some useful standard performance and stability parameters.

Section 11 Contains the trim calculation, which solves equations (4.3.39)–(4.3.45) simultaneously for each velocity step defined in Section 3.

Section 12 Calculates the dependent trim variables, including elevator angle, for the velocity steps defined in Section 3 and using the results of Section 11.

Sections 13 and 14 Contain self-explanatory auxiliary computations.

Section 15 Results. Gives a summary of the flight condition parameters for the chosen application.

Section 16 Results. Gives a tabulated summary of the trim values of all the variables at each velocity step in the range chosen.

Section 17 Results. Shows some plotted variables to illustrate the kind of output *Mathcad* can provide. It is very easy to edit this section to include plots of any variables from the table in Section 16.

Example 4.3.4

To illustrate the use of AeroTrim it is applied to the Cranfield University Jetstream 31 flying laboratory aircraft. Since a comprehensive flight simulation model of the aircraft has been assembled and matched to observed flight behaviour the numerical data are believed to be reasonably representative of the actual aircraft. The sources of data used include manufacturer's published technical information, flight manual, limited original wind tunnel test data and data obtained from flight experiments. Aerodynamic data not provided by any of these sources were estimated using the ESDU Aerodynamics Series (2006) and refined by reference to observed flight behaviour.

The chosen operating condition is typical for the aircraft and the speed range was chosen to vary from the stall, at around 100 kt, to 250 kt in 15 kt increments. Good quality data for the remaining input parameters were available, with the possible exception of the values for wing–body aerodynamic centre position h_0, and the rate of change of downwash at the tail with wing angle of attack $d\varepsilon/d\alpha$. Both parameters were estimated for the aircraft, although the actual value for $d\varepsilon/d\alpha$ is thought to be larger than the value estimated by the programme. Using the value $d\varepsilon/d\alpha = 0.279$ as calculated, the value of $h_0 = -0.08$ was estimated since it returned values for the neutral point

Flight condition

	Units	Value
Aircraft weight	kN	61.8
Altitude	ft	6562
Flight path angle	deg	0
cg position		0.29
Neutral point		0.412
Static margin		0.122
Minimum drag speed	kt	150
Stall speed	kt	116

position h_n and static margin K_n close to their known values. It is likely that this places the aerodynamic centre too far forward in the aircraft. However, with a value of $d\varepsilon/d\alpha$ nearer to its probable value, $d\varepsilon/d\alpha \cong 0.4$, a more aft aerodynamic centre position would return the known stability properties. This illustrates one of the difficulties of getting reliable aerodynamic data together for an aircraft, and for unconventional configurations the difficulties are generally greater.

Running the programme returns trim data for the chosen operating flight condition, of which a reduced selection is shown.

Example trim data

V_{true} (knots)	C_L	C_D	C_τ	L/D	α_e (deg)	η_e (deg)	L (kN)	D (kN)	τ_e (kN)
100	1.799	0.174	0.181	9.409	15.105	−1.208	60.23	5.834	6.042
115	1.374	0.114	0.116	11.017	10.885	−0.460	60.83	5.053	5.146
130	1.081	0.082	0.083	12.106	7.970	0.100	61.15	4.643	4.688
145	0.872	0.064	0.064	12.603	5.885	0.521	61.34	4.494	4.518
160	0.717	0.053	0.053	12.573	4.346	0.842	61.46	4.535	4.548
175	0.600	0.046	0.046	12.154	3.181	1.091	61.54	4.722	4.729
190	0.510	0.042	0.042	11.496	2.277	1.287	61.60	5.025	5.029
205	0.438	0.039	0.039	10.720	1.564	1.444	61.65	5.424	5.426
220	0.381	0.036	0.036	9.912	0.990	1.572	61.70	5.907	5.908
235	0.334	0.035	0.035	9.123	0.523	1.677	61.74	6.465	6.465
250	0.295	0.034	0.034	8.383	0.136	1.764	61.79	7.089	7.089

For the purpose of trim analysis the data can be graphed as required and some examples are given in the *Mathcad* program listing. It follows that the effect of any aerodynamic variable on aircraft design performance can be evaluated quickly using the program. Indeed, this approach was used to identify plausible values for some of the more uncertain values in the model definition.

References

Babister, A.W. 1961: *Aircraft Stability and Control.* Pergamon Press, London.

Duncan, W.J. 1959: *The Principles of the Control and Stability of Aircraft.* Cambridge University Press.

ESDU Aerodynamics Series. 2006. Engineering Sciences Data, ESDU International Ltd., 27 Corsham Street, London. www.esdu.com.

Gates, S.B. and Lyon, H.M. 1944: *A Continuation of Longitudinal Stability and Control Analysis; Part 1, General Theory.* Aeronautical Research Council, Reports and Memoranda No. 2027. Her Majesty's Stationery Office, London.

Mathcad. Adept Scientific, Amor Way, Letchworth, Herts, SG6 1ZA. www.adeptscience co.uk

Shevell, R.S. 1989: *Fundamentals of Flight,* Second Edition. Prentice Hall Inc., New Jersey, USA.

Storey, R.F.R. 1966: *H.P.137. Longitudinal and Lateral Stability Measurements on a 1/6th Scale Model.* W.T. Report No. 3021, BAC (Operating) Ltd., Weybridge, Surrey.

Stribling, C.B. 1984: *BASIC Aerodynamics.* Butterworth & Co (Publishers) Ltd., London.

Problems

1. Explain why the pitching moment coefficient $C_{m_{ac}}$ about the aerodynamic centre of an aerofoil is constant. What is the special condition for $C_{m_{ac}}$ to be zero?

   The NACA 64-412 is a cambered aerofoil with lift coefficient given by

   $$C_L = 0.11\alpha + 0.3$$

   when α is in degree units. What is the value of the constant pitching moment coefficient about the aerodynamic centre? Estimate the position of the centre of pressure for the aerofoil at an angle of attack of 5°. State all assumptions made in answering this question.

   (CU 1998)

2. What are the conditions for the stable longitudinal trim equilibrium of an aircraft? The pitching moment coefficient about the *cg* for a stable aircraft is given by

   $$C_m = C_{m_0} + C_{L_w}(h - h_0) - \overline{V}_T\left(C_{L_w}\frac{a_1}{a}\left(1 - \frac{d\varepsilon}{d\alpha}\right) + a_2\eta\right)$$

where the symbols have the usual meaning. Derive expressions for the controls fixed static margin K_n and the elevator angle to trim as a function of static margin. Explain the physical meaning of the controls fixed neutral point.

(CU 1998)

3. State the conditions required for an aeroplane to remain in longitudinal trimmed equilibrium in steady level flight. The pitching moment equation, referred to the centre of gravity (*cg*), for a canard configured combat aircraft is given by

$$C_m = C_{m_0} + \left(h - h_0\right) C_{L_{wb}} + \overline{V}_f \left(\frac{a_{1f}}{a_{wb}} C_{L_{wb}} + a_{1f}\delta\right)$$

where the symbols have the usual meaning and, additionally, $\overline{V}_f$ is the foreplane volume ratio, a_{1f} is the foreplane lift curve slope and δ is the control angle of the all moving foreplane. Derive expressions for the controls fixed static margin and for the controls fixed neutral point. State any assumptions made.

Given that the mean aerodynamic chord (*mac*) is 4.7 m, the wing–body aerodynamic centre is located at 15% of *mac*, the foreplane volume ratio is 0.12 and the lift curve slope of the wing–body and foreplane are 3.5 and 4.9 1/rad respectively, calculate the aft *cg* limit for the aircraft to remain stable with controls fixed. Calculate also the *cg* location for the aircraft to have a controls fixed static margin of 15%.

(CU 1999)

4. Sketch a typical C_m–α plot and explain the condition for trim, the requirement for static stability and the concept of stability margin. Why is too much stability as hazardous as too little stability?

(CU 2001)

Aerospace Engineering Desk Reference; ISBN 9781856175753

4.4 Longitudinal Dynamics

Michael Cook

4.4.1 Response to Controls

The solution of the longitudinal equations of motion enables the response transfer functions to be obtained. These completely describe the linear dynamic response to a control input in the plane of symmetry. Implicit in the response are the dynamic properties determined by the stability characteristics of the aeroplane. The transfer functions and the response variables described by them are linear since the entire modelling process is based on the assumption that the motion is constrained to small disturbances about an equilibrium trim state. However, it is common practice to assume that the response to controls is valid when the magnitude of the response can hardly be described as "a small perturbation". For many conventional aeroplanes the error incurred by so doing is generally acceptably small as such aeroplanes tend to have substantially linear aerodynamic characteristics over their flight envelopes. For aeroplanes with very large flight envelopes, significant aerodynamic non-linearity and, or, dependence on sophisticated flight control systems, it is advisable not to use the linearised equations of motion for analysis of response other than that which can justifiably be described as being of small magnitude.

It is convenient to review the longitudinal response to elevator about a trim state in which the thrust is held constant. The longitudinal state equation may then be written as

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} x_u & x_w & x_q & x_\theta \\ z_u & z_w & z_q & z_\theta \\ m_u & m_w & m_q & m_\theta \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} x_\eta \\ z_\eta \\ m_\eta \\ 0 \end{bmatrix} \eta \tag{4.4.1}$$

The four response transfer functions obtained in the solution of equation (4.4.1) may conveniently be written as

$$\frac{u(s)}{\eta(s)} \equiv \frac{N_\eta^u(s)}{\Delta(s)} = \frac{k_u(s + 1/T_u)(s^2 + 2\zeta_u\omega_u s + \omega_u^2)}{(s^2 + 2\zeta_p\omega_p s + \omega_p^2)(s^2 + 2\zeta_s\omega_s s + \omega_s^2)} \tag{4.4.2}$$

$$\frac{w(s)}{\eta(s)} \equiv \frac{N_\eta^w(s)}{\Delta(s)} = \frac{k_w(s + 1/T_\alpha)(s^2 + 2\zeta_\alpha\omega_\alpha s + \omega_\alpha^2)}{(s^2 + 2\zeta_p\omega_p s + \omega_p^2)(s^2 + 2\zeta_s\omega_s s + \omega_s^2)} \tag{4.4.3}$$

$$\frac{q(s)}{\eta(s)} \equiv \frac{N_\eta^q(s)}{\Delta(s)} = \frac{k_q s(s + (1/T_{\theta_1}))(s + (1/T_{\theta_2}))}{(s^2 + 2\zeta_p\omega_p s + \omega_p^2)(s^2 + 2\zeta_s\omega_s s + \omega_s^2)} \tag{4.4.4}$$

$$\frac{\theta(s)}{\eta(s)} \equiv \frac{N_\eta^\theta(s)}{\Delta(s)} = \frac{k_\theta(s + (1/T_{\theta_1}))(s + (1/T_{\theta_2}))}{(s^2 + 2\zeta_p\omega_p s + \omega_p^2)(s^2 + 2\zeta_s\omega_s s + \omega_s^2)} \tag{4.4.5}$$

The solution of the equations of motion results in polynomial descriptions of the transfer function numerators and common denominator. The polynomials factorise into real and complex pairs of roots that are most explicitly quoted in the style of equations (4.4.2)–(4.4.5) above. Since the roots are interpreted as time constants, damping ratios and natural frequencies, the above style of writing makes the essential information instantly available. It should also be noted that the numerator and denominator factors are typical for a conventional aeroplane. Sometimes complex pairs of roots may become two real roots and *vice versa*. However, this does not usually mean that the dynamic response characteristics of the aeroplane become dramatically different. Differences in the interpretation of response may be evident but will not necessarily be large.

As has already been indicated, the common denominator of the transfer functions describes the characteristic polynomial which, in turn, describes the stability characteristics of the aeroplane. Thus the response of all variables to an elevator input is dominated by the denominator parameters namely, damping ratios and natural frequencies. The differences between the individual responses are entirely determined by their respective numerators. It is therefore important to fully appreciate the role of the numerator in determining response dynamics. The *response shapes* of the individual variables are determined by the common denominator and "coloured" by their respective numerators. The numerator plays no part in determining stability in a linear system which is how the aeroplane is modelled here.

Example 4.4.1

The equations of motion and aerodynamic data for the Ling-Temco-Vought A-7A Corsair II aircraft were obtained from Teper (1969). The flight condition corresponds to level cruising flight at an altitude of 15,000 ft at Mach 0.3. The equations of motion, referred to a body axis system, arranged in state space format are

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} 0.00501 & 0.00464 & -72.90000 & -31.34000 \\ -0.08570 & -0.54500 & 309.00000 & -7.40000 \\ 0.00185 & -0.00767 & -0.39500 & 0.00132 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} 5.63000 \\ -23.80000 \\ -4.51576 \\ 0 \end{bmatrix} \eta \tag{4.4.6}$$

Since incidence α and flight path angle γ are useful variables in the evaluation of handling qualities, it is convenient to augment the corresponding output equation in order to obtain their response transfer functions in the solution of the equations of motion. The output equation is therefore,

$$\begin{bmatrix} u \\ w \\ q \\ \theta \\ \alpha \\ \gamma \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0.00316 & 0 & 0 \\ 0 & -0.00316 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \eta \tag{4.4.7}$$

Note that all elements in the matrices in equations (4.4.6) and (4.4.7) have been rounded to five decimal places simply to keep the equations to a responsible physical size. This should not be done with the equations used in the actual computation.

Solution of the equations of motion using *Program CC* determines the following response transfer function:

$$\begin{aligned}
\frac{u(s)}{\eta(s)} &= \frac{5.63(s+0.369)(s+0.587)(s+58.437)}{(s^2+0.033s+0.020)(s^2+0.902s+2.666)} \text{ ft/s/rad} \\
\frac{w(s)}{\eta(s)} &= \frac{-23.8(s^2-0.0088s+0.0098)(s+59.048)}{(s^2+0.033s+0.020)(s^2+0.902s+2.666)} \text{ ft/s/rad} \\
\frac{q(s)}{\eta(s)} &= \frac{-4.516s(s-0.008)(s+0.506)}{(s^2+0.033s+0.020)(s^2+0.902s+2.666)} \text{ rad/s/rad (deg/s/deg)} \\
\frac{\theta(s)}{\eta(s)} &= \frac{-4.516(s-0.008)(s+0.506)}{(s^2+0.033s+0.020)(s^2+0.902s+2.666)} \text{ rad/rad (deg/deg)} \\
\frac{\alpha(s)}{\eta(s)} &= \frac{-0.075(s^2-0.0088s+0.0098)(s+59.048)}{(s^2+0.033s+0.020)(s^2+0.902s+2.666)} \text{ rad/rad (deg/deg)} \\
\frac{\gamma(s)}{\eta(s)} &= \frac{0.075(s-0.027)(s+5.004)(s-6.084)}{(s^2+0.033s+0.020)(s^2+0.902s+2.666)} \text{ rad/rad (deg/deg)}
\end{aligned} \tag{4.4.8}$$

All coefficients have again been rounded to a convenient number of decimal places and the above caution should be noted.

The characteristic equation is given by equating the common denominator polynomial to zero:

$$\Delta(s) = (s^2+0.33s+0.020)(s^2+0.902s+2.666) = 0$$

The first pair of complex roots describes the phugoid stability mode, with characteristics:

Damping ratio	$\zeta_p = 0.11$
Undamped natural frequency	$\omega_p = 0.14$ rad/s

The second pair of complex roots describes the short period pitching oscillation, or short period stability mode, with characteristics:

Damping ratio	$\zeta_s = 0.28$
Undamped natural frequency	$\omega_s = 1.63$ rad/s

These mode characteristics indicate that the airframe is aerodynamically stable although it will be shown later that the short-period mode damping ratio is unacceptably low.

The response of the aircraft to a unit step (1°) elevator input is shown in Fig. 4.4-1. All of the variables in the solution of the equations of motion are shown the responses being characterised by the transfer functions, equations (4.4.8).

The responses clearly show both dynamic stability modes, the short period pitching oscillation and the phugoid. However, the magnitude of each stability mode differs in each response variable. For example, the short period pitching oscillation is most visible as the initial transient in the variables w, q and α whereas the phugoid mode is visible in all variables although the relative magnitudes vary considerably. Clearly the stability of the responses is the same, as determined by the common denominator of the transfer functions, equations (4.4.8), but the differences between each of the response variables are determined by the unique numerator of each response transfer function.

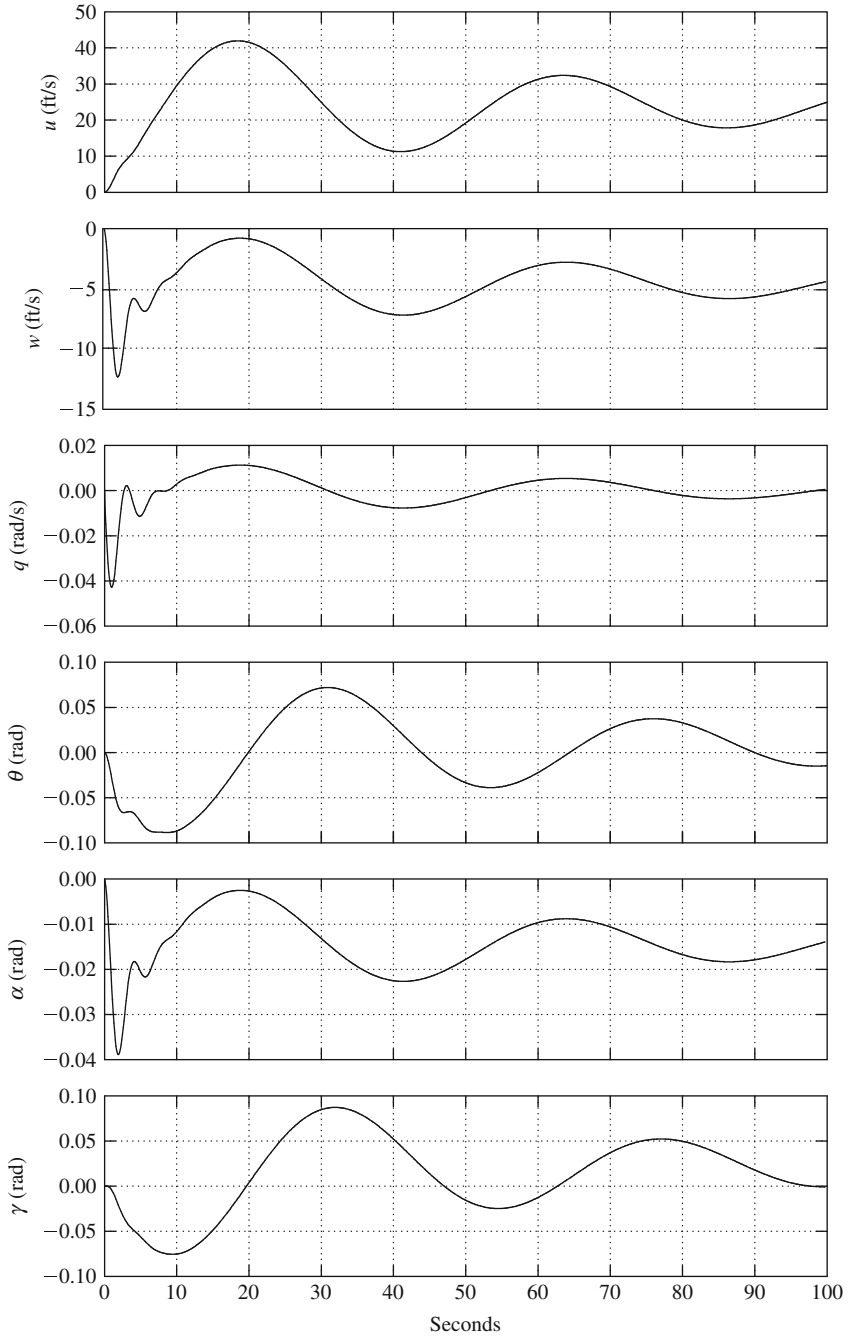


Figure 4.4-1 Aircraft response to 1° elevator step input.

The mode content in each of the motion variables is given most precisely by the eigenvectors. The analytical procedure is applied to the equations of motion for the A-7A. With the aid of *MATLAB* the eigenvector matrix $\mathbf{V}$ is determined as follows:

$$\mathbf{V} = \begin{array}{c} \begin{matrix} \text{Short period mode} & & \text{Phugoid mode} & \end{matrix} \\ \left[\begin{array}{cc|cc} -0.1682-0.1302j & -0.1682+0.1302j & 0.1467+0.9677j & 0.1467-0.9677j \\ 0.2993+0.9301j & 0.2993-0.9301j & 0.0410+0.2008j & 0.0410-0.2008j \\ -0.0046+0.0018j & -0.0046-0.0018j & 0.0001+0.0006j & 0.0001-0.0006j \\ 0.0019+0.0024j & 0.0019-0.0024j & 0.0041-0.0013j & 0.0041+0.0013j \end{array}\right] \end{array} \begin{array}{l} : u \\ : w \\ : q \\ : \theta \end{array} \tag{4.4.9}$$

To facilitate interpretation of the eigenvector matrix, the magnitude of each component eigenvector is calculated as follows:

$$\left[\begin{array}{cc|cc} 0.213 & 0.213 & 0.979 & 0.979 \\ 0.977 & 0.977 & 0.204 & 0.204 \\ 0.0049 & 0.0049 & 0.0006 & 0.0006 \\ 0.0036 & 0.0036 & 0.0043 & 0.0043 \end{array}\right] \begin{array}{l} : u \\ : w \\ : q \\ : \theta \end{array}$$

Clearly, the phugoid mode is dominant in u since $0.979 \gg 0.213$, the short period mode is dominant in w since $0.977 \gg 0.204$, the short period mode is dominant in q since $0.0049 \gg 0.0006$ and the short period and phugoid modes content in θ are of a similar order. These observations accord very well with the responses shown in Fig. 4.4-1.

The steady state values of the variables following a unit step (1°) elevator input may be determined by application of the final value theorem. The transfer functions, equations (4.4.8), assume a unit elevator displacement to mean 1 rad and this has transfer function:

$$\eta(s) = \frac{1}{s} \text{ rad}$$

For a unit step input of 1° the transfer function becomes

$$\eta(s) = \frac{1}{57.3s} = \frac{0.0175}{s} \text{ rad}$$

Thus, for example, the Laplace transform of the speed response to a 1° elevator step input is given by

$$u(s) = \frac{5.63(s+0.369)(s+0.587)(s+58.437)}{(s^2+0.033s+0.020)(s^2+0.902s+2.666)} \frac{0.0175}{s} \text{ ft/s}$$

Applying the final value theorem, equation (5.33):

$$\begin{aligned} u(t)|_{ss} &= \lim_{s\to 0} \left(s \frac{5.63(s+0.369)(s+0.587)(s+58.437)}{(s^2+0.033s+0.020)(s^2+0.902s+2.666)} \frac{0.0175}{s} \right) \text{ft/s} \\ &= 23.39 \text{ft/s} \end{aligned}$$

Since the step input is positive in the nose down sense the response eventually settles to the small steady increase in speed indicated.

In a similar way the steady state response of all the motion variables may be calculated to give

$$\begin{bmatrix} u \\ w \\ q \\ \theta \\ \alpha \\ \gamma \end{bmatrix}_{\text{steady state}} = \begin{bmatrix} 23.39 \text{ ft/s} \\ -4.53 \text{ ft/s} \\ 0 \\ 0.34° \\ -0.81° \\ 1.15° \end{bmatrix} \tag{4.4.10}$$

It is important to remember that the steady state values given in equation (4.4.10) represent the *changes* with respect to the initial equilibrium trim state following the 1° elevator step input. Although the initial response is applied in the nose down sense, inspection of equation (4.4.10) indicates that after the mode transients have damped out the aircraft is left with a small reduction in incidence, a small increase in pitch attitude and is climbing steadily at a flight path angle of 1.15°. This apparent anomaly is due to the fact that at the chosen flight condition the aircraft is operating close to the stall boundary on the *back side* of the drag-speed curve, that is, below the minimum drag speed. Thus the disturbance results in a significant decrease in drag leaving the aircraft with sufficient excess power enabling it to climb gently. It is for the same reason that a number of the transfer functions (4.4.8) have non-minimum phase numerator terms where these would not normally be expected.

4.4.1.1 The Characteristic Equation

The longitudinal characteristic polynomial for a classical aeroplane is fourth order; it determines the common denominator in the longitudinal response transfer functions and, when equated to zero, defines the characteristic equation which may be written as

$$As^4 + Bs^3 + Cs^2 + Ds + E = 0 \tag{4.4.11}$$

The characteristic equation (4.4.11) most commonly factorises into two pairs of complex roots which are most conveniently written as

$$(s^2 + 2\zeta_p\omega_p s + \omega_p^2)(s^2 + 2\zeta_s\omega_s s + \omega_s^2) = 0 \tag{4.4.12}$$

As already explained, the second order characteristics in equation (4.4.12) describe the phugoid and short period stability modes respectively. The stability modes comprising equation (4.4.12) provide a *complete description* of the longitudinal stability properties of the aeroplane subject to the constraint of small perturbation motion. Interpretation of the characteristic equation written in this way is most readily accomplished if reference is first made to the properties of the classical mechanical mass-spring-damper system.

Thus the longitudinal dynamics of the aeroplane may be likened to a pair of loosely coupled mass-spring-damper systems and the interpretation of the motion of the aeroplane following a disturbance from equilibrium may be made by direct comparison with the behaviour of the mechanical mass-spring-damper. However, the damping and frequency characteristics of the aeroplane are obviously not mechanical in origin, they derive entirely from the aerodynamic properties of the airframe. The connection between the observed dynamics of the aeroplane and its aerodynamic characteristics is made by comparing equation (4.4.12) with equation (4.4.11). Clearly, the relationships between the damping ratios and undamped frequencies of equation (4.4.12) and their aerodynamic *drivers* are neither obvious nor simple. Means for dealing with this difficulty are described below in which simplifying approximations are made based on the observation and understanding of the physical behaviour of aeroplane dynamics.

4.4.2 The Dynamic Stability Modes

Both longitudinal dynamic stability modes are excited whenever the aeroplane is disturbed from its equilibrium trim state. A disturbance may be initiated by pilot control inputs, a change in power setting, airframe configuration changes such as flap deployment and by external atmospheric influences such as gusts and turbulence.

4.4.2.1 The Short Period Pitching Oscillation

The short period mode is typically a damped oscillation in pitch about the *oy* axis. Whenever an aircraft is disturbed from its pitch equilibrium state the mode is excited and manifests itself as a classical second order oscillation in which the principal variables are incidence $\alpha(w)$, pitch rate q and pitch attitude θ. This observation is easily confirmed by reference to the eigenvectors in the solution of the equations of motion; this may be seen in Example 4.4.1 and also in Fig. 4.4-1. Typically the undamped natural frequency of the mode is in the range 1–10 rad/s and the damping is usually stabilising although the damping ratio is often lower than desired. A significant feature of the mode is that the speed remains approximately constant ($u = 0$) during a disturbance. As the period of the mode is short, inertia and momentum effects ensure that speed response in the time scale of the mode is negligible.

The physical situation applying can be interpreted by comparison with a torsional mass-spring-damper system. The aircraft behaves as if it were restrained by a torsional spring about the *oy* axis as indicated in Fig. 4.4-2. A pitch disturbance from trim equilibrium causes the "spring" to produce a restoring moment thereby giving rise to an oscillation in pitch. The oscillation is damped and this can be interpreted as a viscous damper as suggested in Fig. 4.4-2. Of course the spring and viscous damping effects are not mechanical. In reality they are produced entirely by aerodynamic mechanisms with contributions from all parts of the airframe, not all of which are necessarily stabilising in effect. However, in the interests of promoting understanding, the stiffness and damping effects are assumed to be dominated by the aerodynamics of the tailplane. The spring stiffness arises from the natural *weathercock* tendency of the tailplane to align with the incident flow. The damping arises from the motion of the tailplane during the oscillation when, clearly, it behaves as a kind of viscous paddle damper. The total observed mode dynamics depend not only on the tailplane contribution, but also on the magnitudes of the additional contributions from other parts of the airframe. When the overall stability is marginal it is implied that the additional contributions are also significant and it becomes much more difficult to identify and quantify the principal aerodynamic mode drivers.

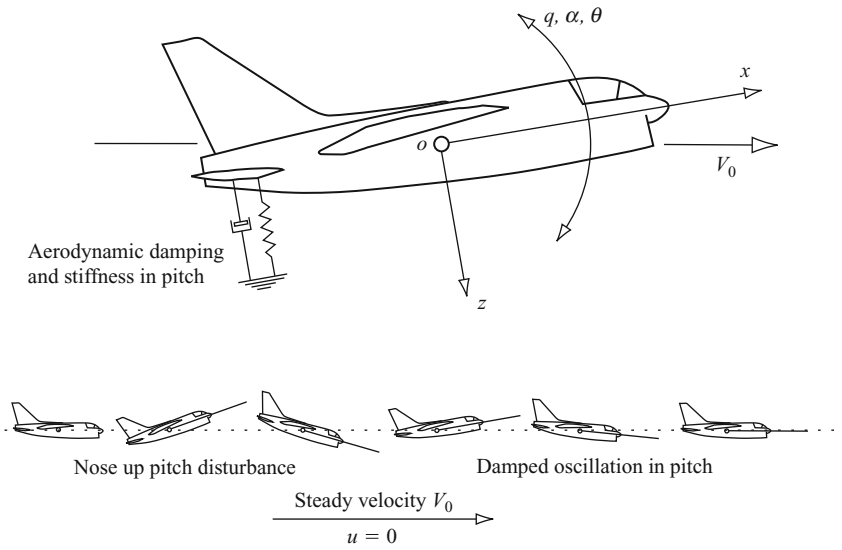


Figure 4.4-2 A stable short period pitching oscillation.

4.4.2.2 The Phugoid

The phugoid mode is most commonly a lightly damped low frequency oscillation in speed u which couples into pitch attitude θ and height h. A significant feature of this mode is that the incidence $\alpha(w)$ remains substantially constant during a disturbance. Again, these observations are easily confirmed by reference to the eigenvectors in the solution of the equations of motion; this may be seen in Example 4.4.1 and also in Fig. 4.4-1. However, it is clear that the phugoid appears, to a greater or lesser extent, in all of the longitudinal motion variables but the relative magnitudes of the phugoid components in incidence $\alpha(w)$ and in pitch rate q are very small. Typically, the undamped natural frequency of the phugoid is in the range 0.1–1 rad/s and the damping ratio is very low. However, the apparent damping characteristics of the mode may be substantially influenced by power effects in some aeroplanes.

Consider the development of classical phugoid motion following a small disturbance in speed as shown in Fig. 4.4-3. Initially the aeroplane is in trimmed level equilibrium flight with steady velocity V_0 such that the lift L and weight mg are equal. Let the aeroplane be disturbed at (a) such that the velocity is reduced by a small amount u. Since the incidence remains substantially constant this results in a small reduction in lift such that the aeroplane is no longer in vertical equilibrium. It therefore starts to lose height and since it is flying "down hill" it starts to accelerate as at (b). The speed continues to build up to a value in excess of V_0 which is accompanied by a build up in lift which eventually exceeds the weight by a significant margin. The build up in speed and lift cause the aircraft to pitch up steadily until at (c) it starts to climb. Since it now has an excess of kinetic energy, inertia and momentum effects cause it to fly up through the nominal trimmed height datum at (d) losing speed and lift as it goes as it is now flying "up hill". As it decelerates it pitches down steadily until at (e) its lift is significantly less than the weight and the accelerating descent starts again. Inertia and momentum effects cause the aeroplane to continue flying down through the nominal trimmed height datum (f) and as the speed and lift continue to build up so it pitches up steadily until at (g) it starts climbing again to commence the next cycle of oscillation. As the motion progresses the effects of drag cause the

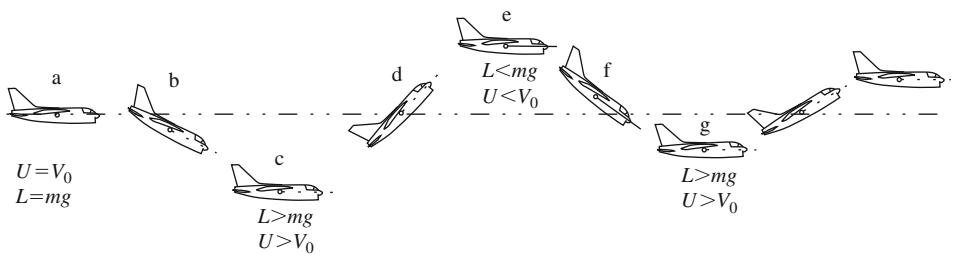


Figure 4.4-3 The development of a stable phugoid.

motion variable maxima and minima at each peak to reduce gradually in magnitude until the motion eventually damps out.

Thus the phugoid is classical damped harmonic motion resulting in the aircraft flying a gentle sinusoidal flight path about the nominal trimmed height datum. As large inertia and momentum effects are involved the motion is necessarily relatively slow such that the angular accelerations, $\dot{q}$ and $\dot{\alpha}(\dot{w})$, are insignificantly small. Consequently, the natural frequency of the mode is low and since drag is designed to be low so the damping is also low. Typically, once excited many cycles of the phugoid may be visible before it eventually damps out. Since the rate of loss of energy is low, a consequence of low drag damping effects, the motion is often approximated by undamped harmonic motion in which potential and kinetic energy are exchanged as the aircraft flies the sinusoidal flight path. This in fact was the basis on which Lanchester (1908) first successfully analysed the motion.

4.4.3 Reduced Order Models

Thus far the emphasis has been on the exact solution of the longitudinal equations of motion which results in an exact description of the stability and response characteristics of the aircraft. Although this is usually the object of a flight dynamics investigation it has two disadvantages. First, a computational facility is required if a very tedious manual solution is to be avoided and, second, it is difficult, if not impossible, to establish the relationships between the stability characteristics and their aerodynamic drivers. Both these disadvantages can be avoided by seeking approximate solutions that can also provide considerable insight into the physical phenomena governing the dynamic behaviour of the aircraft.

For example, an approximate solution of the longitudinal characteristic equation (4.4.11) is based on the fact that the coefficients A, B, C, D and E have relative values that do not change very much for conventional aeroplanes. Generally A, B and C are significantly larger than D and E such that the quartic has the following approximate factors:

$$A\left(s^2+\frac{(CD-BE)}{C^2}s+\frac{E}{C}\right)\left(s^2+\frac{B}{A}s+\frac{C}{A}\right)=0 \tag{4.4.13}$$

Equation (4.4.13) is in fact the first step in the classical manual iterative solution of the quartic; the first pair of complex roots describes the phugoid and the second pair describes the short period mode. As the algebraic expressions are relatively complex further physical insight is not particularly revealing unless simplifying assumptions are made. However, the approximate solution given by equation (4.4.13) is often useful for preliminary mode evaluations, or as a check of computer solutions, when the numerical values of the coefficients A, B, C, D and E are known. For conventional aeroplanes the approximate solution is often surprisingly close to the exact solution of the characteristic equation.

4.4.3.1 The Short Period Mode Approximation

The short term response characteristics of an aircraft are of particular importance in flying and handling qualities considerations for the reasons stated in Section 4.4.5. Since short term behaviour is dominated by the short period mode it is convenient to obtain the *reduced-order* equations of motion in which the phugoid is suppressed or omitted. By observing the nature of the short period pitching oscillation, sometimes called the *rapid incidence adjustment*, it is possible to simplify the longitudinal equations of motion to describe short term dynamics only. The terms remaining in the reduced-order equations of motion are therefore the terms that dominate short term dynamics thereby providing insight into the important aerodynamic drivers governing physical behaviour.

It has already been established that the short period pitching oscillation is almost exclusively an oscillation in which the principal variables are pitch rate q and incidence α, the speed remaining essentially constant; thus, $u=0$. Therefore, the speed equation and the speed dependent terms may be removed from the longitudinal equations of motion 4.4.1; since they are all approximately zero in short term motion, the revised equations may be written as

$$\begin{bmatrix}\dot{w}\\ \dot{q}\\ \dot{\theta}\end{bmatrix}=\begin{bmatrix}z_w & z_q & z_\theta\\ m_w & m_q & m_\theta\\ 0 & 1 & 0\end{bmatrix}\begin{bmatrix}w\\ q\\ \theta\end{bmatrix}+\begin{bmatrix}z_\eta\\ m_\eta\\ 0\end{bmatrix}\eta \tag{4.4.14}$$

Further, assuming the equations of motion are referred to aircraft wind axes and that the aircraft is initially in steady level flight then

$$\theta_e \equiv \alpha_e = 0 \quad \text{and} \quad U_e = V_0$$

and it follows that

$$z_\theta = m_\theta = 0$$

Equation (4.4.14) then reduces to its simplest possible form:

$$\begin{bmatrix} \dot{w} \\ \dot{q} \end{bmatrix} = \begin{bmatrix} z_w & z_q \\ m_w & m_q \end{bmatrix} \begin{bmatrix} w \\ q \end{bmatrix} + \begin{bmatrix} z_\eta \\ m_\eta \end{bmatrix} \eta \tag{4.4.15}$$

where now, the derivatives are referred to a wind axes system. Equation (4.4.15) is sufficiently simple that the transfer function matrix may be calculated manually:

$$\begin{aligned} \mathbf{G}(s) = \frac{\mathbf{N}(s)}{\Delta(s)} &= \frac{\begin{bmatrix} s - m_q & z_q \\ m_w & s - z_w \end{bmatrix} \begin{bmatrix} z_\eta \\ m_\eta \end{bmatrix}}{\begin{vmatrix} s - z_w & -z_q \\ -m_w & s - m_q \end{vmatrix}} \\ &= \frac{\begin{bmatrix} z_\eta \left(s + \left(m_q + z_q \dfrac{m_\eta}{z_\eta} \right) \right) \\ m_\eta \left(s + \left(m_w \dfrac{z_\eta}{m_\eta} - z_w \right) \right) \end{bmatrix}}{(s^2 - (m_q + z_w)s + (m_q z_w - m_w z_q))} \end{aligned} \tag{4.4.16}$$

The transfer functions may be further simplified by noting that

$$z_q \frac{m_\eta}{z_\eta} \gg m_q \quad \text{and} \quad -z_w \gg m_w \frac{z_\eta}{m_\eta}$$

and

$$z_q = \frac{\mathring{Z}_q + mU_e}{m - \mathring{Z}_{\dot{w}}} \cong U_e$$

since

$$\mathring{Z}_q \ll mU_e \quad \text{and} \quad m \gg \mathring{Z}_{\dot{w}}$$

Thus the two short term transfer functions describing response to elevator may be written as

$$\frac{w(s)}{\eta(s)} = \frac{z_\eta \left(s + U_e \dfrac{m_\eta}{z_\eta} \right)}{(s^2 - (m_q + z_w)s + (m_q z_w - m_w U_e))} \equiv \frac{k_w (s + 1/T_\alpha)}{(s^2 + 2\zeta_s \omega_s s + \omega_s^2)} \tag{4.4.17}$$

$$\frac{q(s)}{\eta(s)} = \frac{m_\eta (s - z_w)}{(s^2 - (m_q + z_w)s + (m_q z_w - m_w U_e))} \equiv \frac{k_q (s + 1/T_{\theta_2})}{(s^2 + 2\zeta_s \omega_s s + \omega_s^2)} \tag{4.4.18}$$

where now it is understood that $k_w, k_q, T_\alpha, T_{\theta_2}, \zeta_s$ and ω_s represent approximate values. Clearly it is now very much easier to relate the most important parameters describing longitudinal short term transient dynamics of the aircraft to the aerodynamic properties of the airframe, represented in equations (4.4.17) and (4.4.18) by the concise derivatives.

The reduced order characteristic equation may be written down on inspection of equation (4.4.17) or (4.4.18):

$$\Delta(s) = s^2 + 2\zeta_s \omega_s s + \omega_s^2 = s^2 - (m_q + z_w)s + (m_q z_w - m_w U_e) = 0 \tag{4.4.19}$$

and the damping and natural frequency of the short period mode are given, to a good approximation, by

$$\begin{aligned} 2\zeta_s \omega_s &= -(m_q + z_w) \\ \omega_s &= \sqrt{m_q z_w - m_w U_e} \end{aligned} \tag{4.4.20}$$

It is instructive to write the damping and natural frequency expressions (4.4.20) in terms of the dimensional derivatives. The assumptions made above are applied to give

$$2\zeta_s\omega_s = -\left(\frac{\mathring{M}_q}{I_y} + \frac{\mathring{Z}_w}{m} + \frac{\mathring{M}_{\dot{w}}U_e}{I_y}\right)$$
$$\omega_s = \sqrt{\frac{\mathring{M}_q}{I_y}\frac{\mathring{Z}_w}{m} - \frac{\mathring{M}_w U_e}{I_y}} \tag{4.4.21}$$

Note that the terms on the right hand side of expressions (4.4.21) comprise aerodynamic derivatives divided either by mass or moment of inertia in pitch. These terms may be interpreted in exactly the same way as those of the classical mass-spring-damper. Thus, it becomes apparent that the aerodynamic derivatives are providing stiffness and viscous damping in pitch although there is more than one term contributing to damping and to natural frequency. Therefore the aerodynamic origins of the short period dynamics are a little more complex than those of the classical mass-spring-damper and the various contributions do not always act in the most advantageous way. However, for conventional aeroplanes the overall dynamic characteristics usually describe a stable short period mode.

For a typical conventional aeroplane the relative magnitudes of the aerodynamic derivatives are such that to a crude approximation:

$$2\zeta_s\omega_s = \frac{-\mathring{M}_q}{I_y}$$
$$\omega_s = \sqrt{\frac{-\mathring{M}_w U_e}{I_y}} \tag{4.4.22}$$

which serves only to indicate what are usually regarded as the dominant terms governing the short period mode. Normally the derivative $\mathring{Z}_w$, which is dependent on the lift curve slope of the wing, and the derivative $\mathring{M}_q$, which is determined largely by the viscous "paddle" damping properties of the tailplane, are both negative numbers. The derivative $\mathring{M}_w$ is a measure of the aerodynamic stiffness in pitch and is also dominated by the aerodynamics of the tailplane. The sign of $\mathring{M}_w$ depends on the position of the *cg*, becoming increasingly negative as the *cg* moves forward in the airframe. Thus the short period mode will be stable if the *cg* is far enough forward in the airframe. The *cg* position in the airframe where $\mathring{M}_w$ changes sign is called the *controls fixed neutral point* and $\mathring{M}_w$ is therefore also a measure of the *controls fixed stability margin* of the aircraft. With reference to equation (4.4.19) and expressions (4.4.20), the corresponding *cg* position where $(m_q z_w - m_w U_e)$ changes sign is called the *controls fixed manoeuvre point* and $(m_q z_w - m_w U_e)$ is a measure of the *controls fixed manoeuvre margin* of the aircraft.

4.4.3.2 The Phugoid Mode Approximation

A reduced order model of the aircraft retaining only the phugoid dynamics is very rarely required in flight dynamics studies. However, the greatest usefulness of such a model is to identify those aerodynamic properties of the airframe governing the characteristics of the mode.

4.4.3.2.1 The Lanchester model

Probably the first successful analysis of aeroplane dynamics was made by Lanchester (1908) who devised a mathematical model to describe phugoid motion based on his observations of the behaviour of gliding model aeroplanes. His analysis gives excellent insight into the physical nature of the mode and may be applied to the modern aeroplane by interpreting and restating his original assumptions as follows:

(i) The aircraft is initially in steady level flight.
(ii) The total energy of the aircraft remains constant.
(iii) The incidence α remains constant at its initial trim value.
(iv) The thrust τ balances the drag D.
(v) The motion is sufficiently slow that pitch rate q effects may be ignored.

Referring to Fig. 4.4-4 the aircraft is initially in trimmed straight level flight with velocity V_0. Following a disturbance in speed which excites the phugoid mode, the disturbed speed, pitch attitude and height are denoted V, θ and h respectively. Then based on assumption (ii):

$$\frac{1}{2}mV_0^2 = \frac{1}{2}mV^2 + mgh = \text{constant}$$

whence

$$V^2 = V_0^2 - 2gh \tag{4.4.23}$$

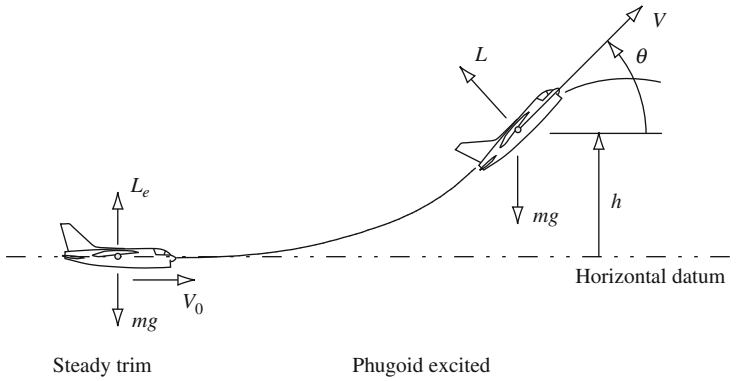


Figure 4.4-4 The phugoid oscillation.

which describes the exchange of kinetic and potential energy as the aeroplane flies the sinusoidal flight path.

In the initial steady trim state the lift and weight are in balance; thus,

$$L_e = \frac{1}{2}\rho V_0^2 S C_L = mg \tag{4.4.24}$$

and in disturbed flight the lift is given by

$$L = \frac{1}{2}\rho V^2 S C_L \tag{4.4.25}$$

As a consequence of assumption (iii) the lift coefficient C_L also remains constant and equations (4.4.23)–(4.4.25) may be combined to give

$$L = mg - \rho g h S C_L \tag{4.4.26}$$

Since simple undamped oscillatory motion is assumed, a consequence of assumption (ii), the single degree of freedom equation of motion in height may be written as

$$m\ddot{h} = L\cos\theta - mg \cong L - mg \tag{4.4.27}$$

since, by definition, θ is a small angle. Substituting for lift L from equation (4.4.26) into equation (4.4.27):

$$\ddot{h} + \left(\frac{\rho g S C_L}{m}\right) h = \ddot{h} + \omega_p^2 h = 0 \tag{4.4.28}$$

Thus, approximately, the frequency of the phugoid mode is given by

$$\omega_p = \sqrt{\frac{\rho g S C_L}{m}} = \frac{g\sqrt{2}}{V_0} \tag{4.4.29}$$

when equation (4.4.24) is used to eliminate the mass.

Thus, to a reasonable approximation, Lanchester's model shows that the phugoid frequency is inversely proportional to the steady trimmed speed about which the mode oscillates and that its damping is zero.

4.4.3.2.2 A reduced order model

A more detailed approximate model of the phugoid mode may be derived from the equations of motion by making simplifications based on assumptions about the nature of the motion. Following a disturbance, the variables $w(\alpha)$ and q respond in the time scale associated with the short period mode; thus, it is reasonable to assume that $w(\alpha)$ and q are quasi-steady in the longer time scale associated with the phugoid. Whence, it follows that

$$\dot{w} = \dot{q} = 0$$

Once again, it is assumed that the equations of motion are referred to aircraft wind axes and since the disturbance takes place about steady level flight then

$$\theta_e \equiv \alpha_e = 0 \quad \text{and} \quad U_e = V_0$$

and it follows that

$$x_\theta = -g \quad \text{and} \quad z_\theta = m_\theta = 0$$

Also, as for the reduced order short period model

$$z_q = \frac{\mathring{Z}_q + mU_e}{m - \mathring{Z}_{\dot{w}}} \cong U_e$$

since

$$\mathring{Z}_q \ll mU_e \quad \text{and} \quad m \gg \mathring{Z}_{\dot{w}}$$

Additionally, it is usually assumed that the aerodynamic derivative x_q is insignificantly small. Thus the equations of motion (4.4.1) may be simplified accordingly:

$$\begin{bmatrix} \dot{u} \\ 0 \\ 0 \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} x_u & x_w & 0 & -g \\ z_u & z_w & U_e & 0 \\ m_u & m_w & m_q & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} x_\eta \\ z_\eta \\ m_\eta \\ 0 \end{bmatrix} \eta \tag{4.4.30}$$

The second and third rows of equation (4.4.30) may be written as

$$\begin{aligned} z_u u + z_w w + U_e q + z_\eta \eta &= 0 \\ m_u u + m_w w + m_q q + m_\eta \eta &= 0 \end{aligned} \tag{4.4.31}$$

Equations (4.4.31) may be solved algebraically to obtain expressions for w and q in terms of u and η:

$$\begin{aligned} w &= \left(\frac{m_u U_e - m_q z_u}{m_q z_w - m_w U_e}\right) u + \left(\frac{m_\eta U_e - m_q z_\eta}{m_q z_w - m_w U_e}\right) \eta \\ q &= \left(\frac{m_w z_u - m_u z_u}{m_q z_w - m_w U_e}\right) u + \left(\frac{m_w z_\eta - m_\eta z_w}{m_q z_w - m_w U_e}\right) \eta \end{aligned} \tag{4.4.32}$$

The expressions for w and q are substituted into rows one and four of equation (4.4.30) and following some rearrangement the reduced order state equation is obtained:

$$\begin{bmatrix} \dot{u} \\ \dot{\theta} \end{bmatrix} = \left[\begin{array}{c|c} x_u - x_w\left(\dfrac{m_u U_e - m_q z_u}{m_w U_e - m_q z_w}\right) & -g \\ \hline \left(\dfrac{m_u z_w - m_w z_u}{m_w U_e - m_q z_w}\right) & 0 \end{array}\right] \begin{bmatrix} u \\ \theta \end{bmatrix} + \begin{bmatrix} x_\eta - \left(\dfrac{m_\eta U_e - m_q z_\eta}{m_w U_e - m_q z_w}\right) \\ \left(\dfrac{m_\eta z_w - m_w z_\eta}{m_w U_e - m_q z_w}\right) \end{bmatrix} \eta \tag{4.4.33}$$

or

$$\dot{\mathbf{x}} = \mathbf{A}_p \mathbf{x} + \mathbf{B}_p \mathbf{u} \tag{4.4.34}$$

Equation (4.4.33) may be solved algebraically to obtain the response transfer functions for the phugoid variables u and θ. However, it is not very meaningful to analyse long term dynamic response to elevator in this way. The characteristic equation describing the reduced order phugoid dynamics is considerably more useful and is given by

$$\Delta(s) = \det[s\mathbf{I} - \mathbf{A}_p] = 0$$

whence

$$\begin{aligned} \Delta(s) &= s^2 + 2\zeta_p w_p s + \omega_p^2 \\ &= s^2 - \left(x_u - x_w\left(\frac{m_u U_e - m_q z_u}{m_w U_e - m_q z_w}\right)\right) s + g\left(\frac{m_u z_w - m_w z_u}{m_w U_e - m_q z_w}\right) \end{aligned} \tag{4.4.35}$$

Thus the approximate damping and natural frequency of the phugoid mode are given in terms of a limited number of aerodynamic derivatives. More explicit, but rather more approximate, insight into the aerodynamic properties of the aeroplane dominating the mode characteristics may be obtained by making some further assumptions. Typically, for conventional aeroplanes in subsonic flight:

$$m_u \to 0, \quad |m_u z_w| \ll |m_w z_u| \quad \text{and} \quad |m_w U_e| \gg |m_q z_w|$$

then the corresponding expressions for the damping and natural frequency become:

$$\begin{aligned} 2\zeta_p\omega_p &= -x_u \\ \omega_p &= \sqrt{\frac{-gz_u}{U_e}} \end{aligned} \tag{4.4.36}$$

Now

$$x_u \cong \frac{\mathring{X}_u}{m} = \frac{\rho V_0 S X_u}{2m} \quad \text{and} \quad z_u \cong \frac{\mathring{Z}_u}{m} = \frac{\rho V_0 S Z_u}{2m} \tag{4.4.37}$$

since $\mathring{X}_{\dot{w}}$ is negligibly small and $m \gg \mathring{Z}_{\dot{w}}$. Expressions for the dimensionless aerodynamic derivatives may be approximated as shown in expressions (4.4.38) when the basic aerodynamic properties are assumed to be independent of speed. This follows from the assumption that the prevailing flight condition is subsonic such that the aerodynamic properties of the airframe are not influenced by compressibility effects:

$$\begin{aligned} X_u &= -2C_D - V_0\frac{\partial C_D}{\partial V} + \left(\frac{1}{\frac{1}{2}\rho V_0 S}\right)\frac{\partial \tau}{\partial V} \cong -2C_D \\ Z_u &= -2C_L - V_0\frac{\partial C_L}{\partial V} \cong -2C_L \end{aligned} \tag{4.4.38}$$

Expressions (4.4.36) may therefore be restated in terms of aerodynamic parameters, assuming again that the trimmed lift is equal to the aircraft weight, to obtain

$$\begin{aligned} \zeta_p\omega_p &= \frac{gC_D}{C_L V_0} \\ \omega_p &= \sqrt{\frac{2g^2}{U_e V_0}} \equiv \frac{g\sqrt{2}}{V_0} \end{aligned} \tag{4.4.39}$$

and a simplified approximate expression for the damping ratio follows:

$$\zeta_p \cong \frac{1}{\sqrt{2}}\left(\frac{C_D}{C_L}\right) \tag{4.4.40}$$

These expressions for damping ratio and natural frequency of the phugoid mode are obviously very approximate since they are the result of many simplifying assumptions. Note that the expression for ω_p is the same as that derived by Lanchester, equation (4.4.29), which indicates that the natural frequency of the phugoid mode is approximately inversely proportional to the trimmed speed. It is also interesting and important to note that the damping ratio of the phugoid mode is approximately inversely proportional to the *lift to drag ratio* of the aeroplane, equation (4.4.40). Since one of the main objectives of aeroplane design is to achieve a high lift to drag ratio it is easy to see why the damping of the phugoid mode is usually very low.

Example 4.4.2

To illustrate the use of reduced order models consider the A-7A Corsair II aircraft of Example 4.4.1 and at the same flight condition. Now the equations of motion in Example 4.4.1 are referred to a body axis system and the use of the reduced order models described above requires the equations of motion referred to a wind, or stability axis system. Thus, using the axis transformation relationships, the stability and control derivatives and inertia parameters referred to wind axes were calculated from the original values, which are of course referred to body axes. The longitudinal state equation was then recalculated to give

$$\begin{bmatrix} \dot{u} \\ \dot{w} \\ \dot{q} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} -0.04225 & -0.11421 & 0 & -32.2 \\ -0.20455 & -0.49774 & 317.48 & 0 \\ 0.00003 & -0.00790 & -0.39499 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} u \\ w \\ q \\ \theta \end{bmatrix} + \begin{bmatrix} 0.00381 \\ -24.4568 \\ -4.51576 \\ 0 \end{bmatrix} \eta$$

The reduced order model corresponding to the short period approximation, as given by equation (4.4.15), is simply taken out of equation (4.4.41) and is written as

$$\begin{bmatrix}\dot{w}\\ \dot{q}\end{bmatrix}=\begin{bmatrix}-0.49774 & 317.48\\ -0.00790 & -0.39499\end{bmatrix}\begin{bmatrix}w\\ q\end{bmatrix}+\begin{bmatrix}-24.4568\\ -4.51576\end{bmatrix}\eta \tag{4.4.41}$$

Solution of the equations of motion 4.4.42 using *Program CC* determines the following reduced order response transfer functions:

$$\begin{aligned}\frac{w(s)}{\eta(s)}&=\frac{-24.457(s+59.015)}{(s^2+0.893s+2.704)}\ \text{ft/s/rad}\\ \frac{q(s)}{\eta(s)}&=\frac{-4.516(s+0.455)}{(s^2+0.893s+2.704)}\ \text{rad/s/rad (deg/s/deg)}\\ \frac{\alpha(s)}{\eta(s)}&=\frac{-0.077(s+59.015)}{(s^2+0.893s+2.704)}\ \text{rad/rad (deg/deg)}\end{aligned} \tag{4.4.42}$$

It is important to remember that these transfer functions describe, approximately, the short term response of those variables that are dominant in short period motion. The corresponding short term pitch attitude response transfer function follows since, for small perturbation motion:

$$\frac{\theta(s)}{\eta(s)}=\frac{1}{s}\frac{q(s)}{\eta(s)}=\frac{-4.516(s+0.455)}{s(s^2+0.893s+2.074)}\ \text{rad/rad (deg/deg)} \tag{4.4.43}$$

From the pitch rate response transfer function in equations (4.4.43) it is readily determined that the steady state pitch rate following a positive unit step elevator input is −0.76 rad/s, which implies that the aircraft pitches continuously until the input is removed. The pitch attitude response transfer function confirms this since, after the short period transient has damped out, the aircraft behaves like a perfect integrator in pitch. This is indicated by the presence of the s term in the denominator of equation (4.4.44). In reality the phugoid dynamics usually prevent this situation developing unless the input is very large and accompanied by a thrust increase that results in a vertical loop manoeuvre. The model described here would be most inappropriate for the analysis of such large amplitude motion.

The common denominator of transfer functions (4.4.43) represents the approximate reduced order short period characteristic polynomial, equation (4.4.19). Thus, approximate values of the damping ratio and undamped natural frequency of the short period mode are easily calculated and are

Damping ratio $\zeta_s=0.27$
Undamping natural frequency $\omega_s=1.64$ rad/s

It will be seen that these values compare very favourably with the exact values given in Example 4.4.1.

Interpretation of the reduced order model is probably best illustrated by observing short term response to an elevator input. The responses to a 1° elevator step input of the variables given in equations (4.4.43) are shown in Fig. 4.4-5. Also shown on the same plots are the corresponding responses of the full aircraft model derived from equation (4.4.41). It is clear that the responses diverge with time, as expected, as no phugoid dynamics are present in the reduced order model. However, for the first ten seconds or so, the comparison is favourable indicating that the reduced order model is acceptable for most short term response studies.

We now turn to the approximate reduced order phugoid mode characteristics. From the state equation referred to wind axes, equation (4.4.41), the required numerical parameters are

$x_u=-0.04225$ 1/s
$z_u=-0.20455$ 1/s
$m_u=0.00003$ rad/ft/s
$U_e\equiv V0=317.48$ ft/s

The simple Lanchester model determines that the damping of the phugoid is zero and that the undamped natural frequency is given by equation (4.4.29). Thus the approximate characteristics of the phugoid mode calculated according to this model are

Damping ratio $\zeta_p=0$
Undamped natural frequency $\omega_p=0.143$ rad/s

The approximate phugoid mode characteristics determined according to the rather more detailed reduced order model are given by equation (4.4.36). Since the chosen flight condition is genuinely subsonic, the derivative m_u is

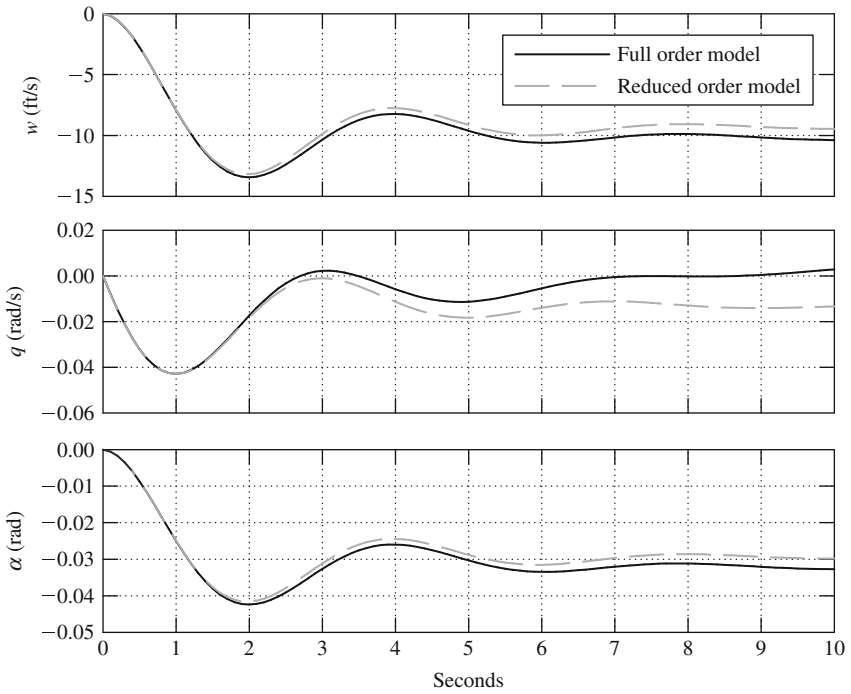


Figure 4.4-5 Reduced order longitudinal response to 1° elevator step input.

very small indeed which matches the constraints of the model well. The approximate characteristics of the phugoid mode calculated according to this model are

Damping ratio $\zeta_p = 0.147$
Undamped natural frequency $\omega_p = 0.144$ rad/s

Again, comparing these approximate values of the phugoid mode characteristics with the exact values in Example 4.4.1 indicates good agreement, especially for the undamped natural frequency. Since the phugoid damping ratio is always small (near to zero) it is very sensitive to computational rounding errors and to the approximating assumptions that make a really good approximate match difficult to achieve. The goodness of the match here is enhanced by the very subsonic flight condition that correlates well with assumptions made in the derivation of the approximate models.

4.4.4 Frequency Response

For the vast majority of flight dynamics investigations time domain analysis is usually adequate, especially when the subject is the classical unaugmented aeroplane. The principal graphical tool used in time domain analysis is, of course, the time history plot showing the response of the aeroplane to controls or to some external disturbance. However, when the subject aeroplane is an advanced modern aeroplane fitted with a flight control system, flight dynamics analysis in the frequency domain can provide additional valuable insight into its behaviour. In recent years frequency domain analysis has made an important contribution to the understanding of the sometimes unconventional handling qualities of aeroplanes whose flying qualities are largely shaped by a flight control system. It is for this reason that a brief review of simple frequency response ideas is considered here. Since frequency response analysis tools are fundamental to classical control engineering their description can be found in almost every book on the subject; very accessible material can be found in Shinners (1980) and Friedland (1987) for example.

Consider the hypothetical situation when the elevator of an otherwise trimmed aeroplane is operated sinusoidally with constant amplitude k and variable frequency ω; the longitudinal input to the aeroplane may therefore be expressed as

$$\eta(t) = k \sin \omega t \tag{4.4.44}$$

It is reasonable to expect that each of the output variables describing aircraft motion will respond sinusoidally to the input. However, the amplitudes of the output variables will not necessarily be the same and they will not

necessarily be in phase with one another or with the input. Thus the general expression describing any output response variable may be written as

$$y(t) = K \sin(\omega t + \phi) \tag{4.4.45}$$

where both the output amplitude K and phase shift ϕ are functions of the exciting frequency ω. As the exciting frequency ω is increased from zero, initially, at low frequencies, the sinusoidal response will be clearly visible in all output variables. As the exciting frequency is increased further, the sinusoidal response will start to diminish in magnitude and will eventually become imperceptible in the outputs. Simultaneously, the phase shift ϕ will indicate an increasingly large lag between the input and output. The reason for these observations is that at sufficiently high frequencies the mass and inertia properties of the aeroplane simply prevent it responding quickly enough to follow the input. The limiting frequency at which the response commences to diminish rapidly is referred to as the *bandwidth* of the aeroplane with respect to the output variable of interest. A more precise definition of bandwidth is given below. Since aeroplanes only respond to frequencies below the bandwidth frequency they have the frequency response properties of a *low pass system*. At exciting frequencies corresponding to the damped natural frequencies of the phugoid and the short period mode, peaks in output magnitude K will be seen together with significant changes in phase shift ϕ. The mode frequencies are described as *resonant frequencies* and the magnitudes of the output parameters K and ϕ at *resonance* are determined by the damping ratios of the modes. The system (or aeroplane) *gain* in any particular response variable is defined as

$$\text{System gain} = \left|\frac{K(\omega)}{k}\right| \tag{4.4.46}$$

where, in normal control system applications, it is usually assumed that the input and output variables have the same units. This is often not the case in aircraft applications and care must be exercised in the interpretation of gain.

A number of graphical tools have been developed for the frequency response analysis of linear systems and include the Nyquist diagram, the Nichols chart and the Bode diagram. All are intended to simplify analytical procedures, the mathematical calculation of which is tedious without a computer, and all plot input–output gain and phase as functions of frequency. Perhaps the simplest of the graphical tools to use and interpret is the Bode diagram although the amount of information it is capable of providing is limited. However, today it is used extensively for flight dynamics analysis, especially in advanced handling qualities studies.

4.4.4.1 The Bode Diagram

The intention here is not to describe the method for constructing a Bode diagram but to describe its application to the aeroplane and to explain its correct interpretation. For an explanation of the method for constructing a Bode diagram the reader should consult a suitable control engineering text, such as either of those referenced above.

To illustrate the application of the Bode diagram to a typical classical aeroplane consider the pitch attitude response to elevator transfer function as given by equation (4.4.5):

$$\frac{\theta(s)}{\eta(s)} = \frac{k_\theta(s + (1/T_{\theta_1}))(s + (1/T_{\theta_2}))}{(s^2 + 2\zeta_p\omega_p S + \omega_p^2)(s^2 + 2\zeta_s\omega_s s + \omega_s^2)} \tag{4.4.47}$$

This response transfer function is of particular relevance to longitudinal handling studies and it has the simplifying advantage that both the input and output variables have the same units. Typically, in frequency response calculation it is usual to assume a sinusoidal input signal of unit magnitude. It is also important to note that whenever the response transfer function is negative, which is often the case in aircraft applications, a negative input is assumed that ensures the correct computation of phase. Therefore, in this particular application, since k_θ is usually a negative number a sinusoidal elevator input of unit magnitude $\eta(t) = -1 \sin \omega t$ is assumed. The pitch attitude frequency response is calculated by writing $s = j\omega$ in equation (4.4.48); the right hand side then becomes a complex number whose magnitude and phase can be evaluated for a suitable range of frequency ω. Since the input magnitude is unity, the *system gain*, equation (4.4.47), is given simply by the absolute value of the magnitude of the complex number representing the right hand side of equation (4.4.48) and is, of course, a function of frequency ω.

Since the calculation of gain and phase involves the products of several complex numbers it is preferred to work in terms of the logarithm of the complex number representing the transfer function. The total gain and phase then become the simple sums of the gain and phase of each factor in the transfer function. For example, each factor in parentheses on the right hand side of equation (4.4.48) may have its gain and phase characteristics calculated separately as a function of frequency; the total gain and phase is then given by summing the contributions from each

factor. However, the system gain is now expressed as a logarithmic function of the gain ratio, equation (4.4.47), and is defined as

$$\text{Logarithmic gain} = 20\log_{10}\left|\frac{K(\omega)}{k}\right|\text{dB} \tag{4.4.48}$$

and has units of *decibels* denoted dB. Fortunately it is no longer necessary to calculate frequency response by hand since many computer software packages, such as *MATLAB*, have this facility and can also provide the desired graphical output. However, as always, some knowledge of the analytical procedure for obtaining frequency response is essential so that the computer output may be correctly interpreted.

The Bode diagram comprises two corresponding plots, the *gain plot* and the *phase plot*. The gain plot shows the logarithmic gain, in dB, plotted against $\log_{10}(\omega)$ and the phase plot shows the phase, in degrees, also plotted against $\log_{10}(\omega)$. To facilitate interpretation the two plots are often superimposed on a common frequency axis. The Bode diagram showing the typical pitch attitude frequency response, as given by transfer function (4.4.48), is shown in Fig. 4.4-6.

Also shown in Fig. 4.4-6 are the asymptotic approximations to the actual gain and phase plots as functions of frequency. The asymptotes can be drawn in simply from inspection of the transfer function, equation (4.4.48), and serve as an aid to interpretation. Quite often the asymptotic approximation is sufficient for the evaluation in hand, thereby dispensing with the need to compute the actual frequency response entirely.

The shape of the gain plot is characterised by the *break frequencies* $\omega 1$ to $\omega 4$ which determine the locations of the discontinuities in the asymptotic gain plot. Each break frequency is defined by a key frequency parameter in the transfer function, namely

$$\begin{aligned}
\omega 1 &= \frac{1}{T_{\theta_1}} \quad \text{with first order phase lead } (+45^\circ)\\
\omega 2 &= \omega_p \quad \text{with second order phase lag } (-90^\circ)\\
\omega 3 &= \frac{1}{T_{\theta_2}} \quad \text{with first order phase lead } (+45^\circ)\\
\omega 4 &= \omega_s \quad \text{with second order phase lag } (-90^\circ)
\end{aligned} \tag{4.4.49}$$

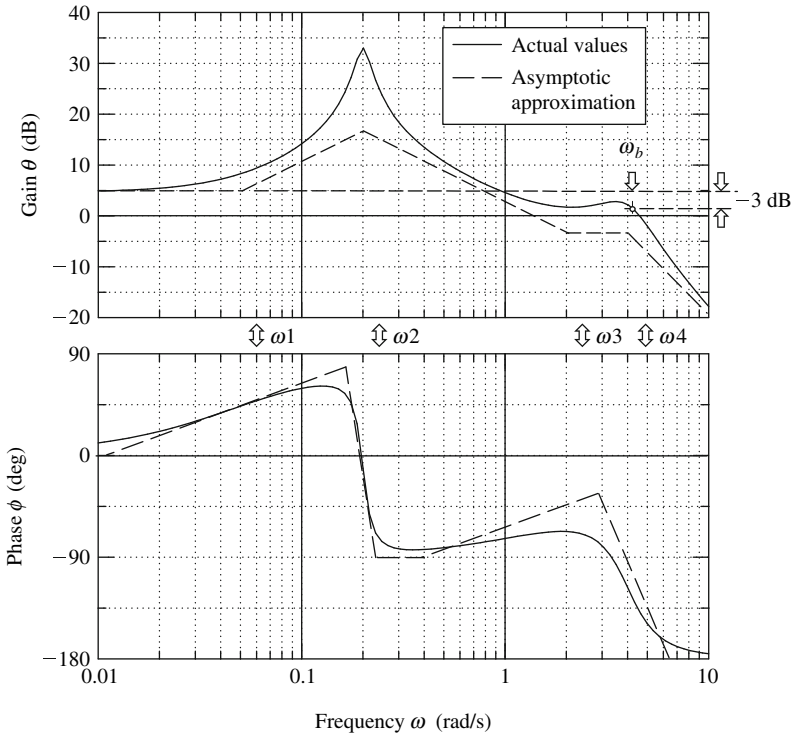


Figure 4.4-6 Bode diagram showing classical pitch attitude frequency response.

Since the transfer function is classical minimum phase, the corresponding phase shift at each break frequency is a *lead* if it arises from a numerator term or a *lag* if it arises from a denominator term. If, as is often the case in aircraft studies, non-minimum phase terms appear in the transfer function, then their frequency response properties are unchanged except that the sign of the phase is reversed. Further, a first order term gives rise to a total phase shift of 90° and a second order term gives rise to a total phase shift of 180°. The characteristic phase response is such that half the total phase shift associated with any particular transfer function factor occurs at the corresponding break frequency. Armed with this limited information a modest interpretation of the pitch attitude frequency response of the aeroplane is possible. The frequency response of the other motion variables may be dealt with in a similar way.

4.4.4.2 Interpretation of the Bode Diagram

With reference to Fig. 4.4-6 it is seen that at very low frequencies, $\omega < 0.01$ rad/s, there is no phase shift between the input and output and the gain remains constant, at a little below 5 dB in this illustration. In other words, the pitch attitude will follow the stick movement more or less precisely. As the input frequency is increased through $\omega 1$ so the pitch response leads the input in phase, the output magnitude increases rapidly and the aeroplane appears to behave like an *amplifier*. At the phugoid frequency the output reaches a substantial peak, consistent with the low damping, and thereafter the gain drops rapidly accompanied by a rapid increase in phase lag. As the input frequency is increased further so the gain continues to reduce gently and the phase settles at –90° until the influence of break frequency $\omega 3$ comes into play. The reduction in gain is arrested and the effect of the phase lead may be seen clearly. However, when the input frequency reaches the short period break frequency a small peak in gain is seen, consistent with the higher damping ratio, and at higher frequencies the gain continues to reduce steadily. Meanwhile, the phase lag associated with the short period mode results in a constant total phase lag of −180° at higher frequencies.

Once the output–input gain ratio drops below unity, or 0 dB, the aeroplane appears to behave like an *attenuator*. The frequency at which the gain becomes sufficiently small that the magnitude of the output response becomes insignificant is called the *bandwidth frequency*, denoted ω_b. There are various definitions of bandwidth, but the definition used here is probably the most common and defines the bandwidth frequency as the frequency at which the gain first drops to −3 dB below the zero frequency, or steady state, gain. The bandwidth frequency is indicated in Fig. 4.4-6 and it is commonly a little higher than the short period frequency. A gain of –3 dB corresponds with a gain ratio of $1/\sqrt{2} = 0.707$. Thus, by definition, the gain at the bandwidth frequency is 0.707 times the steady state gain. Since the pitch attitude bandwidth frequency is close to the short period frequency the latter may sometimes be substituted for the bandwidth frequency which is often good enough for most practical purposes.

The peaks in the gain plot are determined by the characteristics of the stability modes. A very pronounced peak indicates low mode damping and *vice versa*; an infinite peak corresponding with zero damping. The magnitude of the changes in gain and phase occurring in the vicinity of a peak indicates the significance of the mode in the response variable in question. Figure 4.4-6 indicates the magnitude of the phugoid to be much greater than the magnitude of the short period mode in the pitch response of the aeroplane. This would, in fact, be confirmed by response time histories and inspection of the corresponding eigenvectors.

In the classical application of the Bode diagram, as used by the control engineer, inspection of the gain and phase properties in the vicinity of the bandwidth frequency enables conclusions about the stability of the system to be made. Typically, stability is quantified in terms of *gain margin* and *phase margin*. However, such evaluations are only appropriate when the system transfer function is *minimum phase*. Since aircraft transfer functions that are non-minimum phase are frequently encountered, and many also have the added complication that they are negative, it is not usual for aircraft stability to be assessed on the Bode diagram. It is worth noting that, for aircraft augmented with flight control systems, the behaviour of the phase plot in the vicinity of the bandwidth frequency is now known to be linked to the susceptibility of the aircraft to *pilot induced oscillations*, a particularly nasty handling deficiency.

Now the foregoing summary interpretation of frequency response assumes a sinusoidal elevator input to the aircraft. Clearly, this is never likely to occur as a result of normal pilot action. However, normal pilot actions may be interpreted to comprise a mix of many different frequency components. For example, in gentle manoeuvring the frequency content of the input would generally be low whilst in aggressive or high workload situations the frequency content would be higher and might even exceed the bandwidth of the aeroplane. In such a limiting condition the pilot would certainly be aware that the aeroplane could not follow his demands quickly enough and, depending in detail on the gain and phase response properties of the aeroplane, he could well encounter hazardous handling problems. Thus bandwidth is a measure of the *quickness of response* achievable in a given aeroplane. As a general rule it is desirable that flight control system designers should seek the highest response bandwidth consistent with the dynamic capabilities of the airframe.

Example 4.4.3

The longitudinal frequency response of the A-7A Corsair II aircraft is evaluated for the same flight condition as Examples 4.4.1 and 4.4.2. However, the longitudinal response transfer functions used for the evaluations are referred to wind axes and were obtained in the solution of the full order state equation (4.4.41). The transfer functions of primary interest are

$$\frac{u(s)}{\eta(s)} = \frac{0.00381(s+0.214)(s+135.93)(s+598.3)}{(s^2+0.033s+0.02)(s^2+0.902s+2.666)} \text{ ft/s/rad}$$

$$\frac{\theta(s)}{\eta(s)} = \frac{-4.516(s-0.008)(s+0.506)}{(s^2+0.033s+0.02)(s^2+0.902s+2.666)} \text{ rad/rad} \tag{4.4.50}$$

$$\frac{\alpha(s)}{\eta(s)} = \frac{-0.077(s^2+0.042s+0.02)(s+59.016)}{(s^2+0.033s+0.02)(s^2+0.902s+2.666)} \text{ rad/rad}$$

It will be noticed that the values of the various numerator terms in the velocity and incidence transfer functions differ significantly from the values in the corresponding transfer functions in Example 4.4.1, equation (4.4.8). This is due to the different reference axes used and to the fact that the angular difference between body and wind axes is a significant body incidence angle of 13.3°. Such a large angle is consistent with the very low speed flight condition. The frequency response of each transfer function was calculated with the aid of *Program CC* and the Bode diagrams are shown in Figures 4.4-7–4.4-9 respectively. Interpretation of the Bode diagrams for the three variables is straightforward and follows the general interpretation discussed above. However, important or significant differences are commented on as follows.

The frequency response of axial velocity u to elevator input η is shown in Fig. 4.4-7 and it is clear, as might be expected, that it is dominated by the phugoid. The very large low frequency gain values are due entirely to the transfer function units that are ft/s/rad, and a unit radian elevator input is of course unrealistically large! The peak gain of 75 dB at the phugoid frequency corresponds with a gain ratio of approximately 5600 ft/s/rad. However, since

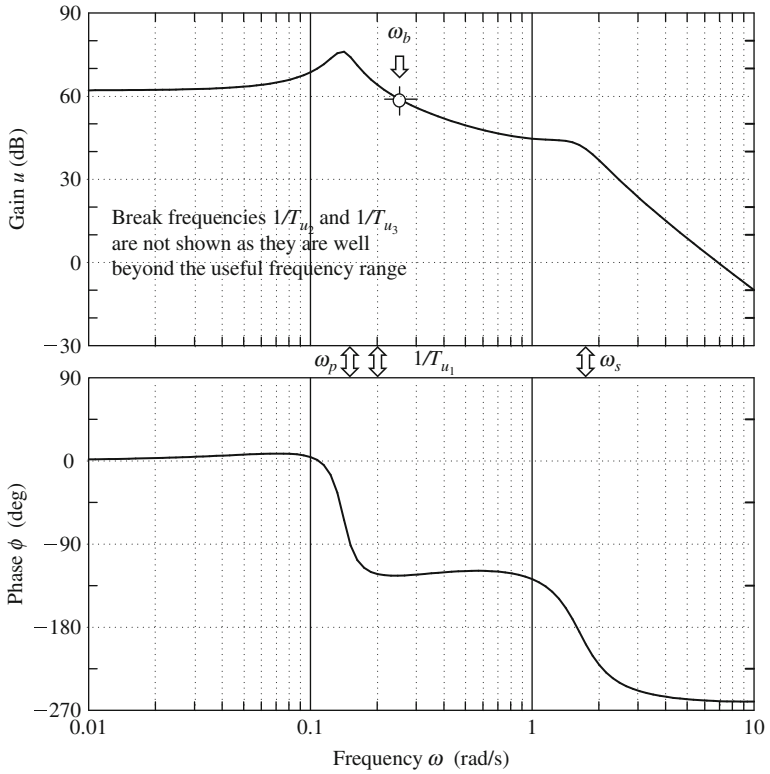


Figure 4.4-7 A-7A velocity frequency response.

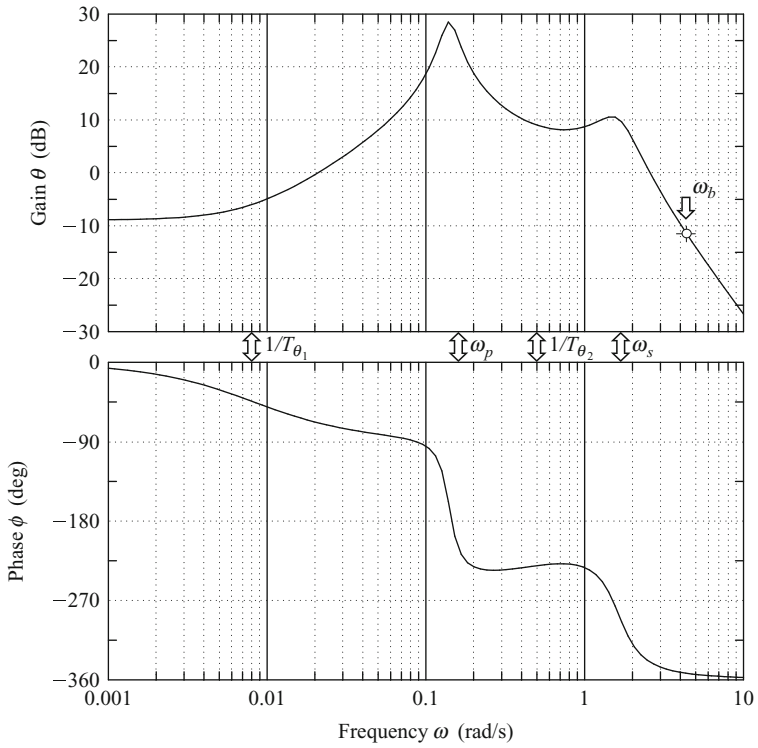


Figure 4.4-8 A-7A pitch attitude frequency response.

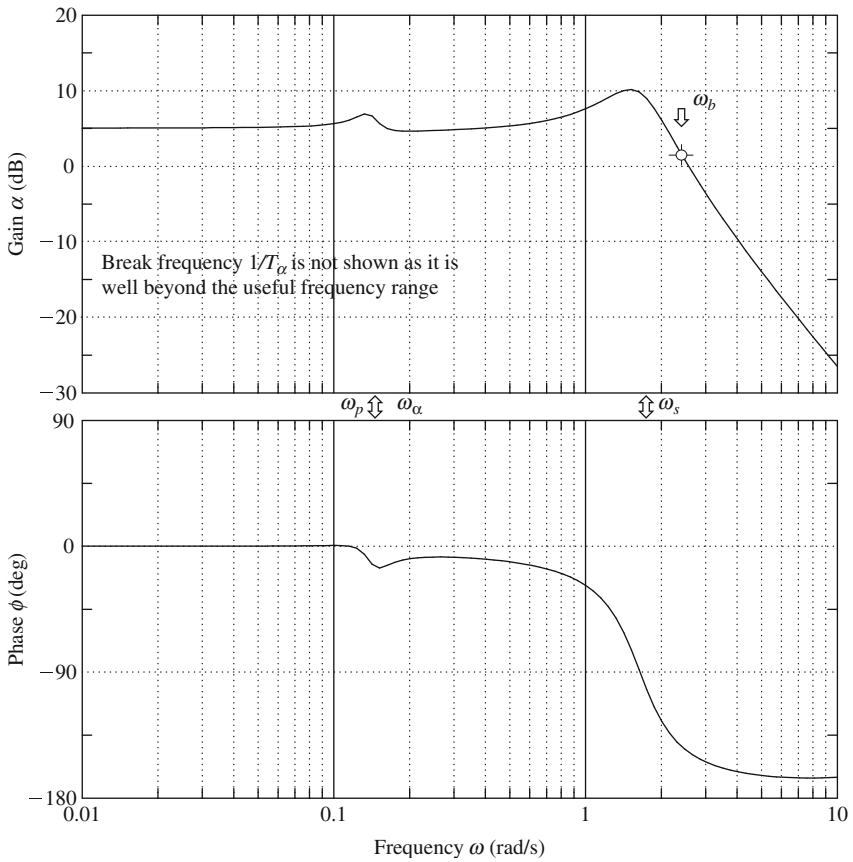


Figure 4.4-9 A-7A body incidence frequency response.

the aircraft model is linear, this very large gain ratio may be interpreted equivalently as approximately 98 ft/s/deg, which is much easier to appreciate physically. Since the gain drops away rapidly as the frequency increases beyond the phugoid frequency, the velocity bandwidth frequency is only a little higher than the phugoid frequency. This accords well with practical observation; velocity perturbations at frequencies in the vicinity of the short period mode are usually insignificantly small. The phase plot indicates that there is no appreciable phase shift between input and output until the frequency exceeds the phugoid frequency when there is a rapid increase in phase lag. This means that for all practical purposes speed changes demanded by the pilot will follow the stick in the usable frequency band.

The pitch attitude θ frequency response to elevator input η is shown in Fig. 4.4-8. Its general interpretation follows the discussion of Fig. 4.4-6 and is not repeated here. However, there are some significant differences which must not be overlooked. The differences are due to the fact that the transfer function is non-minimum phase, a consequence of selecting a very low speed flight condition for the example. Referring to equations (4.4.50), this means that the numerator zero $1/T_{\theta_1}$ is negative, and the reasons for this are discussed in Example 4.4.1. The non-minimum phase effects do not influence the gain plot in any significant way, so its interpretation is quite straightforward. However, the effect of the non-minimum phase numerator zero is to introduce phase lag at very low frequencies rather than the usual phase lead. It is likely that in manoeuvring at this flight condition the pilot would be aware of the pitch attitude lag in response to his stick input.

The body incidence α frequency response to elevator input η is shown in Fig. 4.4-9 and it is clear that, as might be expected, this is dominated by the short period mode. For all practical purposes the influence of the phugoid on both the gain and phase frequency responses is insignificant. This may be confirmed by reference to the appropriate transfer function in equations (4.4.50), where it will be seen that the second order numerator term very nearly cancels the phugoid term in the denominator. This is an important observation since it is quite usual to cancel approximately equal numerator and denominator terms in any response transfer function to simplify it. Simplified transfer functions often provide adequate response models in both the time and frequency domains, and can be extremely useful for explaining and interpreting aircraft dynamic behaviour. In modern control parlance the phugoid dynamics would be said to be *not observable* in this illustration. The frequency response in both gain and phase is more or less *flat* at frequencies up to the short period frequency, or for most of the usable frequency range. In practical terms this means that incidence will follow the stick at constant gain and without appreciable phase lag, which is obviously a desirable state of affairs.

4.4.5 Flying and Handling Qualities

The longitudinal stability modes play an absolutely fundamental part in determining the longitudinal flying and handling qualities of an aircraft and it is essential that their characteristics must be "correct" if the aircraft is to be flown by a human pilot. A simplistic view of the human pilot suggests that he behaves like an adaptive dynamic system and will adapt his dynamics to harmonise with those of the controlled vehicle. Since his dynamics interact and couple with those of the aircraft, he will adapt, within human limits, to produce the best *closed loop* system dynamics compatible with the piloting task. His adaptability enables him to cope well with aircraft with less than desirable flying qualities. However, the problems of coupling between incompatible dynamic systems can be disastrous and it is this latter aspect of the piloting task that has attracted much attention in recent years. Every time the aircraft is disturbed in response to control commands the stability modes are excited and it is not difficult to appreciate why their characteristics are so important. Similarly, the stability modes are equally important in determining *ride quality* when the main concern is response to atmospheric disturbances. In military combat aircraft ride quality determines the effectiveness of the airframe as a weapons platform and in the civil transport aircraft it determines the comfort of passengers.

In general it is essential that the short period mode, which has a natural frequency close to human pilot natural frequency, is adequately damped. Otherwise, dynamic coupling with the pilot may occur under certain conditions leading to severe, or even catastrophic, handling problems. On the other hand, as the phugoid mode is much lower in frequency its impact on the piloting task is much less demanding. The average human pilot can easily control the aircraft even when the phugoid is mildly unstable. The phugoid mode can, typically, manifest itself as a minor trimming problem when poorly damped. Although not in itself hazardous, it can lead to increased pilot workload and for this reason it is desirable to ensure adequate phugoid damping. It is also important that the natural frequencies of the stability modes should be well separated in order to avoid interaction, or coupling, between the modes. Mode coupling may give rise to unusual handling characteristics and is generally regarded as an undesirable feature in longitudinal dynamics.

4.4.6 Mode Excitation

Since the longitudinal stability modes are usually well separated in frequency, it is possible to excite the modes more or less independently for the purposes of demonstration or measurement. Indeed, it is a general flying qualities requirement that the modes be well separated in frequency in order to avoid handling problems arising from dynamic mode coupling. The modes may be excited selectively by the application of a sympathetic elevator input to the trimmed aircraft. The methods developed for in-flight mode excitation reflect an intimate understanding of the dynamics involved and are generally easily adapted to the analytical environment. Because the longitudinal modes are usually well separated in frequency the form of the input disturbance is not, in practice, very critical. However, some consistency in the flight test or analytical procedures adopted is desirable if meaningful comparative studies are to be made.

The short period pitching oscillation may be excited by applying a short duration disturbance in pitch to the trimmed aircraft. This is best achieved with an elevator pulse having a duration of a second or less. Analytically this is adequately approximated by a unit impulse applied to the elevator. The essential feature of the disturbance is that it must be sufficiently short so as not to excite the phugoid significantly. However, as the phugoid damping is usually very low it is almost impossible not to excite the phugoid at the same time but, it does not usually develop fast enough to obscure observation of the short period mode. An example of a short period response recorded during a flight test exercise in a Handley Page Jetstream aircraft is shown in Fig. 4.4-10. In fact two excitations are shown, the first in the nose up sense and the second in the nose down sense. The pilot input "impulse" is clearly visible and represents his best attempt at achieving a clean impulse like input; some practice is required before consistently good results are obtained. Immediately following the input the pilot released the controls to obtain the controls free dynamic response which explains why the elevator angle does not recover its equilibrium trim value until the short period transient has settled. During this short elevator free period its motion is driven by oscillatory aerodynamic

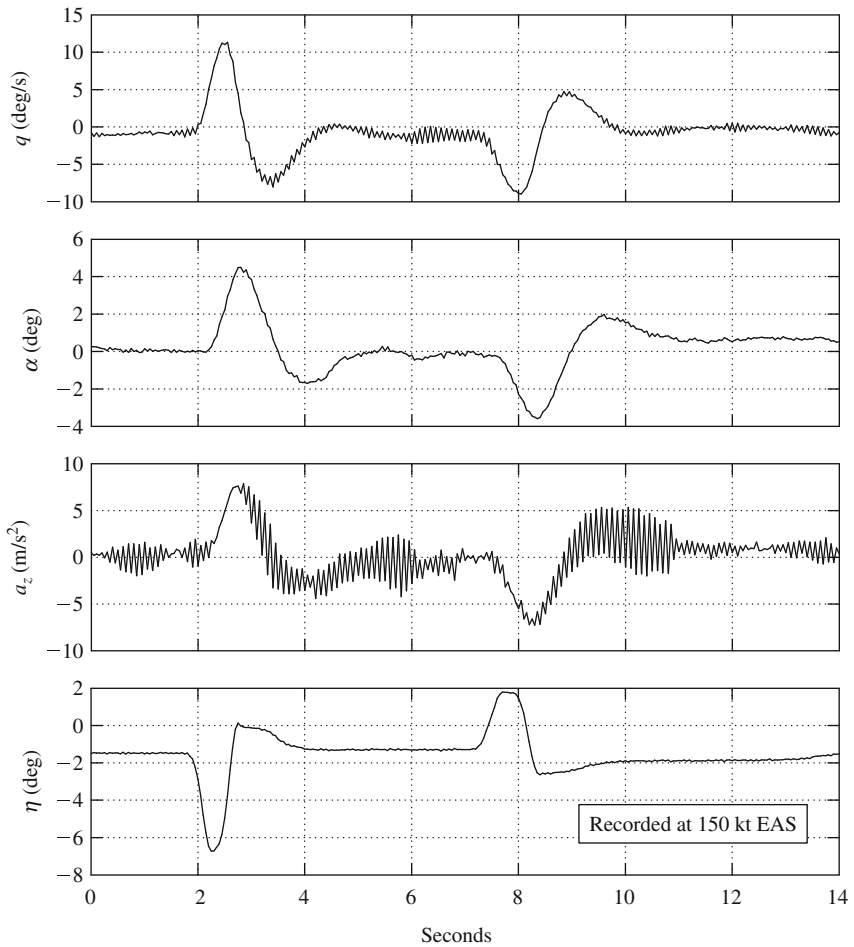


Figure 4.4-10 Flight recording of the short period pitching oscillation.

loading and is also coloured by the control circuit dynamics which can be noticeably intrusive. Otherwise the response is typical of a well damped aeroplane.

The phugoid mode may be excited by applying a small speed disturbance to the aircraft in trimmed flight. This is best achieved by applying a small step input to the elevator which will cause the aircraft to fly up, or down, according to the sign of the input. If the power is left at its trimmed setting then the speed will decrease, or increase, accordingly. When the speed has diverged from its steady trimmed value by about 5% or so, the elevator is returned to its trim setting. This provides the disturbance and a stable aircraft will then execute a phugoid oscillation as it recovers its trim equilibrium. Analytically, the input is equivalent to an elevator pulse of several seconds duration. The magnitude and length of the pulse would normally be established by trial and error since its effect will be very aircraft dependent. However, it should be remembered that for proper interpretation of the resulting response the disturbance should be small in magnitude since a small perturbation model is implied. An example of a phugoid response recorded during a flight test exercise in a Handley Page Jetstream aircraft is shown in Fig. 4.4-11. The pilot input "pulse" is clearly visible and, as for the short period mode, some practice is required before consistently good results are obtained. Again, the controls are released following the input to obtain the controls free dynamic response and the subsequent elevator motion is caused by the sinusoidal aerodynamic loading on the surface itself. The leading and trailing edge steps of the input elevator pulse may excite the short period mode. However, the short period mode transient would normally decay to zero well before the phugoid has properly developed and would not therefore obscure the observation of interest.

It is clear from an inspection of Fig. 4.4-11 that the phugoid damping is significantly higher than might be expected from the previous discussion of the mode characteristics. What is in fact shown is the aerodynamic, or basic airframe, phugoid modified by the inseparable effects of power. The Astazou engines of the Jetstream are governed to run at constant rpm and thrust changes are achieved by varying the propeller blade pitch. Thus as the aircraft flies the sinusoidal flight path during a phugoid disturbance the sinusoidal propeller loading causes the engine to

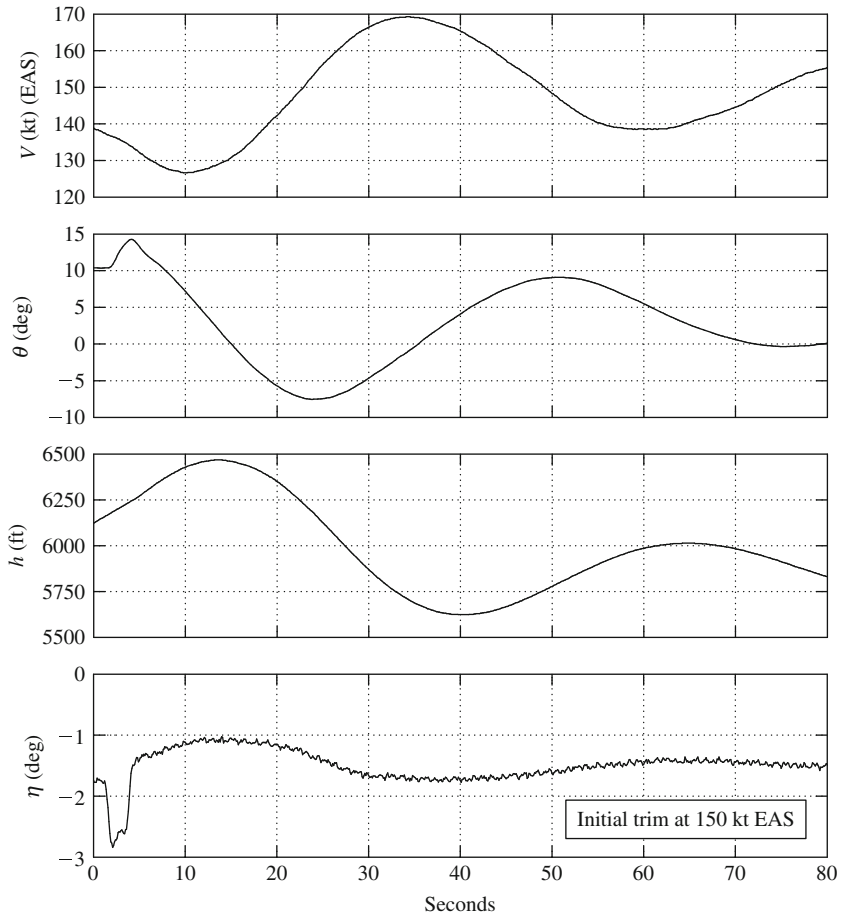


Figure 4.4-11 Flight recording of the phugoid.

automatically adjust its power to maintain constant propeller rpm. This very effectively increases the apparent damping of the phugoid. It is possible to operate the aircraft at a constant power condition when the "power damping" effect is suppressed. Under these circumstances it is found that the aerodynamic phugoid is much less stable, as predicted by the simple theoretical model, and at some flight conditions it is unstable.

The above flight recording of the longitudinal stability modes illustrates the *controls free* dynamic stability characteristics. The same exercise could of course be repeated with the controls held fixed following the disturbing input. In this event the *controls fixed* dynamic stability characteristics would be observed. In general the differences between the responses would be small and not too significant. Now controls free dynamic response is only possible in aeroplanes with reversible controls which includes most small classical aeroplanes. Virtually all larger modern aircraft have powered controls, driven by electronic flight control systems, which are effectively irreversible and which means that they are only capable of exhibiting controls fixed dynamic response. Thus, today, most theoretical modelling and analysis is concerned with controls fixed dynamics only, as is the case throughout this book. However, a discussion of the differences between controls fixed and controls free aeroplane dynamics may be found in Hancock (1995).

When it is required to analyse the dynamics of a single mode in isolation, the best approach is to emulate flight test practice as far as that is possible. It is necessary to choose the most appropriate transfer functions to show the dominant response variables in the mode of interest. For example, as shown in Figs. 4.4-10 and 4.4-11 the short period mode is best observed in the dominant response variables q and $w(\alpha)$ whereas the phugoid is best observed in its dominant response variables u, h and θ. It is necessary to apply a control input disturbance sympathetic to the mode dynamics and it is necessary to observe the response for an appropriate period of time. For example, Fig. 4.4-1 shows both longitudinal modes but the time scale of the illustration reveals the phugoid in much greater detail than the short period mode, whereas the time scale of Fig. 4.4-5 was chosen to reveal the short period mode in detail since that is the mode of interest. The form of the control input is not usually difficult to arrange in analytical work since most software packages have built-in impulse, step and pulse functions, whilst more esoteric functions can usually be programmed by the user. This kind of informed approach to the analysis is required if the best possible visualisation of the longitudinal modes and their associated dynamics is to be obtained.

References

Friedland, B. 1987. *Control System Design*. McGraw-Hill Book Company, New York.
Hancock, G.J. 1995: *An Introduction to the Flight Dynamics of Rigid Aeroplanes*. Ellis Horwood Ltd., Hemel Hempstead.
Lanchester, F.W. 1908: *Aerodonetics*. Macmillan and Co.
Shinners, S.M. 1980: *Modern Control System Theory and Application*. Addison-Wesley Publishing Co, Reading, Massachusetts.
Teper, G.L. 1969: *Aircraft Stability and Control Data*. Systems Technology, Inc., STI Technical Report 176-1. NASA Contractor Report, National Aeronautics and Space Administration, Washington D.C. 20546.

Problems

1. A tailless aeroplane of 9072 kg mass has an aspect ratio 1 delta wing of area 37 m^2. The longitudinal short period motion of the aeroplane is described by the characteristic quadratic:

$$\lambda^2 + B\lambda + C = 0 \quad \text{where } B = \frac{1}{2}\left(\frac{dC_L}{d\alpha}\right)\cos^2\alpha \quad \text{and}$$

$$C = -\frac{1}{2}\left(\frac{\mu_1}{i_y}\right)\left(\frac{dC_m}{d\alpha}\right)\cos\alpha.$$

α is the wing incidence, $\mu_1 = m/\frac{1}{2}\rho S\bar{c}$ is the longitudinal relative density parameter, and $i_y = I_y/m\bar{c}^2$ is the dimensionless moment of inertia in pitch. The aeroplane's moment of inertia in pitch is 1.356×10^5 kg/m^2. The variation of C_L and C_m with incidence $\alpha > 0$ is non-linear for the aspect ratio 1 delta wing:

$$C_L = \frac{1}{2}\pi\alpha + 2\alpha^2$$

$$C_m = C_{m_0} - 0.025\pi\alpha - 0.1\alpha^2$$

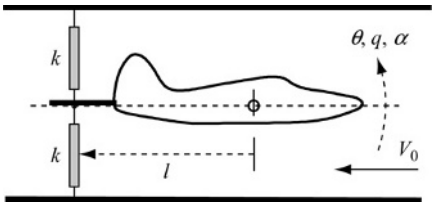
Compare and describe the short period motions when the aeroplane is flying straight and level at 152 m/s at sea level and at 35,000 ft.
$\rho_0 = 1.225$ kg/m^3 at sea level, $\rho/\rho_0 = 0.310$ at 35,000 ft. Characteristic time $\sigma = m/\frac{1}{2}\rho V_0 S$. (CU 1983)

2. (i) List the characteristics of the longitudinal phugoid stability mode.
   (ii) List the characteristics of the longitudinal short period pitching stability mode.
   (iii) The transfer function for the unaugmented McDonnell F-4C Phantom describing the pitch attitude response to elevator when flying at Mach 1.2 at an altitude of 35,000 ft is given by

$$\frac{\theta(s)}{\eta(s)} = \frac{-20.6(s+0.013)(s+0.62)}{(s^2+0.017s+0.002)(s^2+1.74s+29.49)} \text{ rad/rad}$$

   Write down the longitudinal characteristic equation and state whether the aeroplane is stable, or not.
   (iv) What are the numerical parameters describing the longitudinal stability modes of the McDonnell F-4C Phantom? (CU 1999)
3. Describe the longitudinal short period pitching oscillation. On what parameters do its characteristics depend?

A model aircraft is mounted in a wind tunnel such that it is free to pitch about an axis through its *cg* as shown. The model is restrained by two springs attached at a point on a fuselage aft extension which is at a distance $l = 0.5$ m from the *cg*. The model has wing span $b = 0.8$ m, mean aerodynamic chord $\bar{\bar{c}} = 0.15$ m and the air density may be taken as $\rho = 1.225$ kg/m^3.



With the wind off the model is displaced in pitch and released. The frequency of the resulting oscillation is 10 rad/s and the damping ratio 0.1. The experiment is repeated with a wind velocity $V_0 = 30$ m/s, the frequency is now found to be 12 rad/s and the damping ratio 0.3. Given that the spring stiffness $k = 16$ N/m, calculate the moment of inertia in pitch, and values for the dimensionless stability derivatives M_q and $M_{\dot{w}}$. It may be assumed that the influence of the derivative $M_{\dot{w}}$ is negligible. State all assumptions made. (CU 1987)

4. (i) Show that the period of the phugoid is given approximately by, $T_p = \sqrt{2}\pi\frac{V_0}{g}$, and state all assumptions used during the derivation.
   (ii) State which aerodynamic parameters introduce damping into a phugoid, and discuss how varying forward speed whilst on approach to landing may influence phugoid characteristics. (LU 2001)
5. (i) Using a simple physical model, show that the short period pitching oscillation can be approximated to by

$$I_y\ddot{\theta} + \frac{1}{2}\rho V_0 a_1 S_T l_T^2 \dot{\theta} - \frac{1}{2}\rho V_0^2 S\bar{\bar{c}}\frac{dC_m}{d\alpha}\theta = 0$$

   (ii) The aircraft described below is flying at sea level at 90 m/s. Determine the *cg* location at which the short period pitching oscillation ceases to be oscillatory:

Wing lift curve slope	= 5.7 1/rad
Tailplane lift curve slope	= 3.7 1/rad
Horizontal tail arm	= 6 m
Tailplane area	= 5 m^2
$d\varepsilon/d\alpha$	= 0.30
I_y	= 40,000 kg/m^2
Wing area	= 30 m^2
Mean aerodynamic chord	= 1.8 m
Aerodynamic centre	= 0.18 $\bar{\bar{c}}$

   (Hint: Modify the equation in part (i) to include tailplane lag effects.)
   (iii) Determine the period of the short period pitching oscillation if the *cg* location is moved 0.2 $\bar{\bar{c}}$ forward of the position calculated in part (ii). (LU 2001)

6. For a conventional aircraft on an approach to landing, discuss how the aircraft's aerodynamics may influence longitudinal stability. (LU 2002)
7. Determine the time to half amplitude and the period of the short period pitching oscillation. Assume that the short period pitching oscillation can be approximated by

$$I_y\ddot{\theta} - \frac{\partial M}{\partial q}\dot{\theta} - \frac{\partial M}{\partial w}V\theta = 0 \quad \text{and in addition } M_w = \partial C_m/\partial\alpha. \qquad \text{(LU 2003)}$$

Aerospace Engineering Desk Reference; ISBN 9781856175753

4.5 Lateral–Directional Dynamics

Michael Cook

4.5.1 Response to Controls

The procedures for investigating and interpreting the lateral–directional dynamics of an aeroplane are much the same as those used to deal with the longitudinal dynamics and are not repeated at the same level of detail in this chapter. However, some aspects of lateral–directional dynamics, and their interpretation, differ significantly from the longitudinal dynamics and the procedures for interpreting the differences are dealt with appropriately. The lateral–directional response transfer functions are obtained in the solution of the lateral–directional equations of motion. The transfer functions completely describe the linear dynamic asymmetric response in sideslip, roll and yaw to aileron and rudder inputs. As in the longitudinal solution, implicit in the response are the dynamic properties determined by the lateral–directional stability characteristics of the aeroplane. As before, the transfer functions and the response variables described by them are linear since the entire modelling process is based on the assumption that the motion is constrained to small disturbances about an equilibrium trim state. The equilibrium trim state is assumed to mean steady level flight in the first instance and the previously stated caution concerning the magnitude of a small lateral–directional perturbation applies.

The most obvious difference between the solution of the longitudinal equations of motion and the lateral–directional equations of motion is that there is more algebra to deal with. Since two aerodynamic inputs are involved, the ailerons and the rudder, two sets of input–output response transfer functions are produced in the solution of the equations of motion. However, these are no more difficult to deal with than a single input–output set of transfer functions, there are just more of them! The most significant difference between the longitudinal and lateral–directional dynamics of the aeroplane concerns the interpretation. In general the lateral–directional stability modes are not so distinct and tend to exhibit dynamic coupling to a greater extent. Thus some care is needed in the choice of assumptions made to facilitate their interpretation. A mitigating observation is that, unlike the longitudinal dynamics, the lateral–directional dynamics do not change very much with flight condition since most aeroplanes possess aerodynamic symmetry by design.

The lateral–directional equations of motion describing small perturbations about an equilibrium trim condition and referred to wind axes are given by the state equation as follows:

$$\begin{bmatrix} \dot{v} \\ \dot{p} \\ \dot{r} \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} y_v & y_p & y_r & y_\phi \\ l_v & l_p & l_r & l_\phi \\ n_v & n_p & n_r & n_\phi \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \\ r \\ \phi \end{bmatrix} + \begin{bmatrix} y_\xi & y_\zeta \\ l_\xi & l_\zeta \\ n_\xi & n_\zeta \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \zeta \end{bmatrix} \tag{4.5.1}$$

The solution of equation (4.5.1) produces two sets of four response transfer functions, one set describing motion in response to aileron input and a second set describing response to rudder input. As for the longitudinal response transfer functions, it is convenient to adopt a shorthand style of writing the transfer functions. The transfer functions describing response to aileron are conveniently written as

$$\frac{v(s)}{\xi(s)} \equiv \frac{N_\xi^v(s)}{\Delta(s)} = \frac{k_v(s + (1/T_{\beta_1}))(s + (1/T_{\beta_2}))}{(s + (1/T_s))(s + (1/T_r))(s^2 + 2\zeta_d\omega_d s + \omega_d^2)} \tag{4.5.2}$$

$$\frac{p(s)}{\xi(s)} \equiv \frac{N_\xi^p(s)}{\Delta(s)} = \frac{k_p s(s^2 + 2\zeta_\phi\omega_\phi s + \omega_\phi^2)}{(s + (1/T_s))(s + (1/T_r))(s^2 + 2\zeta_d\omega_d s + \omega_d^2)} \tag{4.5.3}$$

$$\frac{r(s)}{\xi(s)} \equiv \frac{N_\xi^r(s)}{\Delta(s)} = \frac{k_r(s + (1/T_\psi))(s^2 + 2\zeta_\psi\omega_\psi s + \omega_\psi^2)}{(s + (1/T_s))(s + (1/T_r))(s^2 + 2\zeta_d\omega_d s + \omega_d^2)} \tag{4.5.4}$$

$$\frac{\phi(s)}{\xi(s)} \equiv \frac{N_\xi^\phi(s)}{\Delta(s)} = \frac{k_\phi(s^2 + 2\zeta_\phi\omega_\phi s + \omega_\phi^2)}{(s + (1/T_s))(s + (1/T_r))(s^2 + 2\zeta_d\omega_d s + \omega_d^2)} \tag{4.5.5}$$

and the transfer functions describing response to rudder are conveniently written as

$$\frac{v(s)}{\zeta(s)} \equiv \frac{N_\zeta^v(s)}{\Delta(s)} = \frac{k_v(s+(1/T_{\beta_1}))(s+(1/T_{\beta_2}))(s+(1/T_{\beta_3}))}{(s+(1/T_s))(s+(1/T_r))(s^2+2\zeta_d\omega_d s+\omega_d^2)} \tag{4.5.6}$$

$$\frac{p(s)}{\zeta(s)} \equiv \frac{N_\zeta^p(s)}{\Delta(s)} = \frac{k_p s(s+(1/T_{\phi_1}))(s+(1/T_{\phi_2}))}{(s+(1/T_s))(s+(1/T_r))(s^2+2\zeta_d\omega_d s+\omega_d^2)} \tag{4.5.7}$$

$$\frac{r(s)}{\zeta(s)} \equiv \frac{N_\zeta^r(s)}{\Delta(s)} = \frac{k_r(s+(1/T_\psi))(s^2+2\zeta_\psi\omega_\psi s+\omega_\psi^2)}{(s+(1/T_s))(s+(1/T_r))(s^2+2\zeta_d\omega_d s+\omega_d^2)} \tag{4.5.8}$$

$$\frac{\phi(s)}{\zeta(s)} \equiv \frac{N_\zeta^\phi(s)}{\Delta(s)} = \frac{k_\phi(s+(1/T_{\phi_1}))(s+1/T_{\phi_2})}{(s+(1/T_s))(s+(1/T_r))(s^2+2\zeta_d\omega_d s+\omega_d^2)} \tag{4.5.9}$$

The solution of the equations of motion results in polynomial descriptions of the transfer function numerators and common denominator. The polynomials factorise into real and complex pairs of roots that are most explicitly quoted in the style of equations (4.5.2)–(4.5.9) above. Since the roots are interpreted as time constants, damping ratios and natural frequencies, the above style of writing makes the essential information instantly available. It should also be noted that the numerator and denominator factors are typical for a conventional aeroplane. Sometimes complex pairs of roots may be replaced with two real roots and *vice versa*. However, this does not usually mean that the dynamic response characteristics of the aeroplane become dramatically different. Differences in the interpretation of response may be evident but will not be necessarily large.

Transfer functions (4.5.2)–(4.5.9) each describe uniquely different, but related, variables in the motion of the aeroplane in response to a control input. However, it will be observed that the notation adopted indicates similar values for some numerator terms in both aileron and rudder response transfer functions, for example, k_r, T_ψ, ζ_ψ and ω_ψ, appear in both $N_\xi^r(s)$ and $N_\zeta^r(s)$. *It must be understood that the numerator parameters are context dependent and usually have a numerical value which is unique to the transfer function in question.* To repeat the comment made above, the notation is a convenience for allocating particular numerator terms and serves only to identify the role of each term as a gain, time constant, damping ratio or frequency.

As before, the denominator of the transfer functions describes the characteristic polynomial which, in turn, describes the lateral–directional stability characteristics of the aeroplane. The transfer function denominator is therefore common to all response transfer functions. Thus the response of all variables to an aileron or to a rudder input is dominated by the denominator parameters namely, time constants, damping ratio and natural frequency. The differences between the individual responses are entirely determined by their respective numerators and the *response shapes* of the individual variables are determined by the common denominator and "coloured" by their respective numerators.

Example 4.5.1

The equations of motion and aerodynamic data for the Douglas DC-8 aircraft were obtained from Teper (1969). At the flight condition of interest the aircraft has a total weight of 190,000 lb and is flying at Mach 0.44 at an altitude of 15,000 ft. The source data are referenced to aircraft body axes and for the purposes of this illustration it has been converted to wind axes reference. The equations of motion, referred to wind axes and quoted in terms of concise derivatives, are in state space format

$$\begin{bmatrix}\dot v\\ \dot p\\ \dot r\\ \dot\phi\end{bmatrix} = \begin{bmatrix}-0.1008 & 0 & -468.2 & 32.2\\ -0.00579 & -1.232 & 0.397 & 0\\ 0.00278 & -0.0346 & -0.257 & 0\\ 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}v\\ p\\ r\\ \phi\end{bmatrix} + \begin{bmatrix}0 & 13.48416\\ -1.62 & 0.392\\ -0.01875 & -0.864\\ 0 & 0\end{bmatrix}\begin{bmatrix}\xi\\ \zeta\end{bmatrix} \tag{4.5.10}$$

Since it is useful to have the transfer function describing sideslip angle β as well as sideslip velocity v, the output equation is augmented. Thus the output equation is

$$\begin{bmatrix} v \\ p \\ r \\ \phi \\ \beta \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0.00214 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \\ r \\ \phi \end{bmatrix} \tag{4.5.11}$$

Again, the numerical values of the matrix elements in equations (4.5.10) and (4.5.11) have been rounded to five decimal places in order to keep the equations to a reasonable written size. This should not be done with the equations used in the actual computation.

Solution of the equations of motion using *Program CC* produced the following two sets of transfer functions: First, the transfer functions describing response to aileron

$$\begin{aligned}
\frac{v(s)}{\xi(s)} &= \frac{8.779(s+0.197)(s-7.896)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ ft/s/rad} \\
\frac{p(s)}{\xi(s)} &= \frac{-1.62s(s^2+0.362s+1.359)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ rad/s/rad (deg/s/deg)} \\
\frac{r(s)}{\xi(s)} &= \frac{-0.0188(s+1.59)(s^2-3.246s+4.982)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ rad/s/rad (deg/s/deg)} \\
\frac{\phi(s)}{\xi(s)} &= \frac{-1.62(s^2+0.362s+1.359)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ rad/rad (deg/deg)} \\
\frac{\beta(s)}{\xi(s)} &= \frac{0.0188(s+0.197)(s-7.896)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ rad/rad (deg/deg)}
\end{aligned} \tag{4.5.12}$$

and second, the transfer functions describing response to rudder

$$\begin{aligned}
\frac{v(s)}{\zeta(s)} &= \frac{13.484(s-0.0148)(s+1.297)(s+30.207)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ ft/s/rad} \\
\frac{p(s)}{\zeta(s)} &= \frac{0.392s(s+1.85)(s-2.566)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ rad/s/rad (deg/s/deg)} \\
\frac{r(s)}{\zeta(s)} &= \frac{-0.864(s+1.335)(s^2-0.03s+0.109)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ rad/s/rad (deg/s/deg)} \\
\frac{\phi(s)}{\zeta(s)} &= \frac{0.392(s+1.85)(s-2.566)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ rad/rad (deg/deg)} \\
\frac{\beta(s)}{\zeta(s)} &= \frac{0.029(s-0.0148)(s+1.297)(s+30.207)}{(s+0.0065)(s+1.329)(s^2+0.254s+1.433)} \text{ rad/rad (deg/deg)}
\end{aligned} \tag{4.5.13}$$

The characteristic equation is given by equating the denominator to zero

$$\Delta(s) = (s+0.0065)(s+1.329)(s^2+0.254s+1.433) = 0 \tag{4.5.14}$$

The first real root describes the *spiral mode* with time constant

$$T_s = \frac{1}{0.0065} \cong 154 \text{ s}$$

the second real root describes the *roll subsidence mode* with time constant

$$T_r = \frac{1}{1.329} = 0.75 \text{ s}$$

and the pair of complex roots describe the oscillatory *dutch roll mode* with characteristics

$$\begin{aligned}
\text{Damping ratio } \zeta_d &= 0.11 \\
\text{Undamped natural frequency } \omega_d &= 1.2 \text{ rad/s}
\end{aligned}$$

Since both real roots are negative and the pair of complex roots have negative real parts then the mode characteristics indicate the airframe to be aerodynamically stable.

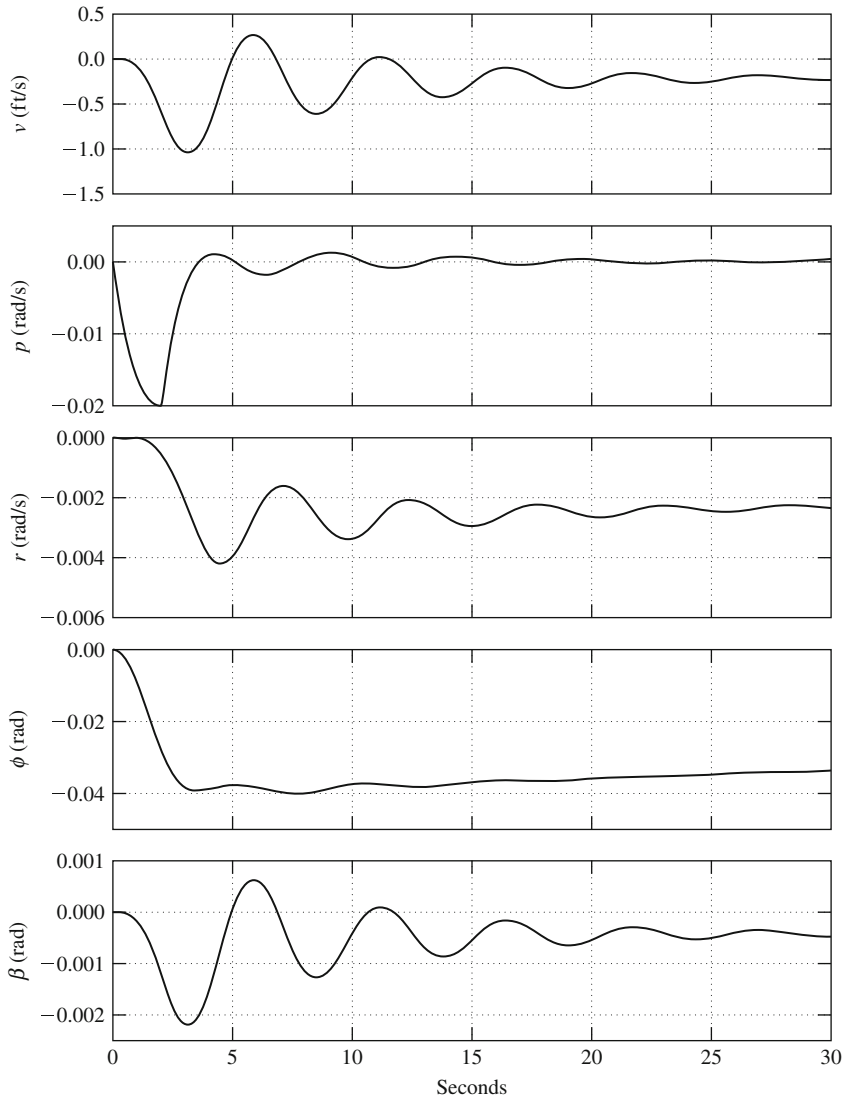


Figure 4.5-1 Aircraft response to 1° 2 s aileron pulse input.

The response of the aeroplane to a unit (1°) aileron pulse, held on for 2 s and then returned to zero, is shown in Fig. 4.5-1. All of the variables obtained in the solution of the equations of motion are shown, the individual responses being characterised by the transfer functions, equations (4.5.12).

The dynamics associated with the three stability modes are visible in the responses although, at first glance, they would appear to be dominated by the oscillatory dutch roll mode since its damping is relatively low. Since the non-oscillatory spiral and roll modes are not so distinct, and since the dynamic coupling between modes is significant it is rather more difficult to expose the modes analytically unless some care is taken in their graphical presentation. This subject is discussed in greater detail in Section 4.5.6. Both the roll and spiral modes appear as exponentially convergent characteristics since they are both stable in this example. The roll mode converges relatively quickly with a time constant of 0.75 s, whereas the spiral mode converges very slowly indeed with a time constant of 154 s. The roll mode is most clearly seen in the roll rate response p where it determines the exponential rise at zero seconds and the exponential recovery when the pulse is removed at 2 s. The spiral mode characteristic is rather more subtle and is most easily seen in the roll attitude response ϕ where it determines the longer term convergence to zero and is fully established at 30 s. Once again, all of the response shapes are determined by the common stability mode dynamics and the obvious differences between them are due to the unique numerators in each transfer function. All of the response variables shown in Fig. 4.5-1 eventually decay to zero in the time scale of the spiral mode (about 200 s) since the aircraft is stable.

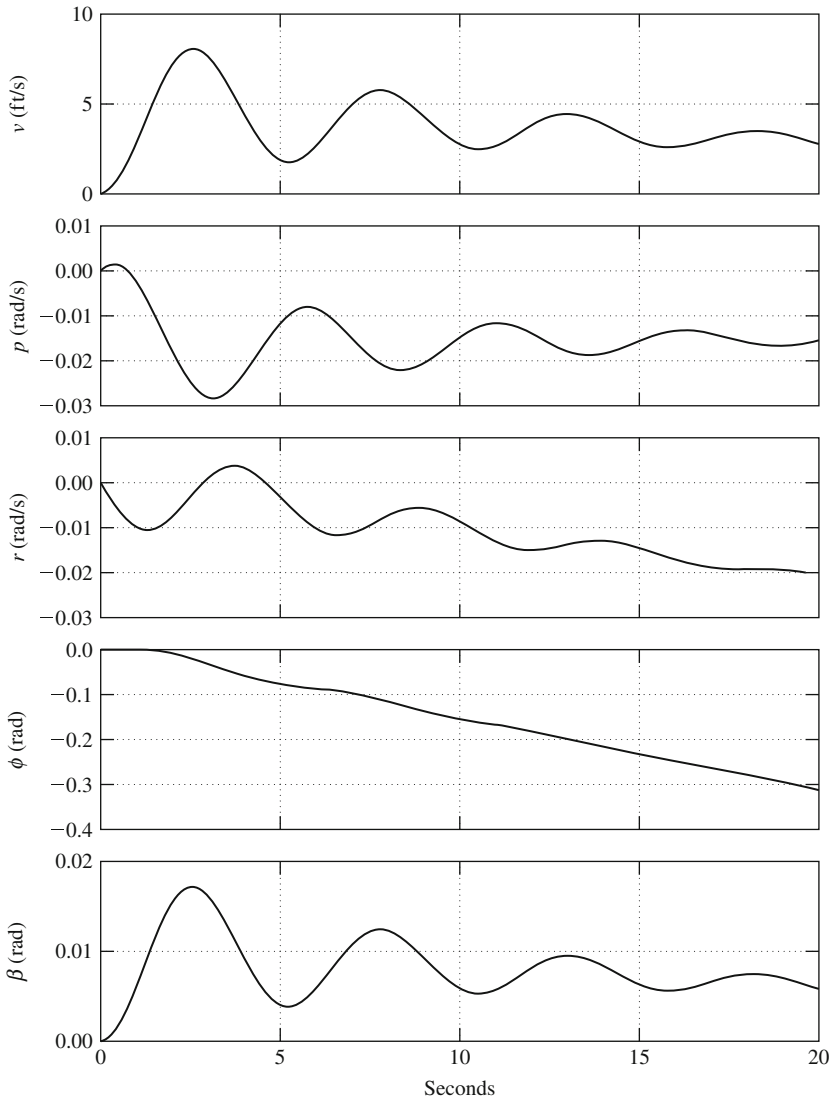


Figure 4.5-2 Aircraft response to 1° rudder step input.

The response of the aeroplane to a unit (1°) rudder step input is shown in Fig. 4.5-2. All of the variables obtained in the solution of the equations of motion are shown, the individual responses being characterised by the transfer functions, equations (4.5.13).

Again, it is very clear that the response is dominated by the oscillatory dutch roll mode. However, unlike the previous illustration, the roll and spiral modes are not discernible in the response. This is due to the fact that a step was chosen as the input which simply causes the aircraft to diverge from its initial equilibrium. This motion, together with the dutch roll oscillation, effectively masks the two non-oscillatory modes. Now it is possible to observe another interesting phenomenon in the response. Inspection of the transfer functions, equations (4.5.12) and (4.5.13), reveals that a number possess non-minimum phase numerator terms. The effect of these non-minimum phase terms would seem to be insignificantly small since they are not detectable in the responses shown in Figs 4.5-1 and 4.5-2, with one exception. The roll rate response p to rudder, shown in Fig. 4.5-2, exhibits a sign reversal for the first second or so of its response and this is the manifestation of the non-minimum phase effect. In aeronautical parlance it is referred to as *adverse roll* in response to rudder.

A positive rudder step input is assumed and this will cause the aircraft to turn to the left, which is a negative response in accordance with the notation. Once the turn is established this results in negative yaw and yaw rate together with negative roll and roll rate induced by yaw-roll coupling. These general effects are correctly portrayed in the responses shown in Fig. 4.5-2. However, when the rudder is deflected initially a substantial side force is generated

at the centre of pressure of the fin which in turn generates the yawing moment causing the aircraft to turn. However, the side force acts at some distance above the roll axis and also generates a rolling moment which causes the aircraft to roll in the opposite sense to that induced by the yawing motion. Since inertia in roll is somewhat lower than inertia in yaw the aircraft responds quicker in roll and starts to roll in the "wrong" direction, but as the yawing motion becomes established the aerodynamically induced rolling moment eventually overcomes the adverse rolling moment and the aircraft then rolls in the "correct" sense. This behaviour is clearly visible in Fig. 4.5-2 and is a characteristic found in most aircraft. The magnitude of the effect is aircraft dependent and if not carefully controlled by design can lead to unpleasant handling characteristics. A similar characteristic, *adverse yaw*, in response to aileron is caused by the differential drag effects associated with aileron deflection giving rise to an adverse yawing moment. This characteristic is also commonly observed in many aircraft; reference to equations (4.5.12) indicates that it is present in the DC-8 but is insignificantly small at the chosen flight condition.

The mode content in each of the motion variables is given most precisely by the eigenvectors. With the aid of *MATLAB* the eigenvector matrix $\mathbf{V}$ was obtained from the state matrix in equation (4.5.10):

$$\mathbf{V} = \begin{array}{c} \begin{matrix} \text{Dutch roll mode} & & \text{Roll mode} & \text{Spiral mode} \end{matrix} \\ \left[\begin{array}{cc|c|c} -0.845+0.5291j & -0.845-0.5291j & -0.9970 & 0.9864 \\ 0.0012-0.0033j & 0.0012+0.0033j & -0.0619 & -0.0011 \\ 0.0011+0.0021j & 0.0011-0.0021j & 0.0006 & 0.0111 \\ -0.0029-0.0007j & -0.0029+0.0007j & 0.0466 & 0.1641 \end{array}\right] \end{array} \begin{matrix} \\ :v \\ :p \\ :r \\ :\phi \end{matrix} \tag{4.5.15}$$

To facilitate interpretation of the eigenvector matrix, the magnitude of each component eigenvector is calculated as follows:

$$|\mathrm{V}| = \left[\begin{array}{cc|c|c} 0.9970 & 0.9970 & 0.9970 & 0.9864 \\ 0.0035 & 0.0035 & 0.0619 & 0.0011 \\ 0.0024 & 0.0024 & 0.0006 & 0.0111 \\ 0.0030 & 0.0030 & 0.0466 & 0.1641 \end{array}\right] \begin{matrix} :v \\ :p \\ :r \\ :\phi \end{matrix}$$

Clearly, the content of all three modes in sideslip velocity v, and hence in β, is of similar order, the roll mode is dominant in roll rate p and the spiral mode is dominant in roll attitude response ϕ. These observations correlate well with the responses shown in Figs 4.5-1 and 4.5-2 although the low dutch roll damping obscures the observation in some response variables. Although not the case in this example, eigenvector analysis can be particularly useful for interpreting lateral–directional response in aircraft where mode coupling is rather more pronounced and the modes are not so distinct.

The steady state values of the motion variables following a unit step (1°) aileron or rudder input may be determined by the application of the final value theorem, equation (4.5.33), to the transfer functions, equations (4.5.12) and (4.5.13). The calculation procedure is shown in Example 4.4.1 and is not repeated here. Thus the steady state response of all the motion variables to an aileron unit step input is

$$\begin{bmatrix} v \\ p \\ r \\ \phi \\ \beta \end{bmatrix}_{\substack{steady \\ state}} = \begin{bmatrix} -19.24 \text{ ft/s} \\ 0 \\ -11.99 \text{ deg/s} \\ -177.84 \text{ deg} \\ -2.35 \text{ deg} \end{bmatrix}_{aileron} \tag{4.5.16}$$

and the steady state response to a rudder unit step input is

$$\begin{bmatrix} v \\ p \\ r \\ \phi \\ \beta \end{bmatrix}_{\substack{steady \\ state}} = \begin{bmatrix} -11.00 \text{ ft/s} \\ 0 \\ -10.18 \text{ deg/s} \\ -150.36 \text{ deg} \\ -1.35 \text{ deg} \end{bmatrix}_{rudder} \tag{4.5.17}$$

It must be realised that the steady state values given in equations (4.5.16) and (4.5.17) serve only to give an indication of the control sensitivity of the aeroplane. At such large roll attitudes the small perturbation model ceases to apply and in practice significant changes in the aerodynamic operating conditions would accompany the response. The actual steady state values would undoubtedly be somewhat different and could only be ascertained with a full non-linear simulation model. This illustration indicates the limiting nature of a small perturbation model for the analysis of lateral–directional dynamics and the need to exercise care in its interpretation.

4.5.1.1 The Characteristic Equation

The lateral–directional characteristic polynomial for a classical aeroplane is fourth order; it determines the common denominator of the lateral and directional response transfer functions and, when equated to zero, defines the characteristic equation which may be written as

$$As^4 + Bs^3 + Cs^2 + Ds + E = 0 \tag{4.5.18}$$

The characteristic equation (4.5.18) most commonly factorises into two real roots and a pair of complex roots which are most conveniently written as

$$(1 + (1/T_s))(1 + (1/T_r))(s^2 + 2\zeta_d\omega_d s + \omega_d^2) = 0 \tag{4.5.19}$$

As indicated previously, the first real root in equation (4.5.19) describes the non-oscillatory spiral mode, the second real root describes the non-oscillatory roll subsidence mode and the pair of complex roots describe the oscillatory dutch roll mode. Now, since the equations of motion from which the characteristic equation is derived are referred to a wind axes reference, the stability modes comprising equation (4.5.19) provide a complete description of the lateral–directional stability properties of the aeroplane with respect to the total steady velocity vector and subject to the constraints of small perturbation motion.

When the equations of motion are referred to a body axes system, the state equation is fifth order and the characteristic equation is also of fifth order. The solution of the characteristic equation then has the following factors:

$$s(1 + (1/T_s))(1 + (1/T_r))(s^2 + 2\zeta_d\omega_d s + \omega_d^2) = 0 \tag{4.5.20}$$

The modes are unchanged except for the addition of a zero root which indicates neutral stability. The zero root results from the addition of yaw angle to the state equation and indicates neutral stability in yaw, or heading. Interpretation of lateral–directional dynamics is unchanged and the additional information indicates the aeroplane to have an indeterminate yaw or heading angle. In other words, lateral–directional dynamics are evaluated about the steady total velocity vector which assumes an arbitrary direction in azimuth, yaw or heading. Interpretation of the non-zero roots of the characteristic equation is most easily accomplished if reference is first made to the properties of the classical mass–spring–damper system.

Unlike the longitudinal dynamics, interpretation of the lateral–directional dynamics is not quite so straightforward as the stability modes are not so distinct; there usually exists a significantly greater degree of mode coupling, or interaction. This tends to make the necessary simplifying assumptions less appropriate with a consequent reduction of confidence in the observations. However, an assortment of well tried procedures for interpreting the dynamic characteristics of the well behaved aeroplane exist and these will be discussed below. The principal objective of course is to identify the aerodynamic drivers for each of the stability modes. The connection between the observed dynamics of the aeroplane and its aerodynamic characteristics is made by comparing equation (4.5.18) with either of equation (4.5.19) or (4.5.20), and then referring to the definitions of the coefficients in equation (4.5.18) in terms of aerodynamic stability derivatives. It will be appreciated immediately that further analytical progress is impossibly difficult unless some gross simplifying assumptions are made. Means for dealing with this difficulty requires the derivation of reduced order models as described in Section 4.5.3.

4.5.2 The Dynamic Stability Modes

As for the longitudinal stability modes, whenever the aeroplane is disturbed from its equilibrium trim state the lateral–directional stability modes will also be excited. Again, the disturbance may be initiated by pilot control action, a change in power setting, airframe configuration changes, such as flap deployment, and by external influences such as gusts and turbulence.

4.5.2.1 The Roll Subsidence Mode

The *roll subsidence mode*, or simply the *roll mode*, is a non-oscillatory lateral characteristic which is usually substantially decoupled from the spiral and dutch roll modes. Since it is non-oscillatory it is described by a single real root of the characteristic polynomial, and it manifests itself as an exponential lag characteristic in rolling motion. The aero-mechanical principles governing the behaviour of the mode are shown in Fig. 4.5-3.

With reference to Fig. 4.5-3, the aircraft is viewed from the rear so the indicated motion is shown in the same sense as it would be experienced by the pilot. Assume that the aircraft is constrained to the single degree of freedom

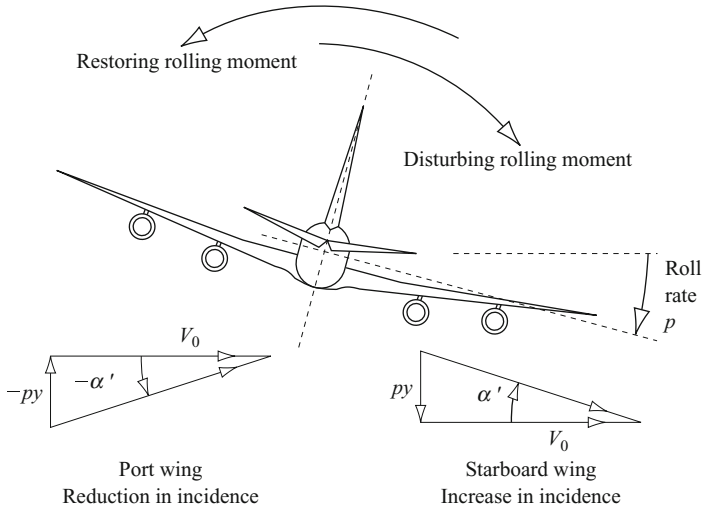


Figure 4.5-3 The roll subsidence mode.

motion in roll about the *ox* axis only, and that it is initially in trimmed wings level flight. If then, the aeroplane experiences a positive disturbing rolling moment it will commence to roll with an angular acceleration in accordance with Newton's second law of motion. In rolling motion the wing experiences a component of velocity normal to the wing py, where y is the spanwise coordinate measured from the roll axis *ox*. As indicated in Fig. 4.5-3 this results in a small increase in incidence on the down-going starboard wing and a small decrease in incidence on the up-going port wing. The resulting differential lift gives rise to a restoring rolling moment as indicated. The corresponding resulting differential induced drag would also give rise to a yawing moment, but this is usually sufficiently small that it is ignored. Thus following a disturbance the roll rate builds up exponentially until the restoring moment balances the disturbing moment and a steady roll rate is established. In practice, of course, this kind of behaviour would be transient rather than continuous as implied in the figure. The physical behaviour explained is simple "paddle" damping and is stabilising in effect in all aeroplanes operating in normal, aerodynamically linear, flight regimes. For this reason, the stability mode is sometimes referred to as the *damping in roll*.

In some modern combat aeroplanes which are designed to operate in seriously non-linear aerodynamic conditions, for example, at angles of attack approaching 90°, it is possible for the physical conditions governing the roll mode to break down completely. The consequent loss of roll stability can result in rapid roll departure followed by complex lateral–directional motion of a hazardous nature. However, in the conventional aeroplane the roll mode appears to the pilot as a lag in roll response to controls. The lag time constant is largely dependent on the moment of inertia in roll and the aerodynamic properties of the wing, and is typically around 1 s or less.

4.5.2.2 The Spiral Mode

The *spiral mode* is also non-oscillatory and is determined by the other real root in the characteristic polynomial. When excited, the mode dynamics are usually slow to develop and involve complex coupled motion in roll, yaw and sideslip. The dominant aeromechanical principles governing the mode dynamics are shown in Fig. 4.5-4. The mode characteristics are very dependent on the lateral static stability and on the directional static stability of the aeroplane and these topics are discussed in Sections 4.3.4 and 4.3.5.

The mode is usually excited by a disturbance in sideslip which typically follows a disturbance in roll causing a wing to drop. Assume that the aircraft is initially in trimmed wings level flight and that a disturbance causes a small positive roll angle ϕ to develop; left unchecked this results in a small positive sideslip velocity v as indicated at (a) in Fig. 4.5-4. The sideslip puts the fin at incidence β which produces lift, and which in turn generates a yawing moment to turn the aircraft into the direction of the sideslip. The yawing motion produces differential lift across the wing span which, in turn, results in a rolling moment causing the starboard wing to drop further thereby exacerbating the situation. This developing divergence is indicated at (b) and (c) in Fig. 4.5-4. Simultaneously, the dihedral effect of the wing generates a negative restoring rolling moment due to sideslip which acts to return the wing to a level attitude. Some additional restoring rolling moment is also generated by the fin lift force when it acts at a point above the roll axis *ox*, which is usual.

Therefore, the situation is one in which the fin effect, or directional static stability, and the dihedral effect, or lateral static stability, act in opposition to create this interesting dynamic condition. Typically, the requirements for

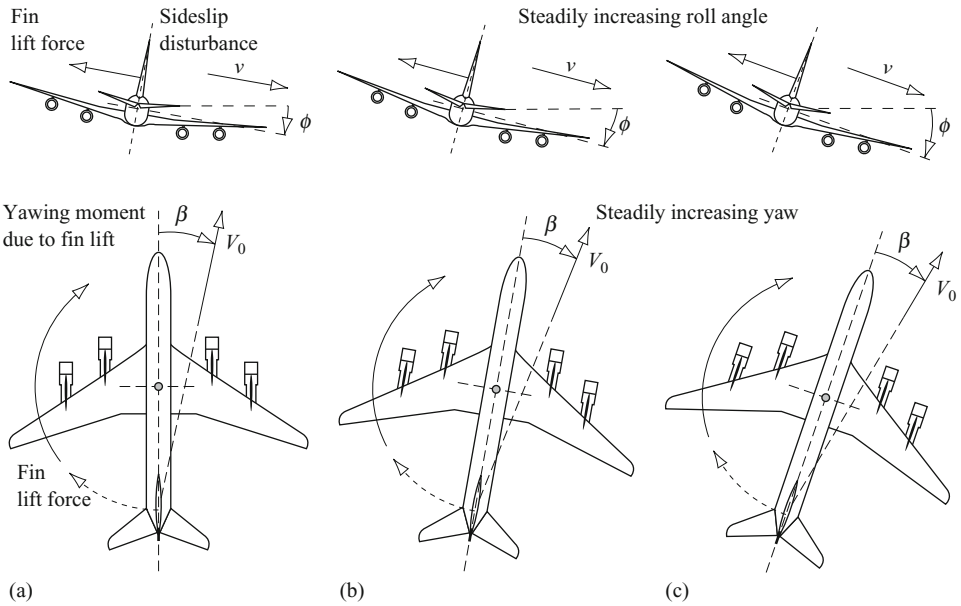


Figure 4.5-4 The spiral mode development.

lateral and directional static stability are such that the opposing effects are very nearly equal. When dihedral effect is greater the spiral mode is stable, and hence convergent, and when the fin effect is greater the spiral mode is unstable, and hence divergent. Since these effects are nearly equal the spiral mode will be nearly neutrally stable, and sometimes it may even be neutrally stable, that is, it will be neither convergent nor divergent. Since the mode is non-oscillatory it manifests itself as a classical exponential convergence nor divergence and, since it is nearly neutral, the time constant is very large, typically 100 s or more. This means that when the mode is stable the wing is slow to recover a level attitude following a disturbance and when it is unstable the rate at which it diverges is also very slow. When it is neutral the aircraft simply flies a turn at constant roll attitude.

Now it is the unstable condition which attracts most attention for obvious reasons. Once the mode is excited the aircraft flies a slowly diverging path in both roll and yaw and since the vertical forces are no longer in equilibrium the aircraft will also lose height. Thus the unstable flight path is a spiral descent which left unchecked will end when the aircraft hits the ground! However, since the rate at which the mode diverges is usually very slow most pilots can cope with it. Consequently, an unstable spiral mode is permitted provided its time constant is sufficiently large. Because the mode is very slow to develop the accelerations in the resulting motion are insignificantly small and the motion cues available to the pilot are almost imperceptible. In a spiral departure the visual cues become the most important cues to the pilot. It is also important to appreciate that a spiral departure is not the same as a spin. Spinning motion is a fully stalled flight condition whereas in a spiral descent the wing continues to fly in the usual sense.

4.5.2.3 The Dutch Roll Mode

The *dutch roll* mode is a classical damped oscillation in yaw, about the oz axis of the aircraft, which couples into roll and, to a lesser extent, into sideslip. The motion described by the dutch roll mode is therefore a complex interaction between all three lateral–directional degrees of freedom. Its characteristics are described by the pair of complex roots in the characteristic polynomial. Fundamentally, the dutch roll mode is the lateral–directional equivalent of the longitudinal short period mode. Since the moments of inertia in pitch and yaw are of similar magnitude the frequency of the dutch roll mode and the longitudinal short period mode are of similar order. However, the fin is generally less effective than the tailplane as a damper and the damping of the dutch roll mode is often inadequate. The dutch roll mode is so called since the motion of the aeroplane following its excitation is said to resemble the rhythmical flowing motion of a dutch skater on a frozen canal. One cycle of a typical dutch rolling motion is shown in Fig. 4.5-5.

The physical situation applying can be appreciated by imagining that the aircraft is restrained in yaw by a torsional spring acting about the yaw axis oz, the spring stiffness being aerodynamic and determined largely by the fin. Thus when in straight, level trimmed equilibrium flight a disturbance in yaw causes the "aerodynamic spring" to produce a restoring yawing moment which results in classical oscillatory motion. However, once the yaw oscillation is established the relative velocity of the air over the port and starboard wing also varies in an oscillatory manner giving rise to oscillatory differential lift and drag perturbations. This aerodynamic coupling gives rise in turn to an oscillation

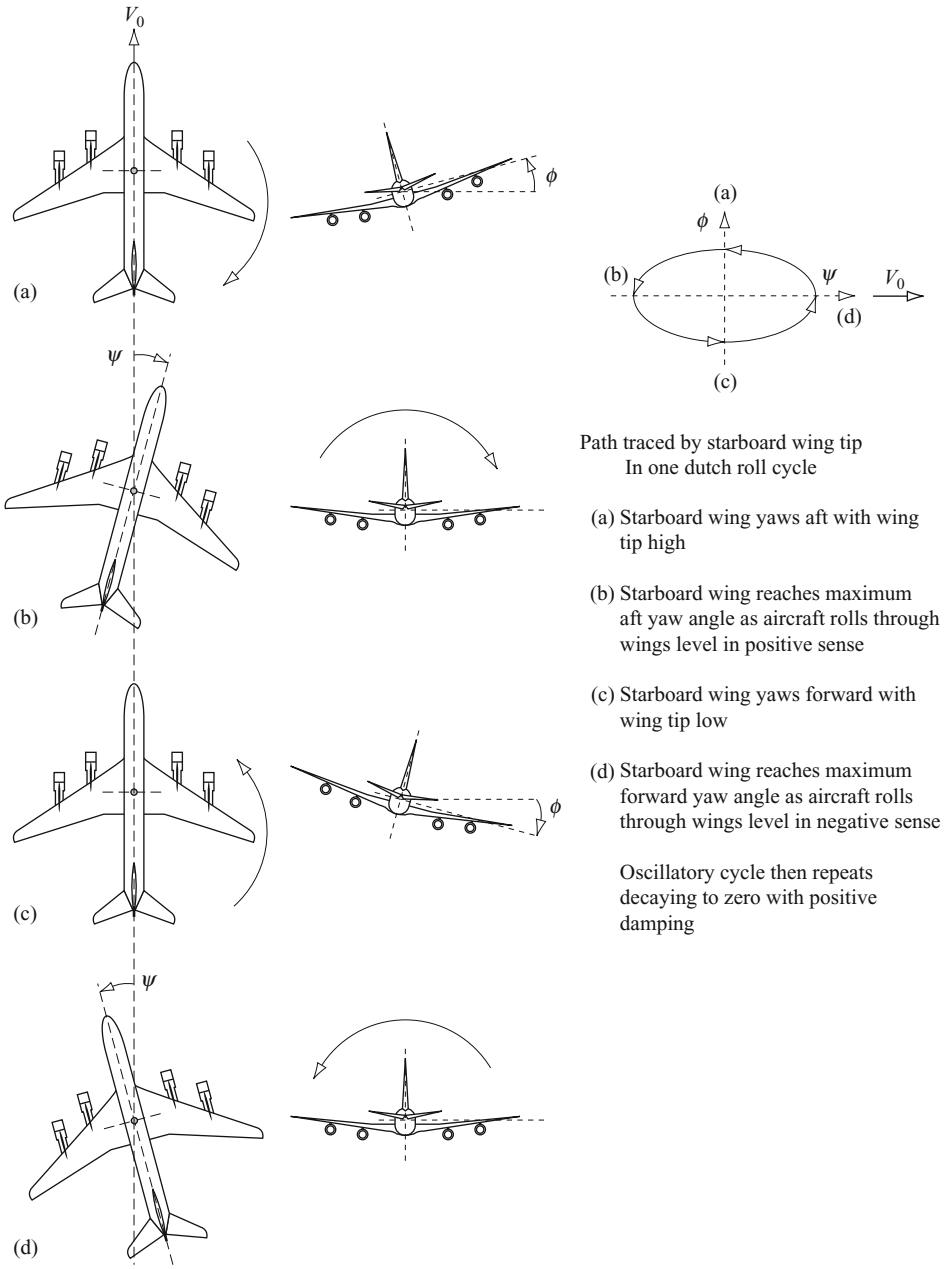


Figure 4.5-5 The oscillatory dutch roll mode.

in roll which lags the oscillation in yaw by approximately 90°. This phase difference between yawing and rolling motion means that the forward going wing panel is low and the aft going wing panel is high as indicated in Fig. 4.5-5.

Consequently, the classical manifestation of the dutch roll mode is given by the path described by the wing tips relative to the horizon and which is usually elliptical, also shown in Fig. 4.5-5. The peak roll to peak yaw ratio is usually less than 1, as indicated, and is usually associated with a stable dutch roll mode. However, when the peak roll to peak yaw ratio is greater than 1 an unstable dutch roll mode is more likely.

Whenever the wing is disturbed from level trim, left to its own devices the aeroplane starts to slip sideways in the direction of the low wing. Thus the oscillatory rolling motion leads to some oscillatory sideslipping motion in dutch rolling motion although the sideslip velocity is generally small. Thus it is fairly easy to build up a visual picture of the complex interactions involved in the dutch roll mode. In fact the motion experienced in a dutch rolling aircraft would seem to be analogous to that of a ball bearing dropped into an inclined channel having a semi-circular cross section. The ball bearing rolls down the inclined channel whilst oscillating from side to side on the circular surface.

Both the damping and stiffness in yaw, which determine the characteristics of the mode, are largely determined by the aerodynamic properties of the fin, a large fin being desirable for a well behaved stable dutch roll mode.

Unfortunately this contradicts the requirement for a stable spiral mode. The resulting aerodynamic design compromise usually results in aeroplanes with a mildly unstable spiral mode and a poorly damped dutch roll mode. Of course, the complexity of the dynamics associated with the dutch roll mode suggests that there must be other aerodynamic contributions to the mode characteristics in addition to the fin. This is generally the case and it is quite possible for the additional aerodynamic effects to be as significant as the aerodynamic properties of the fin if not more so. However, one thing is quite certain; it is very difficult to quantify all the aerodynamic contributions to the dutch roll mode characteristics with any degree of confidence.

4.5.3 Reduced Order Models

Unlike the longitudinal equations of motion it is more difficult to solve the lateral–directional equations of motion approximately. Because of the motion coupling present, to a greater or lesser extent, in all three modes dynamics, the modes are not so distinct and simplifying approximations are less relevant with the consequent loss of accuracy. Response transfer functions derived from reduced order models based on simplified approximate equations of motion are generally insufficiently accurate to be of any real use other than as a means for providing enhanced understanding of the aeromechanics of lateral–directional motion.

The simplest, and most approximate, solution of the characteristic equation provides an initial estimate for the two real roots only. This approximate solution of the lateral–directional characteristic equation (4.5.18) is based on the observation that conventional aeroplanes give rise to coefficients A, B, C, D and E that have relative values which do not change very much with flight condition. Typically, A and B are relatively large whilst D and E are relatively small, in fact E is very often close to zero. Further, it is observed that $B \gg A$ and $E \ll D$ suggesting the following real roots as approximate solutions of the characteristic equation

$$\begin{aligned}(s + (1/T_r)) &\cong (s + B/A))\\ (s + (1/T_s)) &\cong (s + (E/D))\end{aligned} \tag{4.5.21}$$

No such simple approximation for the pair of complex roots describing the dutch roll mode may be determined. Further insight into the aerodynamic drivers governing the characteristics of the roll and spiral modes may be made, with some difficulty, by applying assumptions based on the observed behaviour of the modes to the polynomial expressions for A, B, D and E. Fortunately, the same information may be deduced by a rather more orderly process involving a reduction in order of the equations of motion. The approximate solutions for the non-oscillatory modes as given by equations (4.5.21) are only useful for preliminary mode evaluations, or as a check of computer solutions, when the numerical values of the coefficients in the characteristic equation are known.

4.5.3.1 The Roll Mode Approximation

Provided the perturbation is small, the roll subsidence mode is observed to involve almost pure rolling motion with little coupling into sideslip or yaw. Thus a reduced order model of the lateral–directional dynamics retaining only the roll mode follows by removing the side force and yawing moment equations from the lateral–directional state equation (4.5.1) to give

$$\begin{bmatrix}\dot{p}\\ \dot{\phi}\end{bmatrix} = \begin{bmatrix}l_p & l_\phi\\ 1 & 0\end{bmatrix}\begin{bmatrix}p\\ \phi\end{bmatrix} + \begin{bmatrix}l_\xi & l_\zeta\\ 0 & 0\end{bmatrix}\begin{bmatrix}\xi\\ \zeta\end{bmatrix} \tag{4.5.22}$$

Further, if aircraft wind axes are assumed then $l_\phi = 0$ and equation (4.5.22) reduces to the single degree of freedom rolling moment equation

$$\dot{p} = l_p p + l_\xi \xi + l_\zeta \zeta \tag{4.5.23}$$

The roll response to aileron transfer function is easily derived from equation (4.5.23). Taking the Laplace transform of equation (4.5.23), assuming zero initial conditions and assuming that the rudder is held fixed, $\zeta = 0$, then

$$sp(s) = l_p p(s) + l_\xi \xi(s) \tag{4.5.24}$$

which on rearranging may be written as

$$\frac{p(s)}{\xi(s)} = \frac{l_\xi}{(s - l_p)} \equiv \frac{k_p}{(s + (1/T_r))} \tag{4.5.25}$$

The transfer function given by equation (4.5.25) is the approximate reduced order equivalent to the transfer function given by equation (4.5.3) and is the transfer function of a simple first order lag with time constant T_r. For small

perturbation motion equation (4.5.25) describes the first second or two of roll response to aileron with a reasonable degree of accuracy and is especially valuable as a means for identifying the dominant physical properties of the airframe which determine the roll mode time constant. With reference to the definitions of the concise aerodynamic stability derivatives, the roll mode time constant is determined approximately by

$$T_r \cong -\frac{1}{l_p} = -\frac{(I_x I_z - I_{xz}^2)}{(I_z \mathring{L}_p + I_{xz} \mathring{N}_p)} \tag{4.5.26}$$

Since $I_x \gg I_{xz}$ and $I_z \gg I_{xz}$ then equation (4.5.26) may be further simplified to give the classical approximate expression for the roll mode time constant

$$T_r \cong -\frac{I_x}{\mathring{L}_p} \tag{4.5.27}$$

where I_x is the moment of inertia in roll and $\mathring{L}_p$ is the dimensional derivative describing the aerodynamic damping in roll.

4.5.3.2 The Spiral Mode Approximation

Since the spiral mode is very slow to develop following a disturbance, it is usual to assume that the motion variables v, p and r are quasi-steady relative to the time scale of the mode. Whence $\dot{v} = \dot{p} = \dot{r} = 0$ and the lateral–directional state equation (4.5.1) may be written as

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} y_v & y_p & y_r & y_\phi \\ l_v & l_p & l_r & l_\phi \\ n_v & n_p & n_r & n_\phi \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \\ r \\ \phi \end{bmatrix} + \begin{bmatrix} y_\xi & y_\zeta \\ l_\xi & l_\zeta \\ n_\xi & n_\zeta \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \xi \\ \zeta \end{bmatrix} \tag{4.5.28}$$

Further, if aircraft wind axes are assumed $l_\phi = n_\phi = 0$ and if the controls are assumed fixed such that unforced motion only is considered $\xi = \zeta = 0$ then equation (4.5.28) simplifies to

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} y_v & y_p & y_r & y_\phi \\ l_v & l_p & l_r & 0 \\ n_v & n_p & n_r & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v \\ p \\ r \\ \phi \end{bmatrix} \tag{4.5.29}$$

The first three rows in equation (4.5.29) may be rearranged to eliminate the variables v and r to give a reduced order equation in which the variables are roll rate p and roll angle ϕ only

$$\begin{bmatrix} 0 \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} y_v \dfrac{(l_p n_r - l_r n_p)}{(l_r n_v - l_v n_r)} + y_p + y_r \dfrac{(l_v n_p - l_p n_v)}{(l_r n_v - l_v n_r)} & y_\phi \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p \\ \phi \end{bmatrix} \tag{4.5.30}$$

The first element of the first row of the reduced order state matrix in equation (4.5.30) may be simplified since the terms involving y_v and y_p are assumed to be insignificantly small compared with the term involving y_r. Thus equation (4.5.30) may be rewritten as

$$\begin{bmatrix} 0 \\ \dot{\phi} \end{bmatrix} = \begin{bmatrix} y_r \dfrac{(l_v n_p - l_p n_v)}{(l_r n_v - l_v n_r)} & y_\phi \\ 1 & 0 \end{bmatrix} \begin{bmatrix} p \\ \phi \end{bmatrix} \tag{4.5.31}$$

Since $\dot{\phi} = p$, equation (4.5.31) may be reduced to the single degree of freedom equation describing, approximately, the unforced rolling motion involved in the spiral mode

$$\dot{\phi} + \left(\frac{y_\phi (l_r n_v - l_v n_r)}{y_r (l_v n_p - l_p n_v)} \right) \phi = 0 \tag{4.5.32}$$

The Laplace transform of equation (4.5.32), assuming zero initial conditions, is

$$\phi(s) \left(s + \left(\frac{y_\phi (l_r n_v - l_v n_r)}{y_r (l_v n_p - l_p n_v)} \right) \right) \equiv \phi(s)(s + (1/T_s)) = 0 \tag{4.5.33}$$

It should be noted that equation (4.5.33) is the reduced order lateral–directional characteristic equation retaining a very approximate description of the spiral mode characteristics only, whence an approximate expression for the time constant of the spiral mode is defined as

$$T_s \cong \frac{y_r(l_v n_p - l_p n_v)}{y_\phi(l_r n_v - l_v n_r)} \tag{4.5.34}$$

The spiral mode time constant 4.5.34 may be expressed conveniently in terms of the dimensional or dimensionless aerodynamic stability derivatives to provide a more direct link with the aerodynamic mode drivers. Noting that $\mathring{Y}_r << mU_e$, so $y_r \cong -U_e \equiv -V_0$, and that $y_\phi = g$ since aircraft wind axes are assumed, then equation (4.5.34) may be re-stated as

$$T_s \cong -\frac{U_e\left(\mathring{L}_v\mathring{N}_p - \mathring{L}_p\mathring{N}_v\right)}{g\left(\mathring{L}_r\mathring{N}_v - \mathring{L}_v\mathring{N}_r\right)} \equiv -\frac{V_0(L_vN_p - L_pN_v)}{g(L_rN_v - L_vN_r)} \tag{4.5.35}$$

Now a stable spiral mode requires that the time constant T_s is positive. Typically for most aeroplanes, especially in sub-sonic flight

$$(L_vN_p - L_pN_v) > 0$$

and the condition for the mode to be stable simplifies to the approximate classical requirement that

$$L_vN_r > L_rN_v \tag{4.5.36}$$

Further analysis of this requirement is only possible if the derivatives in equation (4.5.36) are expressed in terms of the aerodynamic properties of the airframe. This means that L_v, dihedral effect, and N_r (damping in yaw) should be large whilst N_v (the yaw stiffness), should be small. Rolling moment due to yaw rate, L_r, is usually significant in magnitude and positive. In very simple terms aeroplanes with small fins and reasonable dihedral are more likely to have a stable spiral mode.

4.5.3.3 The Dutch Roll Mode Approximation

For the purpose of creating a reduced order model to describe the dutch roll mode it is usual to make the rather gross assumption that dutch rolling motion involves no rolling motion at all. Clearly this is contradictory, but it is based on the fact that the mode is firstly a yawing oscillation and aerodynamic coupling causes rolling motion as a secondary effect. It is probably true that for most aeroplanes the roll to yaw ratio in dutch rolling motion is less than 1, and in some cases it may be much less than 1, which gives the assumption some small credibility from which the lateral–directional state equation (4.5.1) may be simplified by writing

$$\dot{p} = p = \dot{\phi} = \phi = 0$$

As before, if aircraft wind axes are assumed $l_\phi = n_\phi = 0$ and if the controls are assumed fixed such that unforced motion only is considered $\xi = \zeta = 0$ then equation (4.5.1) simplifies to

$$\begin{bmatrix}\dot{v}\\ \dot{r}\end{bmatrix} = \begin{bmatrix}y_v & y_r\\ n_v & n_r\end{bmatrix}\begin{bmatrix}v\\ r\end{bmatrix} \tag{4.5.37}$$

If equation (4.5.37) is written as

$$\dot{\mathbf{x}}_d = \mathbf{A}_d\mathbf{x}_d$$

then the reduced order characteristic equation describing the approximate dynamic characteristics of the dutch roll mode is given by

$$\Delta_d(s) = \det[s\mathbf{I} - \mathbf{A}_d] = \begin{vmatrix} s - y_v & -y_r\\ -n_v & s - n_r\end{vmatrix} = 0$$

or

$$\Delta_d(s) = s^2 - (n_r + y_v)s + (n_r y_v - n_v y_r) = 0 \tag{4.5.38}$$

Therefore the damping and frequency properties of the mode are given approximately by

$$\begin{aligned} 2\zeta_d\omega_d &\cong -(n_r + y_v)\\ \omega_d^2 &\cong (n_r y_v - n_v y_r)\end{aligned} \tag{4.5.39}$$

The expressions given by equations (4.5.39) can be re-stated in terms of dimensional aerodynamic stability derivatives. Further approximating simplifications are made by assuming $\mathring{Y}_r << mU_e$, so that $y_r \cong -U_e \equiv -V_0$, and by assuming, quite correctly, that both I_x and I_z are usually much greater than I_{xz}. It then follows that

$$2\zeta_d\omega_d \cong -\left(\frac{\mathring{N}_r}{I_z}+\frac{\mathring{Y}_v}{m}\right)$$
$$\omega_d^2 \cong \left(\frac{\mathring{N}_r}{I_z}\frac{\mathring{Y}_v}{m}+V_0\frac{\mathring{N}_v}{I_z}\right)\cong V_0\frac{\mathring{N}_v}{I_z} \tag{4.5.40}$$

Comparing the damping and frequency terms in the expressions in equations (4.5.40) with those of the mass–spring–damper it is easy to identify the roles of those aerodynamic stability derivatives which are dominant in determining the characteristics of the dutch roll mode. For example, $\mathring{N}_r$ is referred to as the yaw damping derivative and $\mathring{N}_v$ is referred to as the yaw stiffness derivative, and both are very dependent on the aerodynamic design of the fin and the fin volume ratio.

Although the dutch roll mode approximation gives a rather poor impression of the real thing, it is useful as a means for gaining insight into the physical behaviour of the mode and its governing aerodynamics.

Example 4.5.2

It has been stated that the principal use of the lateral–directional reduced order models is for providing insight into the aerodynamic mode drivers. With the exception of the transfer function describing roll rate response to aileron, transfer functions derived from the reduced order models are not commonly used in analytical work as their accuracy is generally poor. However, it is instructive to compare the values of the modes characteristics obtained from reduced order models with those obtained in the solution of the full order equations of motion.

Consider the Douglas DC-8 aircraft of Example 4.5.1. The equations of motion referred to wind axes are given by equation (4.5.10) and the solution gives the characteristic equation (4.5.14). The unfactorised characteristic equation is

$$\Delta(s) = s^4 + 1.5898s^3 + 1.7820s^2 + 1.9200s + 0.0125 = 0 \tag{4.5.41}$$

In accordance with the expression given in equations (4.5.21), approximate values for the roll mode and spiral mode time constants are given by

$$T_r \cong \frac{A}{B} = \frac{1}{1.5898} = 0.629\text{s}$$
$$T_s \cong \frac{D}{E} = \frac{1.9200}{0.0125} = 153.6\text{s} \tag{4.5.42}$$

The approximate roll mode time constant does not compare particularly well with the exact value of 0.75 s, whereas the spiral mode time constant compares extremely well with the exact value of 154 s.

The approximate roll rate response to aileron transfer function, given by equation (4.5.25), may be evaluated by obtaining the values for the concise derivatives l_p and l_ξ from equation (4.5.10) whence

$$\frac{p(s)}{\xi(s)} = \frac{-1.62}{(s+1.232)}\ \text{deg/s/deg} \tag{4.5.43}$$

With reference to equation (4.5.25) an approximate value for the roll mode time constant is given by

$$T_r \cong \frac{1}{1.232} = 0.812\,\text{s} \tag{4.5.44}$$

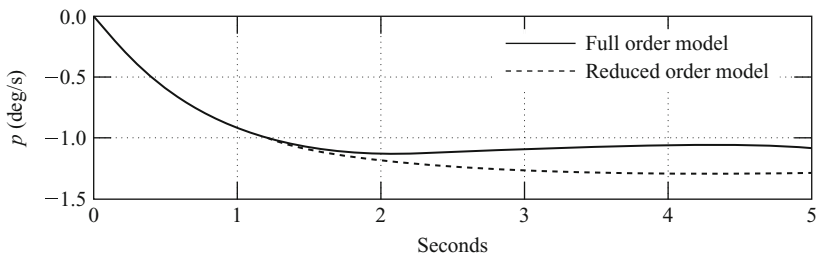


Figure 4.5-6 Roll rate response to 1° aileron step input.

and this value compares rather more favourably with the exact value. The short term roll rate response of the DC-8 to a 1° aileron step input as given by equation (4.5.43) is shown in Fig. 4.5-6 where it is compared with the exact response of the full order model as given by equations (4.5.12).

Clearly, for the first 2 s, or so, the match is extremely good which confirms the assumptions made about the mode to be valid provided the period of observation of roll behaviour is limited to the time scale of the roll mode. The approximate roll mode time constant calculated by substituting the appropriate derivative and roll inertia values, given in the aircraft data, into the expression given by equation (4.5.27) results in a value almost the same as that given by equation (4.5.44). This simply serves to confirm the validity of the assumptions made about the roll mode.

With reference to equations (4.5.34) and (4.5.35) the approximate spiral mode time constant may be written in terms of concise derivatives as

$$T_s \cong -\frac{U_e(l_v n_p - l_p n_v)}{g(l_r n_v - l_v n_r)} \tag{4.5.45}$$

Substituting values for the concise derivatives obtained from equation (4.5.10), the velocity U_e and g, then

$$T_s \cong -\frac{468.2(0.0002 + 0.00343)}{32.2(0.0011 - 0.00149)} = 135.34 \text{ s} \tag{4.5.46}$$

Clearly this approximate value of the spiral mode time constant does not compare so well with the exact value of 154 s. However, this is not so important since the mode is very *slow* in the context of normal piloted manoeuvring activity. The classical requirement for spiral mode stability given by the inequality condition of equation (4.5.36) is satisfied since

$$0.00149 > 0.0011$$

Notice how close the values of the two numbers are, suggesting the mode to be close to neutrally stable in the time scale of normal transient response. This observation is quite typical of a conventional aeroplane like the DC-8.

Approximate values for the dutch roll mode damping ratio and undamped natural frequency are obtained by substituting the relevant values for the concise derivatives, obtained from equation (4.5.10) into the expressions given by equations (4.5.39) Thus, approximately

$$\omega_d \cong 1.152 \text{ rad/s}$$
$$\zeta_d \cong 0.135$$

These approximate values compare reasonably well with the exact values which are a natural frequency of 1.2 rad/s and a damping ratio of 0.11. Such a good comparison is not always achieved and merely emphasises once more the validity of the assumptions about the dutch roll mode in this particular application. The implication is that at the flight condition of interest the roll to yaw ratio of the dutch roll mode in the DC-8 is significantly less than 1 and, indeed, this may be inferred from either Fig. 4.5-1 or 4.5-2.

4.5.4 Frequency Response

It is useful, and sometimes necessary, to investigate the lateral–directional response properties of an aeroplane in the frequency domain. The reasons why such an investigation might be made are much the same as those given for the longitudinal case in Section 4.4.4. Again, the Bode diagram is the most commonly used graphical tool for lateral–directional frequency response analysis. The method of construction of the Bode diagram and its interpretation follow the general principles described in Section 4.4.4 and are not repeated here. Since it is difficult to generalise, a typical illustration of lateral–directional frequency response analysis is given in the following example.

Example 4.5.3

The lateral–directional frequency response of the Douglas DC-8 aircraft is evaluated for the same flight condition as Examples 4.5.1 and 4.5.2. The total number of transfer functions which could be evaluated on a Bode diagram is 10, given by equations (4.5.12) and (4.5.13), and to create ten Bode diagrams would be prohibitively lengthy in the present context. Since the essential frequency response information can be obtained from a much smaller number of transfer functions the present example is limited to four transfer functions only. The chosen transfer functions were selected from equations (4.5.12) and (4.5.13) all are referred to aircraft wind axes and are repeated here for convenience:

$$\frac{\phi(s)}{\xi(s)} = \frac{-1.62(s^2 + 0.362s + 1.359)}{(s + 0.0065)(s + 1.329)(s^2 + 0.254s + 1.433)} \text{ rad/rad (deg/deg)}$$

$$\frac{\beta(s)}{\xi(s)} = \frac{0.0188(s + 0.197)(s - 7.896)}{(s + 0.0065)(s + 1.329)(s^2 + 0.254s + 1.433)} \text{ rad/rad (deg/deg)}$$

$$\frac{r(s)}{\varsigma(s)} = \frac{-0.864(s + 1.335)(s^2 - 0.03s + 0.109)}{(s + 0.0065)(s + 1.329)(s^2 + 0.254s + 1.433)} \text{ rad/s/rad (deg/s/deg)} \tag{4.5.47}$$

$$\frac{p(s)}{\varsigma(s)} = \frac{0.392s(s + 1.85)(s - 2.566)}{(s + 0.0065)(s + 1.329)(s^2 + 0.254s + 1.433)} \text{ rad/s/rad (deg/s/deg)}$$

The first two transfer functions (4.5.47) describe lateral response to the lateral command (aileron) variable, the third transfer function describes directional response to the directional command (rudder) variable, the last transfer function was chosen to illustrate cross-coupling and describes lateral response to the directional command variable. Now consider the frequency response of each transfer function in turn.

The frequency response of roll attitude ϕ to aileron input ξ is shown in Fig. 4.5-7. The most obvious features of the Bode diagram are the very high steady state gain, 45 dB, and the very small peak at the dutch roll frequency. The *roll-off* in phase behaves quite conventionally in accordance with the transfer function properties. The high zero frequency gain corresponds with a gain ratio of approximately 180. This means that following a 1° aileron step input the aeroplane will settle at a roll attitude of −180°, in other words inverted! Clearly, this is most inappropriate for a large civil transport aeroplane and serves as yet another illustration of the limitations of linear system modelling. Such a large amplitude excursion is definitely not a small perturbation and should not be regarded as such. However, the high zero frequency, or steady state, gain provides a good indication of the roll control sensitivity. As the control input frequency is increased the attitude response attenuates steadily with increasing phase lag, the useful bandwidth being a little above the spiral mode break frequency $1/T_s$. However, at all frequencies up to that corresponding with the roll subsidence mode break frequency, $1/T_r$, the aeroplane will respond to aileron since the gain is always greater than 0 dB; it is the steady reduction in control sensitivity that will be noticed by the pilot. Since the dutch roll damping ratio is relatively low at 0.11, an obvious peak might be expected in the gain plot at the dutch roll frequency. Clearly this is not the case. Inspection of the relevant transfer function in equation (4.5.47) shows that the second

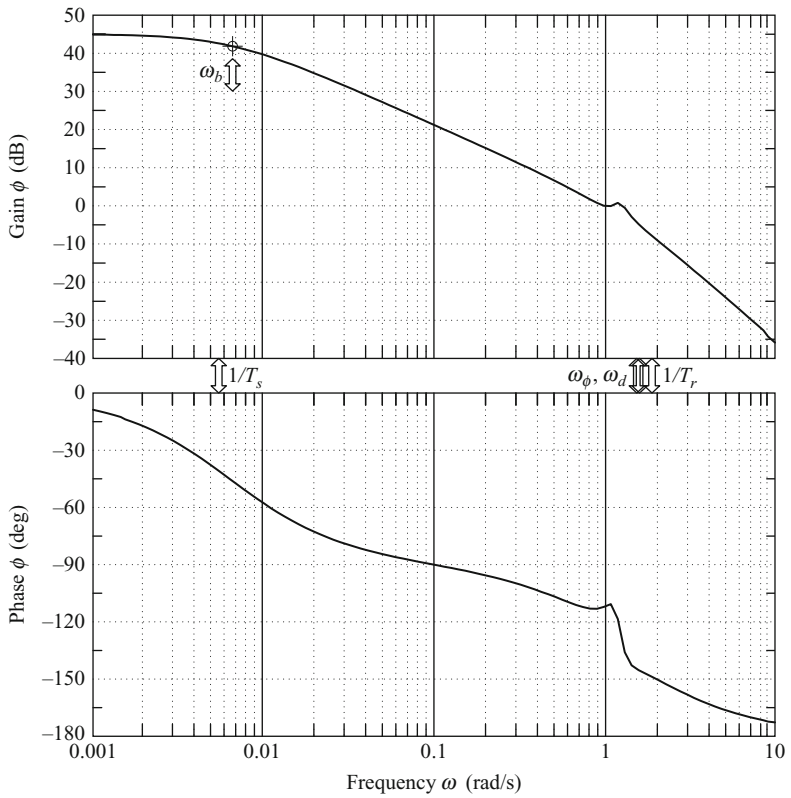


Figure 4.5-7 DC-8 roll attitude frequency response to aileron.

order numerator factor very nearly cancels the dutch roll roots in the denominator. This means that the dutch roll dynamics will not be very obvious in the roll attitude response to aileron in accordance with the observation. This conclusion is also confirmed by the time history response shown in Fig. 4.5-1. In fact the dutch roll cancellation is sufficiently close that it is permissible to write the transfer function in approximate form

$$\frac{\phi(s)}{\xi(s)} = \frac{-1.62}{(s+0.0065)(s+1.329)} \text{ rad/rad (deg/deg)} \tag{4.5.48}$$

with little loss of meaning. The time response plot and the Bode diagram derived from this approximate transfer function correspond closely with those derived from the full transfer function and may be interpreted to achieve the same conclusions for all practical purposes.

The frequency response of sideslip angle β to aileron input ξ is shown in Fig. 4.5-8 and corresponds with the second transfer function given in equation (4.5.47). Again, there are no real surprises here. The transfer function is non-minimum phase since the numerator term $1/T_{\beta_2}$ is negative which introduces 90° of phase lag at the corresponding break frequency. In this response variable the dutch roll gain peak is clearly visible although at the dutch roll frequency the gain is attenuated by about −20 dB which means that the pilot would see no significant oscillatory sideslip behaviour. Again, it is established that the usable bandwidth is a little higher than the spiral mode break frequency $1/T_s$.

The frequency response of yaw rate r to rudder input ζ is shown in Fig. 4.5-9. This transfer function describes the typical classical directional response to control and the frequency response, shown in Fig. 4.5-9, has some interesting features. The gain plot shows a steady but significant attenuation with increasing frequency to reach a minimum of about −30 dB at ω_ψ, the resonant frequency of the second order numerator factor. The gain rises rapidly with a further increase in frequency to reach a maximum of 10 dB at the dutch roll frequency only to decrease rapidly thereafter. At very low input frequencies the phase lag increases gently in accordance with the spiral mode dynamics until the effect of the second order numerator term becomes apparent. The rate of change of phase is then very dramatic since the effective damping ratio of the second order numerator term is very small and negative. At the dutch roll frequency, approximately, the phase reaches −360° and the response appears to be in phase again only to

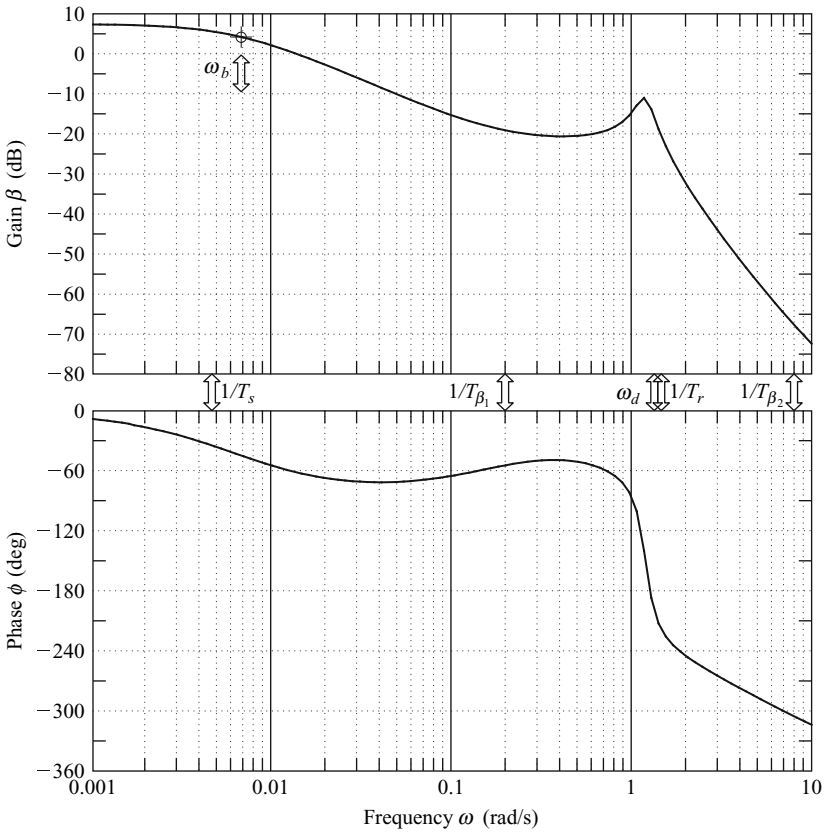


Figure 4.5-8 DC-8 sideslip angle frequency response to aileron.

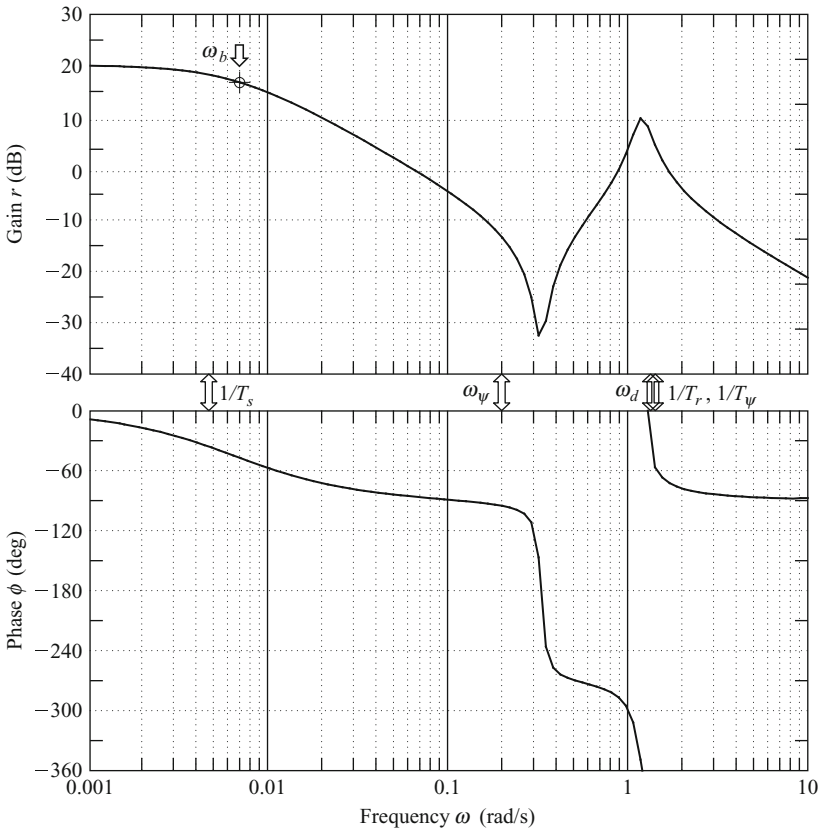


Figure 4.5-9 DC-8 yaw rate frequency response to rudder.

roll off smartly at higher frequency. Again, the effective bandwidth is a little higher than the spiral mode break frequency $1/T_s$. These unusual frequency response characteristics are easily appreciated in a flight demonstration.

If the pilot approximates a sinusoidal rudder input by pedalling gently on the rudder pedals, then at very low frequencies approaching the steady state the yaw rate response will follow the input easily and obviously, since the gain is approximately 20 dB, and with very little phase lag. As he increases the frequency of his pedalling the response will lag the input and the magnitude of the response will reduce very quickly until there is no significant observable response. If he increases the frequency of his forcing yet further, then the aircraft will spring into life again as the dutch roll frequency (resonance) is reached when the yaw rate response will be approximately in phase with the input. At higher frequencies still the response will rapidly attenuate for good. The substantial dip in both gain and phase response with frequency, caused by the second order numerator factor, effectively isolates the dutch roll mode to a small window in the frequency band. This then makes it very easy for the pilot to identify and excite the dutch roll mode by rudder pedalling. This is very good for flight demonstration but may not be so good for handling if the dutch roll damping is low and the second order numerator factor is not too close in frequency to that of the dutch roll mode.

The frequency response of roll rate p to rudder input ζ is shown in Fig. 4.5-10. This frequency response example is interesting since it represents a cross-coupling case. In the steady state, or equivalently at zero frequency, roll rate in response to a rudder input would not be expected. This is clearly evident on the gain plot where the gain is $-\infty$ dB at zero frequency. This observation is driven by the zero in the numerator which also introduces 90° of phase lead at the very lowest frequencies. This zero also very nearly cancels with the spiral mode denominator root such that at input frequencies above the spiral mode break frequency $1/T_s$ the response in both gain and phase is essentially *flat* until the effects of the remaining numerator and denominator roots come into play, all at frequencies around the dutch roll frequency. The dutch roll resonant peak in gain and the subsequent roll off in both gain and phase is absolutely classical and is easily interpreted. These frequency response observations correspond well with the response time history shown in Fig. 4.5-2 where the effects of the roll subsidence mode and the dutch roll mode are clearly visible, whilst the longer term convergence associated with the spiral mode is not visible at all. In this example bandwidth tends to lose its meaning. However, it would not be unrealistic to suggest that the usable bandwidth is a little higher

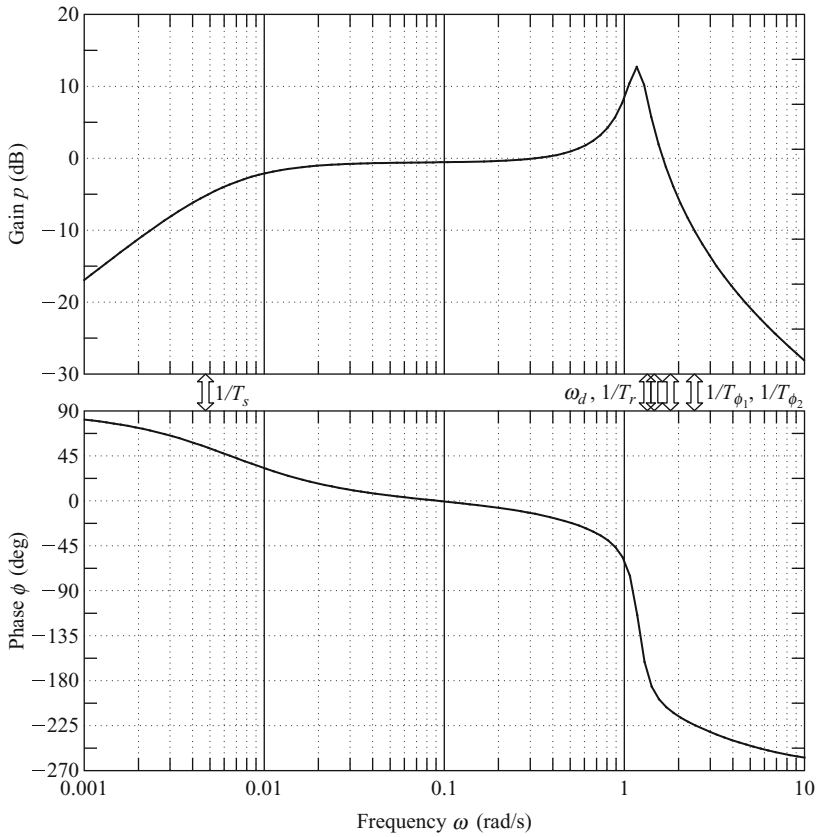


Figure 4.5-10 DC-8 roll rate frequency response to rudder.

than the dutch roll mode frequency, provided the effects at very low frequency are ignored. This then assumes that the zero numerator factor cancels with the spiral mode denominator factor to give the approximate transfer function

$$\frac{p(s)}{\zeta(s)} = \frac{0.392(s+1.85)(s-2.566)}{(s+1.329)(s^2+0.254s+1.433)} \text{ rad/s/rad (deg/s/deg)} \tag{4.5.49}$$

As before, this approximate transfer function may be interpreted in both the time domain and in the frequency domain with little loss of meaning over the usable frequency band.

4.5.5 Flying and Handling Qualities

As with longitudinal stability the lateral–directional stability characteristics of the aeroplane are critically important in the determination of its flying and handling qualities and there is no doubt that they must be correct. Traditionally the emphasis on lateral–directional flying and handling qualities has been much less than the emphasis on the longitudinal flying and handling qualities. Unlike the longitudinal flying and handling qualities the lateral–directional flying and handling qualities do not usually change significantly with flight condition, especially in the context of small perturbation modelling. So once they have been fixed by the aerodynamic design of the airframe they tend to remain more or less constant irrespective of flight condition. Any major lateral–directional departures from nominally small perturbations about trim are likely to be transient, under full pilot control and, consequently, unlikely to give rise to serious handling problems. However, this is not necessarily a safe assumption to make when considering highly augmented aircraft, a topic which is beyond the scope of the present discussion.

It is a recurrent theme in handling qualities work that short term dynamics are properly controlled by design. The typical frequencies involved in short term dynamics are similar to human pilot frequencies and their inadvertent mismatch is a sure recipe for potential handling problems. So for reasons similar to those discussed in greater detail in Section 4.4.5 referring to longitudinal dynamics, it is equally important that the lateral–directional short period stability modes be properly controlled. This may be interpreted to mean that the damping of both the roll subsidence mode and the dutch roll mode should be adequate.

The roll subsidence mode appears to the pilot as a lag in the response to control and, clearly, if the time constant should become too large roll response to control would become too sluggish. A large roll mode time constant is the direct result of low roll stability although the mode is usually stable as discussed in Section 4.5.2.1. Generally, acceptable levels of roll mode stability result in a time constant, or roll response lag which is almost imperceptible to the pilot. However, it is quite common to find aircraft in which the roll mode damping is inadequate but it is unusual to find over damped aircraft.

The spiral mode, being a long period mode, does not usually influence short term handling significantly. When it is stable and its time constant is sufficiently long it has little, or no impact on flying and handling qualities. However, when it is unstable it manifests itself as a trimming problem since the aeroplane will continually attempt to diverge laterally. When the time constant of the mode is short it is more unstable, the rate of divergence becomes faster with a corresponding increase in pilot workload. Since the mode is generally so slow to develop the motion cues associated with it may well be imperceptible to the pilot. Thus a hazardous situation may easily arise if the external visual cues available to the pilot are poor or absent altogether, such as in IMC flight conditions. It is not unknown for inexperienced pilots to become disorientated in such circumstances with the inevitable outcome! Therefore the general requirement is that, the spiral mode should preferably be stable but, since this is difficult to achieve in many aeroplanes, when it is unstable the time constant should be greater than a defined minimum.

Since the dutch roll mode is a short period mode and is the directional equivalent of the longitudinal short period mode its importance to handling is similarly critical. Generally, it is essential that the dutch roll mode is stable and that its damping is greater than a defined minimum. Similarly tight constraints are placed on the permitted range of combinations of frequency and damping. However, a level of damping lower than that of the longitudinal short period mode is permitted. This is perhaps convenient but is more likely to result from the design conflict with the spiral mode which must not have more than a limited degree of instability.

4.5.6 Mode Excitation

Unlike the longitudinal stability modes the lateral–directional stability modes usually exhibit a significant level of dynamic coupling and as a result it is more difficult to excite the modes independently for the purposes of demonstration or measurement. However, the lateral–directional stability modes may be excited selectively by the careful application of a sympathetic aileron or rudder input to the trimmed aircraft. Again, the methods developed for in-flight mode excitation reflect an intimate understanding of the dynamics involved and are generally easily adapted to the analytical environment. Because the lateral–directional stability modes usually exhibit a degree of dynamic

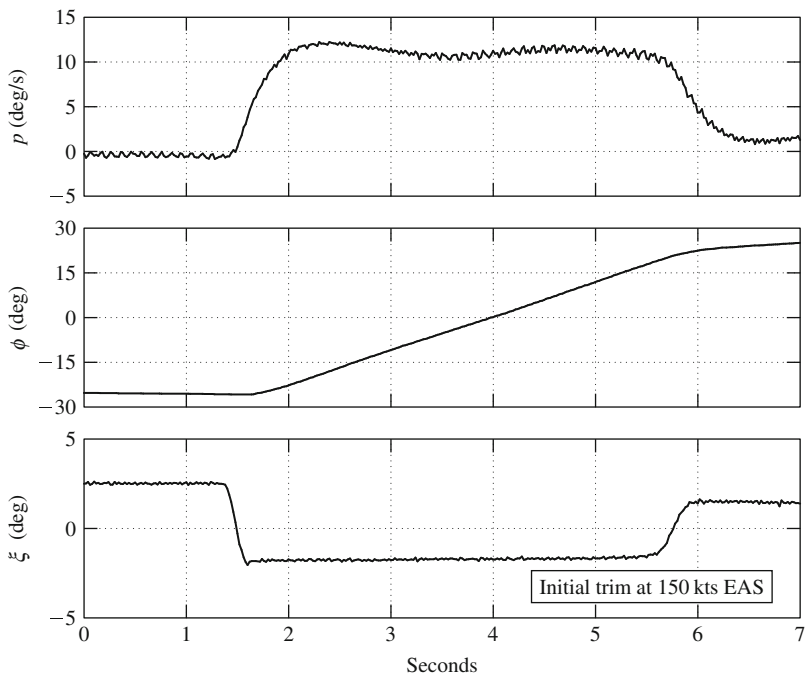


Figure 4.5-11 Flight recording of the roll subsidence mode.

coupling, the choice and shape of the disturbing input is critical to the mode under investigation. As always, standard experimental procedures have been developed in order to achieve consistency in the flight test or analytical process so that meaningful comparative studies may be made.

The roll subsidence mode may be excited by applying a short duration square pulse to the aileron, the other controls remaining fixed at their trim settings. The magnitude and duration of the pulse must be carefully chosen if the aeroplane is not to roll too rapidly through a large attitude change and thereby exceed the limit of small perturbation motion. Since the mode involves almost pure rolling motion only no significant motion coupling will be seen in the relatively short time scale of the mode. Therefore, to see the classical characteristics of the roll subsidence mode it is only necessary to observe roll response for a few seconds. An example of a roll response showing the roll subsidence mode recorded during a flight test exercise in a Handley Page Jetstream aircraft is shown in Fig. 4.5-11. The input aileron pulse is clearly seen and has a magnitude of about 4° and duratin of about 4 s. The shape of this input will have been established by the pilot by trial and error since the ideal input is very much aircraft dependent. The effect of the roll mode time constant is clearly visible since it governs the exponential rise in roll rate p as the response attempts to follow the leading edge of the input ξ. The same effect is seen again in reverse when the input is returned to its datum at the end of the pulse. The barely perceptible oscillation in roll rate during the "steady part" of the response is, in fact, due to a small degree of coupling with the dutch roll mode.

In order to conduct the flight experiment without large excursions in roll attitude ϕ it is usual to first establish the aircraft in a steady turn with, in this illustration, $-30°$ of roll attitude. On application of the input pulse the aircraft rolls steadily through to $+30°$ of roll attitude when the motion is terminated by returning the aileron to datum. This is also clearly visible in Fig. 4.5-11. The effect of the roll mode time constant on the roll attitude response is to smooth the entry to, and exit from the steady part of the response. Since the roll mode time constant is small, around 0.4 s for the Jetstream, its effect is only just visible in the roll attitude response. It is interesting to observe that the steady part of the roll response is achieved when the moment due to the damping in roll becomes established at a value equal and opposite to the disturbing moment in roll caused by the aileron deflection. Clearly, therefore, the roll subsidence mode governs the transient entry to, and exit from all rolling motion.

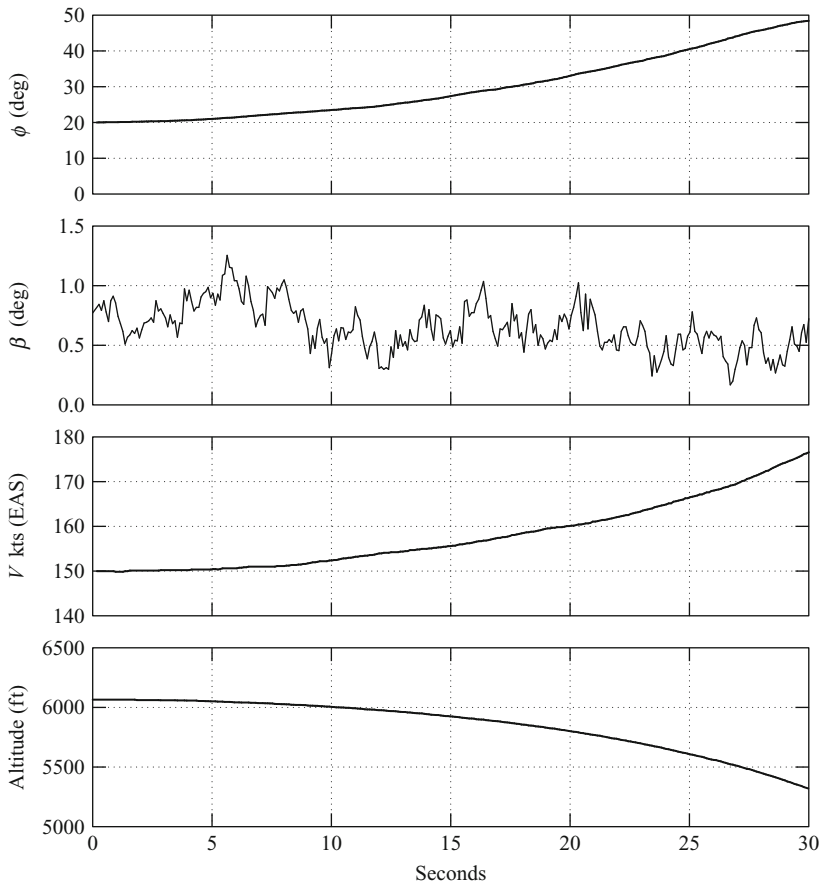


Figure 4.5-12 Flight recording of the spiral mode departure.

The spiral mode may be excited by applying a small step input to rudder ζ, the remaining controls being held at their trim settings. The aeroplane responds by starting to turn, the wing on the inside of the turn starts to drop and sideslip develops in the direction of the turn. When the roll attitude has reached about 20° the rudder is gently returned to datum and the aeroplane left to its own devices. When the spiral mode is stable the aeroplane will slowly recover wings level flight, the recovery being exponential with spiral mode time constant. When the mode is unstable the coupled roll-yaw-sideslip departure will continue to develop exponentially with spiral mode time constant. An example of an unstable spiral mode, captured from the time the disturbing rudder input is returned gently to datum, and recorded during a flight test exercise in a Handley Page Jetstream aircraft is shown in Fig. 4.5-12. The slow exponential divergence is clearly visible in all recorded variables, with the possible exception of sideslip angle β which is rather noisy. In any event the magnitude of sideslip would normally be limited to a small value by the weathercock effect of the fin. Although speed and altitude play no part in determining the characteristic of the mode, the exponential departure in these variables is a classical, and very visible, consequence of an unstable spiral mode. Once excited, since the aircraft is no longer in wings level flight, lift is insufficient to maintain altitude and so an accelerating descent follows and the spiral flight path is determined by the aeromechanics of the mode. The first 30 s of the descent is shown in Fig. 4.5-12. Obviously, the departure must be terminated after a short time if the safety of the aeroplane and its occupants is not to be jeopardised.

Ideally, the dutch roll mode may be excited by applying a doublet to the rudder pedals with a period matched to that of the mode, all other controls remaining at their trim settings. In practice the pilot pedals continuously and cyclically on the rudder pedal and by adjusting the frequency it is easy to find the resonant condition. See the related comments in Example 4.5.3 and note that the dutch roll frequency is comfortably within the human bandwidth. In this manner a forced oscillation may easily be sustained. On ceasing the forcing input the free transient characteristics of the dutch roll mode may be seen. This free response is shown in the flight recording in Fig. 4.5-13 which was made in a Handley Page Jetstream aircraft. The rudder input ζ shows the final doublet before ceasing the forcing at about 5 s, the obvious oscillatory rudder motion after 5 s is due to the cyclic aerodynamic load on the free rudder.

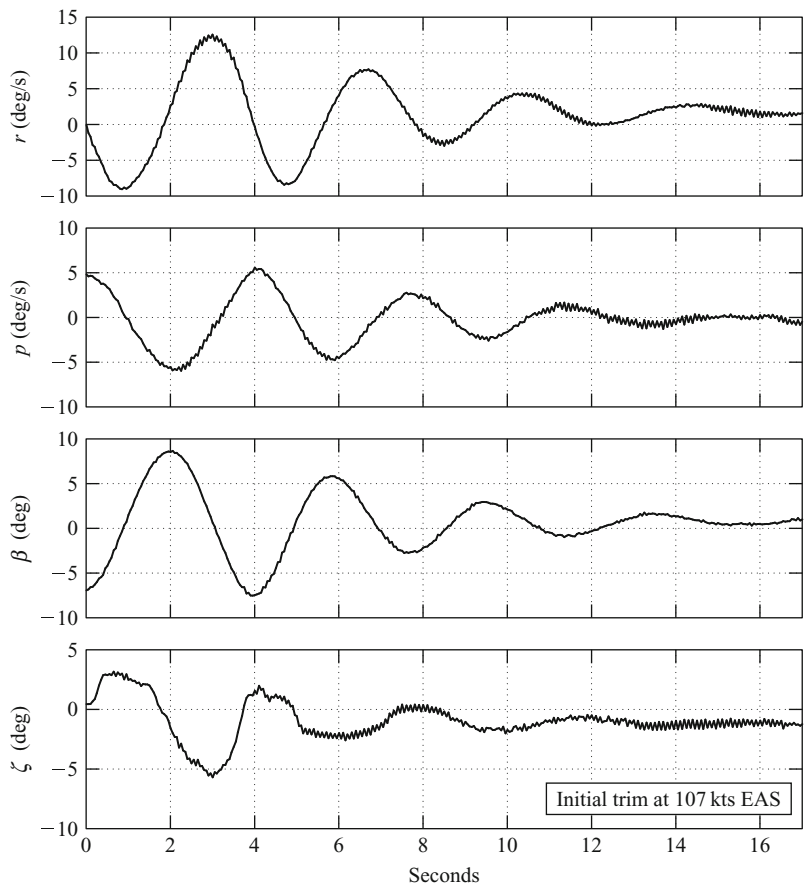


Figure 4.5-13 Flight recording of the dutch roll mode.

The classical damped oscillatory motion is clearly visible in the variables shown, yaw rate r, roll rate p and sideslip angle β. The motion would also be clearly evident in both roll and yaw attitude variables which are not shown. Note the relative magnitudes of, and the phase shift between yaw rate r and roll rate p, observations which are consistent with the classical physical explanation of the mode dynamics.

As for the longitudinal modes discussed in Section 4.4.6 the above flight recordings of the lateral–directional stability modes illustrate the *controls free* dynamic stability characteristics. The same exercise could be repeated with the controls held fixed following the disturbing input. Obviously, in this event the *controls fixed* dynamic stability characteristics would be observed and, in general, the differences between the responses would be small. To reiterate the important comments made in Section 4.4.6, controls free dynamic response is only possible in aeroplanes with reversible controls which includes most small classical aeroplanes. Virtually all larger modern aircraft have powered controls, driven by electronic flight control systems, which are effectively irreversible and which means that they are only capable of exhibiting controls fixed dynamic response. Thus, today, most theoretical modelling and analysis is concerned with controls fixed dynamics only, as is the case throughout this book. However, a discussion of the differences between controls fixed and controls free aeroplane dynamics may be found in Hancock (1995).

When it is required to investigate the dynamics of a single mode in isolation analytically, the best approach is to emulate flight test practice as far as that is possible. It is necessary to choose the most appropriate transfer functions to show the dominant response variables in the mode of interest. For example, the roll subsidence mode may only be observed sensibly in the dominant response variable p and, to a lesser extent, in ϕ. Similarly for the spiral and dutch roll modes, it is important to observe the motion in those variables which are dominant, and hence most visible in the mode dynamics. It is also essential to apply a control input disturbance sympathetic to the mode dynamics and it is essential to observe the response for an appropriate period of time. Otherwise the dynamics of interest will inevitably be obscured by motion coupling effects. For example, Fig. 4.5-11 shows both the roll subsidence mode and the dutch roll mode but, the excitation, choice of output variables and time scale were chosen to optimise the recording of the roll subsidence mode. The form of the control input is not usually difficult to arrange in analytical work since most software packages have built-in impulse, step and pulse functions, whilst more esoteric functions can usually be programmed by the user. For the analysis of the lateral–directional mode dynamics especially, this kind of informed approach is critically important if the best possible visualisation of the modes and their associated dynamics are to be obtained.

References

Hancock, G.J. 1995: *An Introduction to the Flight Dynamics of Rigid Aeroplanes*. Ellis Horwood Ltd., Hemel Hempstead.
Teper, G.L. 1969: *Aircraft Stability and Control Data*. Systems Technology, Inc., STI Technical Report 176-1. NASA Contractor Report, National Aeronautics and Space Administration, Washington D.C. 20546.

Problems

1. Describe the possible modes of lateral–directional motion of an aircraft when disturbed slightly from steady flight.

   An aircraft in steady horizontal flight is disturbed slightly in the lateral plane. If the inertia forces associated with the angular accelerations in the resulting motion are neglected, as well as the components of the acceleration and aerodynamic forces along the oy axis, show that the resulting motion is either a divergence or a subsidence depending in general on the sign of $(L_v N_r - L_r, -N_v)$. Describe how the stability of an aircraft in this mode will change with increase of fin size. (CU 1979)

2. A transport aircraft whose wing span is 35.8 m is flying at 262 kt at an altitude where the lateral relative density parameter $\mu_2 = 24.4$. The dimensionless controls fixed lateral–directional characteristic equation is

   $$\lambda^4 + 5.8\lambda^3 + 20.3\lambda^2 + 79.0\lambda + 0.37 = 0$$

   (i) What can be deduced about the lateral–directional stability of the aircraft from inspection of the characteristic equation?
   (ii) Solve the characteristic equation approximately; determine estimates for the time constants of the non-oscillatory modes and the frequency and damping ratio of the oscillatory mode.
   (iii) Comment on the acceptability of this aircraft. (CU 1980)

3. (i) What is the lateral–directional weathercock stability of an aircraft?
   (ii) State the main aerodynamic contributions to weathercock stability. (CU 1982)

4. The Navion is a small light aeroplane of conventional layout and in a low speed level flight condition the coefficients of the dimensionless lateral–directional stability quartic are given by

$$\lambda^4 + B_2\lambda^3 + C_2\lambda^2 + D_2\lambda + E_2 = 0$$

where

$$B_2 = 20.889$$
$$C_2 = 46.714 - k_v$$
$$D_2 = 115.120 - 18.636k_v$$
$$E_2 = 55.570 + 1.994k_v$$
$$\text{and} \quad k_v = -\frac{\mu_2 N_v}{i_z}$$

The lateral relative density parameter $\mu_2 = 11.937$, and the dimensionless moment of inertia in yaw $i_z = 0.037$. The quartic factorises to

$$(\lambda + B_2)\left(\lambda + \frac{E_2}{D^2}\right)(\lambda^2 + k_1\lambda + k_2) = 0$$

Show that if the fin were made too large the aircraft would become dynamically unstable. What would happen to the aircraft if a critical value were exceeded? (CU 1983)

5. Describe and explain the physical characteristics of the roll subsidence stability mode.

Assuming the motion associated with the mode comprises pure rolling only write down the equation of motion assuming the rudder to be fixed ($\zeta = 0$). By taking the Laplace transform of this equation show that the roll control transfer function is given by

$$\frac{p(s)}{\xi(s)} = \frac{-k}{(1 + sT_r)}$$

where $k = \mathring{L}_\xi/\mathring{L}_p$ and $T_r = -I_x/\mathring{L}_p$. State any assumptions made in obtaining the transfer function.

Obtain the inverse Laplace transform of the transfer function to show that the roll rate response to a unit step of aileron is given by

$$p(t) = -k(1 - e^{\frac{t}{T_r}})$$

The Republic F-105 Thunderchief aircraft has a wing span of 10.4 m and moment of inertia in roll of 13965 kg m^2. In a cruise flight condition at Mach 0.9 at an altitude of 35,000 ft, the dimensionless derivatives have the following values, $L_p = -0.191$ and $L_\xi = -0.029$. Sketch the roll rate response to a 1° step of aileron deflection and comment on the roll handling of the aircraft. (CU 1986)

6. The aircraft described below is flying at a true airspeed of 150 m/s at sea level. At this flight condition the aircraft is required to have a steady roll rate of 60 deg/s, when each aileron is deflected through 10 deg. Assuming that the outboard edges of the ailerons are at the wing tip, calculate the required aileron span. If the ailerons produce 17,500 N m of adverse yawing moment, calculate the rudder deflection required for trim.

Aircraft data:

Rectangular unswept wing		Fin	
	Span = 15 m		Area = 3 m^2
	Area = 27 m^2		Moment arm from cg = 6 m
	$L_p = -0.2$	Rudder	Area = 1.2 m^2
Aileron	$dC_l/d\xi = 2$ 1/rad		$dC_l/d\zeta = 2.3$ 1/rad

(LU 2001)

7. Using a simple model show that the time to half amplitude of the roll subsidence mode may be approximated by

$$t_{1/2} - \frac{I_x}{\mathring{L}_p}\ln(2)$$

Given that the rolling moment due to roll rate derivative may be written as

$$\mathring{L}_p = -\rho V_0 \int_0^s \left(C_D + \frac{dC_L}{d\alpha}\right)_y c_y y^2 dy$$

Determine the time to half amplitude of the roll subsidence mode for an aircraft with the following characteristics, when it is flying at sea level at 100 m/s.

Wing span = 10 m	$dC_L/d\alpha$ at root = 5.7 1/rad
Wing root chord = 1.5 m	$dC_L/d\alpha$ at tip = 5.7 1/rad
Wing tip chord = 0.75 m	$C_D = 0.005$ (constant)
Inertia in roll = 8000 kg m^2	

Assume $dC_L/d\alpha$ varies linearly along the span. (LU 2002)

8. For the aircraft described below, determine the value of wing dihedral required to make the spiral mode neutrally stable. The rolling moment due to sideslip derivative is given by

$$L_v = -\frac{1}{Ss}\int_0^s c_y a_y \Gamma y \, dy$$

and the time to half (double) amplitude for the spiral mode is given by

$$t_{1/2} = \frac{V_0}{g}\left(\frac{L_v N_p - L_p N_v}{L_v N_r - L_r N_v}\right)\ln(2)$$

Wing area	$S = 52$ m^2
Wing span	$B = 14.8$ m
Wing root chord	5.0 m
Wing tip chord	2.0 m
Fin area	$S_F = 8.4$ m^2
Fin roll arm	$h_F = 1.8$ m
Wing lift-curve slope	$a_y = 3.84$ 1/rad
Fin lift curve slope	$a_{1F} = 2.2$ 1/rad
L_r	−0.120
N_r	−0.120
N_v	0.158

Aircraft data:

Discuss how the geometry of the wing and fin influence the stability of the spiral mode. (LU 2003)

This page intentionally left blank

SECTION 5

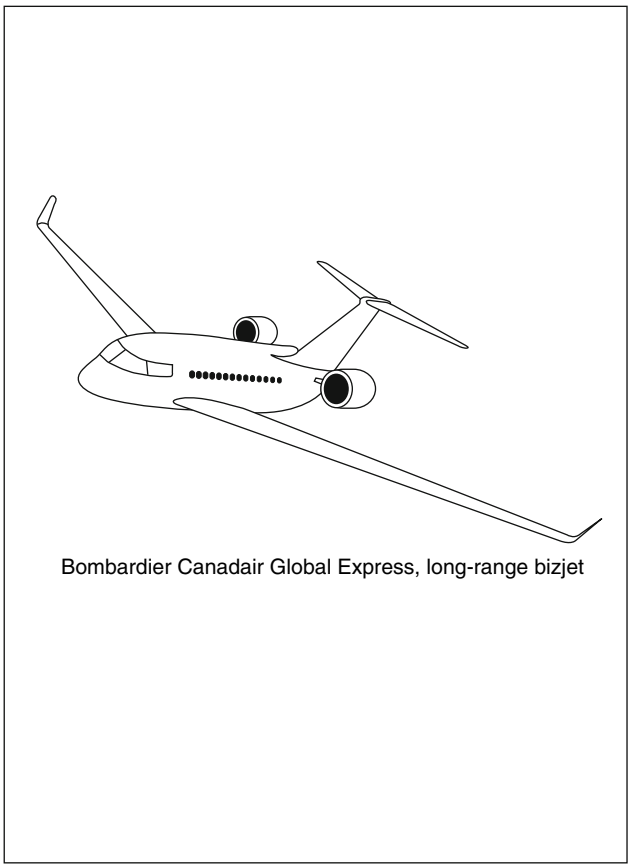
Aircraft Design Projects

This page intentionally left blank

Aerospace Engineering Desk Reference; ISBN 9781856175753

5.1 Long-Range Business Jet

Lloyd Jenkinson and Jim Marchman



Bombardier Canadair Global Express, long-range bizjet

5.1.1 Introduction

Up to the events of 11 September 2001, all of the professional aeronautical industry market analysts predicted that scheduled airline business over the next 20 years was likely to increase at an annual growth rate of between 3.5 and 5 per cent. Such unexpected and tragic events illustrate the vulnerability of airline market projections to influences outside the control of the industry. However, it is expected that after a period of industrial recession the previous projections will be resumed. Although this is welcome news for the aircraft manufacturers and airlines, as more passenger movements equate to a growth in business, it also means that existing airports and associated infrastructures will become increasingly inadequate to satisfy this expansion. Already many of the world's international airports are working beyond capacity at peak operating periods. The expected doubling of demand over the next 15 to 20 years is generally incompatible with the planning approval and building timescales for airport expansion. The political, social and economic factors that accompany airport building projects lie outside the control of the aeronautical industries. In the past, planning enquires and environmental pressure groups have delayed many of the proposed airport development projects. There is no evidence that this situation will improve in the future.

Some of the problems at airports may improve when the new, supercapacity aircraft are introduced but even this development will not solve the passenger capacity problems at airports. Moving airline operations to larger aircraft is not new. Most airlines now use larger capacity aircraft on services that smaller types satisfied a few years earlier. This trend is likely to continue. This development allows an increase in passenger movements without increasing aircraft movements (i.e. increases passengers per flight 'slot'). However, this practice does not solve the problems of

increased passenger demand on the airport terminal facilities. Handling larger aircraft and greater numbers of passengers requires an associated expansion of airport infrastructure.

Analysis shows that although the main airports are working at full capacity, over 70 per cent of all aircraft movements involve relatively small aircraft. These aircraft do not need the service provided at the large airports (i.e. runway length for take-off and landing, and terminal lounge capacity). Many of these flights are related to regional 'feeder' services that provide linking flights to international scheduled services. The mixture of small and large capacity services at airports leads to an inefficient use of the facilities available. This inefficiency is the source of many of the delays and disruption currently endemic at large airports.

Business surveys show that delays at airports will increase as demand on the services increases in the future. Delays and disruptions in the service affect all passengers. Airlines provide exclusive facilities at airports for their business travellers but this does not pacify a customer who misses an important meeting because of a flight delay. Such passengers demand more certainty in their travel arrangements than can be provided by the current and future operations. An expensive alternative to the current situation is for the business traveller to use a small, exclusive business jet for the journey but this may not be within the budget of most commercial travellers.

It has been suggested by researchers that the current problems at large airports could be eased if the feeder services were transferred to satellite airports. Such developments would potentially increase the capacity provision at the larger airport without the need to make changes to the present runway or terminal facility. However, as the traveller will need to transfer to and from the large airport there will be a requirement to provide or improve the ground transport provision between the two airports. This type of development is slowly taking place at the main 'hub' airports. The downside to this scheme is that the traveller is then subjected to extra potential delays from congestion at both airports and the ground interchange.

An essential element to any airline's success is the ability to attract the 'business' traveller. Business travellers pay significantly higher prices for travel than tourist-class passengers and represent a more dependable source of income than travellers who opt for first class. Capturing the loyalty of the business traveller is high on the agenda of every major airline. This is often accomplished on longer-range flights, where there is sufficient space, by creating a separate 'business-class' cabin. In this the seat widths, seat pitch and cabin amenities are set between those of the first-class and tourist-class sections. Business-class passengers are allowed early boarding and a wider choice of in-flight movies and passenger services, etc. Business travellers, of course, pay for these advantages but ticket-pricing schemes are devised in which the business-class ticket costs little more than the 'list' price of one in economy class. And, unless one wishes to purchase his or her ticket a couple of weeks in advance of the flight and is willing to stay at his or her destination over the weekend, the 'list' price is the best available. These special business-class amenities are usually provided only on longer-range flights since it is assumed that on trips of a couple of hours or less the benefits of such service are questionable.

There have been a few attempts to create specialised, all business-class airlines using aircraft such as the Boeing 727 or 737 fitted with only about half to two-thirds the usual number of seats. However, these and similar aircraft like the Airbus 320 class are usually limited in range and are unsuited to the long transcontinental or international flights where business-class amenities can really make a difference to the target group of travellers. So far, these attempts have failed to attract enough customers to make a profit. This may be due to insufficient perceived advantage in the wider seats and better service against the higher price on shorter-range flights. Alternatively, it may be due to the insufficient flight frequency of the special services compared to flight schedules of the existing airlines.

Is there a market for a business-class only aircraft or airline? For success, it must meet the preferences of the business traveller, which include the following amenities:

- Larger, more comfortable seats with more leg room than those in tourist class.
- Pemium in-flight service (better meals, free drinks, more selection of movies and a wider choice of entertainment options).
- Separation from tourist-class passengers in airport lounges during boarding, and on board the aircraft (for mixed-class operations).
- Faster flight check-in and post-flight luggage retrieval.
- Direct flights without delays at airports, especially on longer journeys.

The first two of the above points are currently available to business-class travellers on larger, longer-range flights with most airlines. The next two preferences can only be achieved with special 'business-class only' flights. In addition, the last of these may require a reduced dependence on hub airports.

Most airlines depend on a mix of passenger classes and fare levels to operate profitably. First-class passengers or their company pay dearly for their extra comfort and amenities. However, on many flights, the first-class seats are filled with business- or tourist-class passengers who have used accumulated frequent flyer miles to 'upgrade' their

seats. At the other end of the airplane, the tourist-class passenger may have paid anything from a couple of hundred dollars to over two thousand dollars for a transoceanic or transcontinental flight. The 'list' price for tourist class is often very near that for the discounted business class. This serves as an inducement for passengers who cannot meet the requirements for discount tickets (typically 14 to 17 days' advance purchase and travel which includes a weekend stay) to purchase business-class tickets. Those who can meet the discount requirements can often fly for very low cost. Airlines use sophisticated seat management software to optimise the price of each ticket; seats in 'economy' are sold at a different price depending on availability and demand. The goal is to fill every seat, and having a 90 per cent discount passenger is better than an empty seat. Like soft fruits, scheduled airline seats are perishable goods that must be sold before the 'shop shuts' or aircraft departs!

There is a question as to whether an 'all business-class' airline can fill enough seats on enough flights to make a profit. If not, it must become like every other airline and offer either managed discounts or some other 'class' of seating with a scheme to fill these in order to make each flight at least break even in cost. Such an airline would also have to come close to matching the flight frequency of regular airlines and it may need to offer at least slightly lower ticket prices to induce business travellers away from their frequent flyer club loyalties. Success would probably also require a new class of airliner which could fly transcontinental and transocean ranges with passenger capacities similar to today's B-737s and A-320 class aircraft. Most existing airliners are designed as either a long-range/large-capacity, or a short-range/small-capacity operation, compatible with the current spoke and hub system of flight routing.

Meeting the last of the comments above will also require either a departure from the hub and spoke route model, or the use of aircraft designed to fly faster. The ideal airliner for this goal is probably a B-737, A-320 size aircraft with a range of 7000 nm and a cruise speed of Mach 0.9 or higher.

Any departure from the hub and spoke system may prove problematic given the saturated state of hub airports from which longer-range flights generally operate. A solution may be found in the use of other airports. Perhaps former metropolitan airports, which are currently used primarily for private and corporate aircraft operations, could be used. This would require the aircraft to be able to take off and land on shorter runways. It would also require the construction of high-speed ground transportation systems which could move business travellers between airports. This would further need the provision of gate-to-gate transfer of passenger and baggage without requiring additional baggage or security checks.

5.1.2 Project Brief

From the analysis of existing and future air travel conditions above, it is possible to postulate a new type of airline service; one that is aimed at the profitable business travel market. This project study involves the development of a new scheduled business-exclusive international service from smaller airports.

An initial survey of the location of airports in the developed world was undertaken. This showed that within a radius of 50 miles around most current international airports, there is at least one regional airport that could be used for such a service. However, this may require the establishment of facilities to deal with international flights.

For existing airlines such a service would provide an improved and exclusive premium-class service and would allow an expansion of economy-class business at existing busy airports. New airlines may be set up to exploit the perceived market opportunity. Over the past decade, with the relaxed 'deregulatory' airline service, several new airlines have evolved to provide quick, easy and cheap (bus-type) alternative scheduled services. Some of these failed to achieve profitable operation but a significant number survived to compete with the older and larger established airlines. Such developments show that the airline business is dynamic enough to respond to novel market opportunities. A new aircraft type would create a unique, convenient, exclusive high-class business service that would compete with the current business-class sections in existing mixed-class scheduled services.

5.1.2.1 Project Requirements

The following design requirements and research studies are set for the project:

- To design an aircraft that will transport 80 business-class passengers and their associated baggage over a design range of 7000 nm at a cruise speed equal or better than existing competitive services.
- To provide the passengers with equivalent, or preferably better, comfort and service levels to those currently provided for business travellers in mixed-class operations.
- To operate from regional airports.
- To use advanced technologies to reduce operating costs.
- To offer a unique and competitive service to existing scheduled operations.
- To investigate alternative roles for the aircraft.

Table 5.1-1

	PAX	Range (nm)	Field length (m)
Small aircraft			
728Jet	70	1430	1676
CRJ 700	70	1690	1880
F 70	79	1870	1583
928Jet	95	1900	1950
RJ 100	112	2090	1275
F 100	107	1680	1715
B717-200	106	1460	1480
A318-100	107	2350	1400
Medium aircraft			
A310-300	218	5180	2300
B767-200	186	3000	2100
B767-300	245	3460	2550
A300-600R	266	4160	2280
Large aircraft			
A340-500	313	8504	3050
B777-200ER	305	7775	3020
MD11	285	6910	3110
B747-400	416	6177	3020
A380	555	7676	2900

- To assess the development potential in the primary role of the aircraft.
- To produce a commercial analysis of the aircraft project.

5.1.3 Project Analysis

Project analysis will consider, in detail, each of the design requirements described in the previous section.

5.1.3.1 Payload/Range

With only 80 seats, the aircraft is considered as a small aircraft in commercial transport operations. This size of aircraft is normally used only on short-range, regional routes. Table 5.1-1 (data from a *Flight International* survey of world airliners) shows the existing relationship between aircraft size (number of passengers (PAX)), design range and field capability.

By considering the range requirement of 7000 nm, the new aircraft falls into the large aircraft category but the passenger capacity of only 80 defines it as a small aircraft. This contradiction defines the unique performance of the new aircraft. The closest comparison to the specification is with the corporate business jet. However, this type of aircraft has a much smaller capacity (usually up to a maximum of 20 seats).

To investigate the significance of the 7000 nm range requirement, an analysis of the 50 busiest international airports was undertaken. This compared the great circle distances between airport pairs. It showed that very few scheduled routes exceeded 7000 nm. The list below shows the exceptions:

US east coast – Sydney, Singapore, Thailand
US central – Sydney, Thailand
US west coast – Singapore
Europe central – Sydney

All of the above routes could be flown with a refuelling stop at Honolulu for the US flights and Asia for the European flights. This analysis showed that 7000 nm was a reasonable initial assumption. This distance could be reduced if the design was shown, in subsequent trade-off studies, to be too sensitive to the range specification.

5.1.3.2 Passenger Comfort

Long-range flights obviously equate to long duration. At an average speed of 500 kt, the 7000 nm journey will take 14 hours. Increasing the speed by only 5 per cent (e.g. from M0.84 to M0.88) will reduce this by 45 minutes. Anyone who has travelled on a long flight will agree that this reduction would be very welcome. Business travellers may accept

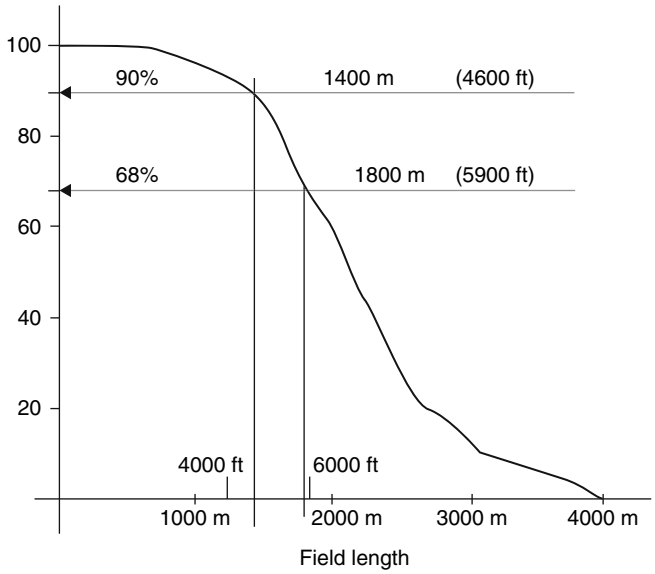


Figure 5.1-1 Runway length survey.

a premium on the fare for such a saving in time and discomfort. As flight duration and comfort are interrelated, it is desirable to provide a high cruise speed for long-range operations. Reduced aircraft block time will also provide an advantage in the aircraft direct operating costs (DOCs) providing that extra fuel is not required for the flight.

As journey time relates directly to perceived comfort level, airlines have traditionally provided more space for business-class travellers. In the highly competitive air transport industry many other facilities and inducements have been used to attract this high value sector of the market. A new aircraft design will need to anticipate this practice and offer, at least, equivalent standards. This will impact directly on the design of the aircraft fuselage and in the provision of cabin services and associated systems.

5.1.3.3 Field Requirements

The requirement to operate from regional airports effectively dictates the aircraft maximum take-off and landing performance (see Table 5.1-1). Operation from smaller airports will also affect the aircraft compatibility to the available airport facilities.

To understand this in more detail a survey is required to determine the available runway length of regional airports. Such a survey was undertaken in an aircraft design study[1] for a conventional feederliner. Figure 5.1-1 shows the results of this survey.

The frequency distribution of major European, regional airport, runway lengths (mostly UK, France and Germany) indicates that the 90 percentile equates to a minimum field length of 1400 m (4600 ft). Many of the aircraft operating within this field requirement are general aviation types. For an aircraft of 80 or more seats, this short distance may be regarded as too demanding on the aircraft design. It would force the wing to be too large or require a complex flap system. Both, or either, of these would increase drag in the cruise phase and thereby the aircraft DOCs. A sensitivity study on this aspect of the design could be conducted later in the design process when more details of the aircraft are available. Increasing the field length to 1800 m (5900 ft) will allow operation from 70 per cent of the airports surveyed. Comparing this choice to current, regional aircraft characteristics shows it is equivalent to the Avro RJ, Fokker and Boeing types. For this reason, the longer (1800 m) length will be specified for the design.

5.1.3.4 Technology Assessments

The requirement to incorporate advanced technology into the design raises several questions relating to commercial risk, technical viability and economics. A design study that included a detailed assessment of new technologies applied to regional aircraft was presented by a Virginia Tech (VT) team at an AIAA meeting in 1995.[2] This considered some emerging technologies in propulsion, aerodynamics, materials and systems. In the final configuration of their aircraft, they selected ducted-direct-drive prop-fans as the powerplant. This showed substantial fuel saving over normal, high-bypass turbofans. They accepted the relatively slow cruise speed (M0.7) because their specification only called for a 3000 nm range. As much of the flight duration on short stage distances is spent in climb and decent,

a reduced cruise speed is not too critical. For our design, such a slow cruise speed would not be acceptable, as it would significantly compromise the performance (flight duration) against existing scheduled services. For this reason, the prop-fan engine is not a suitable choice for our aircraft. A conventional high-bypass turbofan engine that is already certified and in use on other aircraft types will be our preferred choice. Although this will not show the fuel savings identified in the VT study, it will be comparable to the competitive aircraft. In addition, adopting a fully developed engine will reduce commercial risk and lower DOCs.

From an aerodynamic standpoint, the VT study proposed the incorporation of natural laminar flow aerofoil sections with boundary layer suction on the upper leading edge profile. Research results from NASA Langley were quoted to validate this approach. The hybrid laminar flow control system was shown to reduce aircraft drag and therefore fuel consumption. The study proposed the use of wing tip vortex turbines to power the boundary layer suction system. As such devices have not been developed in the time since the report was published, it is not considered wise to adopt this concept for our design. This will leave the wing tips clear for winglets to reduce induced drag in cruise. These are now well established on many long-range aircraft; therefore, the technology is well understood. Boundary layer suction will need to be provided from bleeds from the engines. Later in the design process, a study will need to be undertaken to determine the effectiveness of the laminar flow system against the reduction in engine thrust in cruise caused by the demand from the air bleed system. On the turbulent flow parts of the aerofoil, it is proposed to incorporate the surface striation researched by Airbus and NASA in the late 1990s.

The use of new materials in the construction of civil aircraft is now becoming commonplace. To continue this trend composite materials will be used for wing skins, control surfaces, bulkheads and access panels. Advanced metallic materials will be used in high load areas (landing gear, flap mechanisms, engine and wing attachment structures). As proposed in the VT study, micro-perforated titanium, wing-leading-edge skins will be used for the boundary layer suction structure. A conventional, aluminium-alloy, fuselage pressure shell will be proposed as this is well proven and adds confidence to the aircraft structural framework. Filament wound composite structures may offer mass reductions for the pressure cabin but this technology is still unproven in airliner manufacture, so it will not be used on our aircraft.

Aircraft systems will follow current technology trends. This will include a modern flight deck arrangement. Aircraft system demand will increase due to the improvement in provision for the passenger services and comfort. This will include better air conditioning in the cabin to provide an increase in the percentage of fresh air feed into the system, more electronic in-flight passenger services and business (computing and communication) facilities. The aircraft will be neutrally stabilised to reduce trim drag in cruise and therefore require redundancy in flight control systems.

5.1.3.5 Marketing

Our aircraft type lies between the conventional mixed-class scheduled service and the exclusive corporate jet. The aircraft and operator will be offering a unique service. A comparison to the old 'Pullman'-class service operated by the railways at the beginning of the last century is appropriate. Avoiding major airports and the associated, and increasing, congestion and delays will be a significant feature of the service. Segmentation of the premium ticket passengers away from the low-cost travellers will be another positive marketing feature. Providing commercial/office facilities and a quieter environment during the flight will be another improvement over the existing mixed-class operations. All of these advantages will need to be set against the premium fare that the service will need to charge to offset the higher cost of operating the aircraft compared to existing services. In an analysis of the pricing policy of the new service it may be difficult to assess the elasticity of the ticket price because the service is new and untried. In the past, a sector of the travelling public has been attracted to the Concorde service. The reason that the extra ticket price was accepted is not clear. Either the time saving from supersonic flight or the exclusivity of the service, or both, may have been the feature that the customer was attracted to. It is felt that a premium above the existing business-class fare of 30 per cent is probably the limit of acceptance by the market sector. At this stage in the development of the project, this is only a 'guesstimate'. Market research would be necessary to identify the exact premium. A more in-depth market analysis will be needed before confidence in this figure is possible. There will always be a number of people who would use such a service. But as the ticket price rises, this number reduces. The number of passengers willing to pay the extra price must be seen to be greater than the number required to make the service commercially viable. The price at which companies regard the airfare as excessive must be determined.

5.1.3.6 Alternative Roles

Developing an aircraft exclusively for a specialised role in civil aviation would be regarded as commercial madness. All aircraft projects should consider other roles the aircraft may fulfil. Our aircraft will have a fuselage size that is more spacious than normally associated with an 80-seat airliner. The long-range requirement will demand a high fuel load and this will make the aircraft maximum design weight heavier than normal for 80-seat aircraft. Both of these aspects

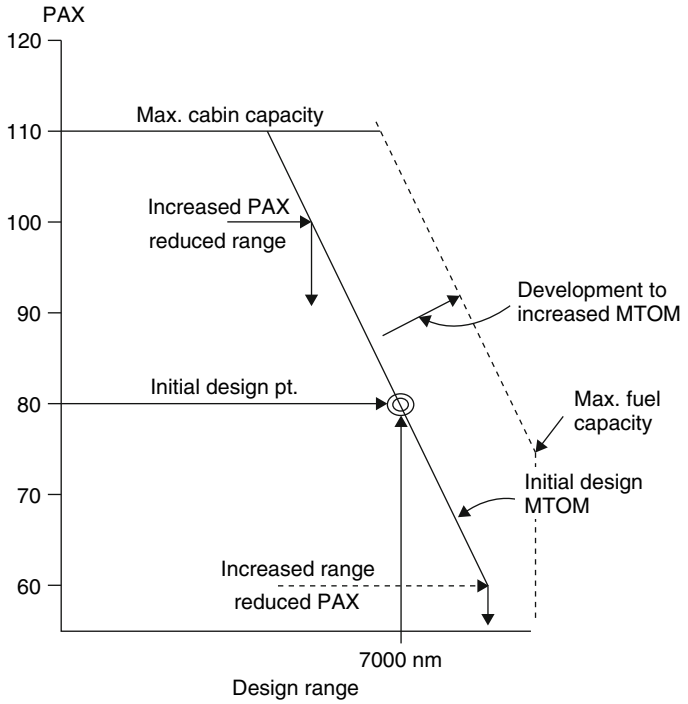


Figure 5.1-2 Aircraft development (payload/range) options.

suggest that the aircraft could be transformed into a conventional higher capacity, shorter-range airliner. A study will be required to investigate such variants. This type of investigation may result in recommendations to change the baseline aircraft geometry to make such developments easier to achieve. For example, increasing the fuselage diameter may allow a change from five to six abreast seating in the higher capacity aircraft to be made. Without such a change, six abreast seating may be unfeasible.

Other variants of the aircraft could be envisaged for military use. The long-range and small field features of the design are compatible with troop and light equipment transport operations. The ability to move military personnel without the need to refuel would avoid some diplomatic problems that have arisen in the past. The long endurance feature would make the aircraft suitable for maritime patrol, reconnaissance, surveillance and communication roles. The military variants should not be considered in the design of the baseline aircraft, as this would unduly complicate the conceptual design process. Such considerations should be left until the current design specification is better realised.

5.1.3.7 Aircraft Developments

All aircraft projects must consider future development strategies to avoid complicated and expensive modifications in the development process. In modern civil aircraft design, it is common practice to consider the aircraft type as a 'family'. Airbus and Boeing use this approach successfully in their product lines. Stretching, and in some cases shrinking, the original design is now normal development practice. All new aircraft projects consider this in their definition of the initial design. It is essential to consider the consequences of this approach in the conceptual design phase. In this way, constraints to the development of the aircraft are reduced.

Apart from making geometrical changes around the initial, maximum design mass, it is common to expect a growth in this limit over the lifetime of the aircraft type.

Typically a 35 per cent growth in max. take-off mass may be expected over the lifetime of the type. Figure 5.1-2 shows how such developments are planned. The payload (PAX) range (nm) diagram shows the initial design specification of the aircraft. The sloping maximum design mass line shows the initial layout options (trading passengers for range and vice versa). The dashed line represents a developed higher mass aircraft. This shows the growth (PAX and range) potential for a maximum take-off mass (MTOM) increase. Such investigations are required in the early conceptual design phase to guide the aircraft development path. It may be found necessary to slightly compromise the best layout of the initial aircraft to provide for such developments.

Table 5.1-2

	PAX	Range (nm)	M_{TO} (kg, lb)	*T/W*	*W/S* (kg/sq.m, lb/sq.ft)	M_E/M_{TO}
Business jets						
Falcon 2000	19	3000	15 875, 35 000	0.327	323, 66.3	0.563
Gulfstream V	14	6500	40 370, 89 000	0.332	382, 78.3	0.526
Learjet 45	10	2200	8 845, 19 500	0.359	359, 73.6	0.600
Canadair RJER	50	2270	23 133, 48 800	–	478, 98.1	0.591
Beechcft 400A	8	1690	7 303, 16 100	0.360	326, 66.8	0.624
Hawker 100	10	3010	14 061, 31 000	0.340	404, 82.8	0.581
Citation	11	3300	15 650, 34 500	0.371	–	0.586
Commercial jets						
Fokker 100	107	1680	44 450, 98 000	0.308	475, 97.3	0.556
Romero 1-11	109	1480	47 400, 104 500	0.289	494, 101.2	0.500
RJ100	112	2090	44 000, 97 000	0.290	572, 117.2	0.573
B717-200	106	1460	49 895, 110 000	0.291	536, 110.0	0.614
A318-100	107	2350	64 500, 142 200	0.330	526, 107.8	0.627*

*A derivative of a larger aircraft.

5.1.3.8 Commercial Analysis

This last topic in the analysis of the aircraft project considers the commercial viability of the whole project. Although this cannot be assessed in detail at the start of the project due to a lack of technical data, it is possible to prepare for a commercial analysis later in the design process.

This preparation will identify the potential market for the aircraft, the potential customers for the aircraft, and the main competitors. The design team will need to know what are the principal commercial parameters that potential customers (airlines and passengers) will use to judge the attractiveness of the new service in the total market. One of the obvious issues to be considered is aircraft costs. This includes the purchase price and various DOC parameters.

Finally, assessment of the operating issues relating to the new service will need to be understood. This will include the customer service for both pre- and in-flight parts of the operation.

5.1.4 Information Retrieval

As mentioned earlier in this chapter, this aircraft specification lies between long-range bizjets and regional feeder liners. The aircraft specified range is similar to the Gulfstream V but this bizjet only carries up to 15 passengers. The passenger capacity is similar to regional jets but they only fly about 1300 nm. To assess the design parameters that might be used in later sizing studies, Table 5.1-2 has been compiled, which shows some of the details of these two different types of aircraft.

Table 5.1-2 shows that the thrust to weight ratios (T/W) for the two types are significantly different. The reasons for this lie in the requirements for higher climb/cruise performance and short field performance for the bizjets. These are parameters that our aircraft should have, so a thrust/weight ratio of 0.32 (the lower value for bizjets and the upper one for regionals) will initially be assumed for our aircraft.

Wing loading (W/S) is also seen from the data in the table to be statistically different between the two aircraft groups. There may be a variety of operational criteria for this division but for the same reason as above, a value lying between the two sets will be selected. A value of 450 kg/sq. m, being low for regional jets but high for bizjets, will be used. This decision may mean that 'high-performance' flaps will be required.

Mass ratios are always difficult to assess from published data as there are often conflicting variations in the definition of terms. For example, empty weight ratio will be higher for smaller aircraft and smaller for long-range aircraft. It should be relatively easy to reassess the selected mass ratio following the first detailed mass estimations. Until this data is available it is necessary to make sensible 'guesstimates'. Values of 0.52 for the empty mass fraction and 0.35 for the fuel fraction seem reasonable, at this time.

For comparison, the values for these parameters for the VT study aircraft[2] are quoted as: 0.32, 535, 0.42, 0.32. Some of these differences can be explained by the larger size (165 PAX), shorter-range design specification of the VT study.

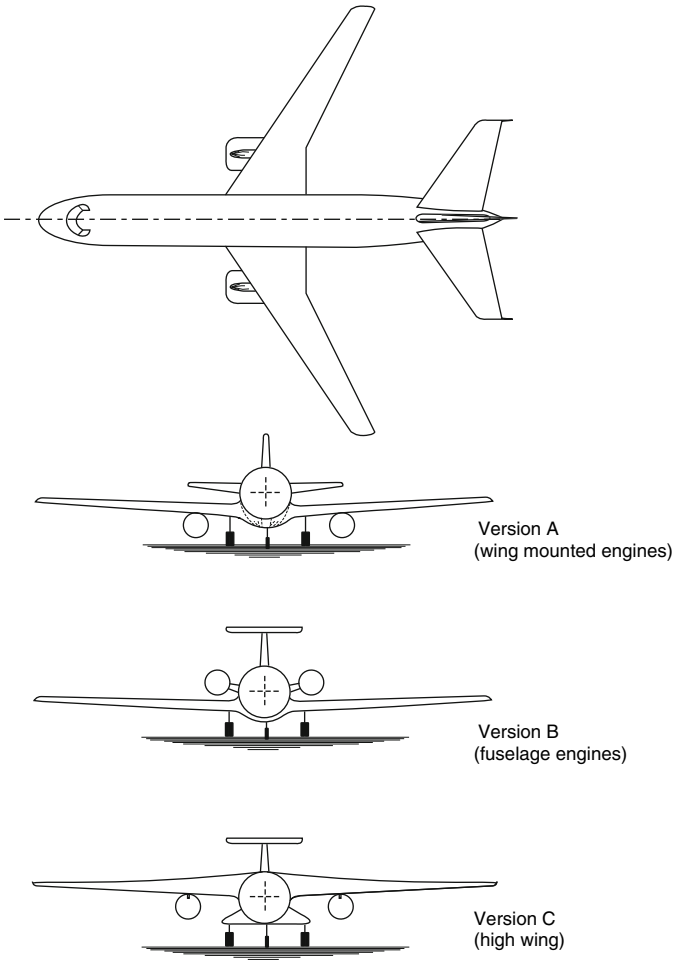


Figure 5.1-3 Conventional layouts.

5.1.5 Design Concepts

The previous section has shown that all of the potential competitors to the new design are of conventional configuration. They have trapezoidal, swept, low-mounted wings, with twin turbofan engines and tail control surfaces. Obviously, one of the concepts to consider is to follow this arrangement. The conservative airline industry may prefer such a choice. An alternative strategy is to adopt a novel/radical layout.

The 'new look' would set the aircraft apart from the competition and offer a marketing opportunity. In adopting such a design strategy, care must be taken to reduce technical risk and to show improved operational efficiency over the conventional layout.

Four design options are to be considered:

- Conventional layout
- Braced wing canard layout
- Three-surface layout
- Blended body layout

5.1.5.1 Conventional Layout(s) (Figure 5.1-3)

This must be regarded as a strong candidate for our baseline aircraft configuration as it is a well-proven, low-risk option. The technical analysis is relatively straightforward and has a high confidence level in the accuracy of the results. Its main advantage is that it is similar to the competitor aircraft and thereby with airport existing facilities and operations.

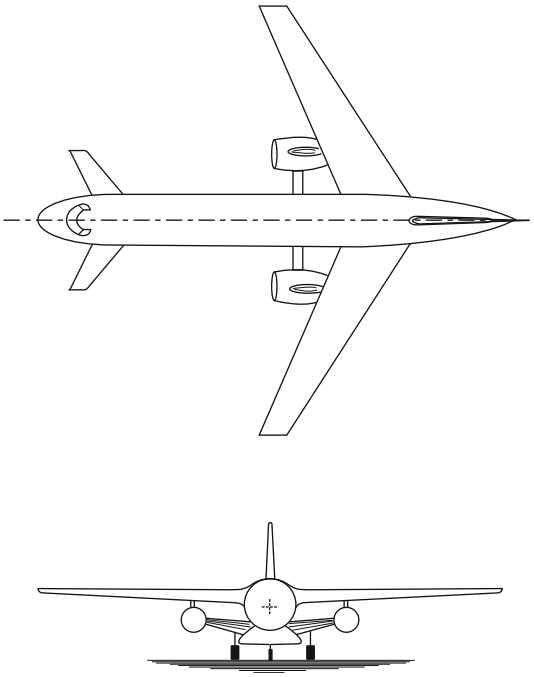


Figure 5.1-4 Braced wing layout.

There are some drawbacks to choosing this layout. These relate to the geometrical difficulties of mounting a high-bypass engine on a relatively small aircraft wing (relating mainly to ground clearance below the engine nacelle). This is illustrated in drawing A in Figure 5.1-3. There are two possible, alternative aircraft arrangements that could overcome this problem. Version B, in Figure 5.1-3, shows the engines mounted at the rear of the fuselage structure. This avoids the ground clearance problem but introduces other difficulties. Since a large component of aircraft mass is moved rearwards the aircraft centre of gravity also moves aft. This requires the wing to be moved back to balance the aircraft. The movement of the wing lift vector rearwards shortens the tail arm and consequently demands larger control surfaces. This increases profile drag and possibly trim drag in cruise. The second alternative layout is shown in version C in Figure 5.1-3. In this option the wing is moved to the top of the fuselage section (a high mounted wing). This lifts the engine away from the runway and provides adequate ground clearance. The high wing position, although used on some aircraft, is regarded as less crashworthy. The fuselage and therefore the passengers are not cushioned by the wing structure in the event of a forced landing. This is regarded as particularly significant in the case of ditching into water, as the fuselage would be below the floating wing structure. For an aircraft that is likely to spend long periods over water, airworthiness considerations may deter airlines from this type of layout.

A problem not necessarily restricted to the conventional layout is the potential lack of fuel tankage. A long-range aircraft will require substantial fuel storage and this may not be available in a conventional wing layout.

5.1.5.2 Braced Wing/Canard Layout (Figure 5.1-4)

Although this configuration looks radical, it is technically straightforward with well-proven, and understood, analysis that provides technical confidence.

The canard and swept forward wings offer low cruise drag possibilities. The rearward positioning of the engines reduce cabin noise. The bracing structure should reduce wing loads and allow a thinner wing section to be used. This, in combination, may reduce wing structural mass and aircraft drag. The main weakness of the layout lies in the uncertainty of the positioning of the canard, wing and engine components, and the interference effects of the airflow at the brace structure junctions. There is also some uncertainty about the effect of the brace on future stretch capability.

5.1.5.3 Three-Surface Layout (Figure 5.1-5)

This configuration has the advantage of low trim drag in cruise. The combination of forward and rear control and stability surfaces can be used to trim the aircraft in cruise with an upward (lift) force which will unload the wing. Two different wing layouts can be considered – swept forward or swept back. These options are shown in Figure 5.1-5.

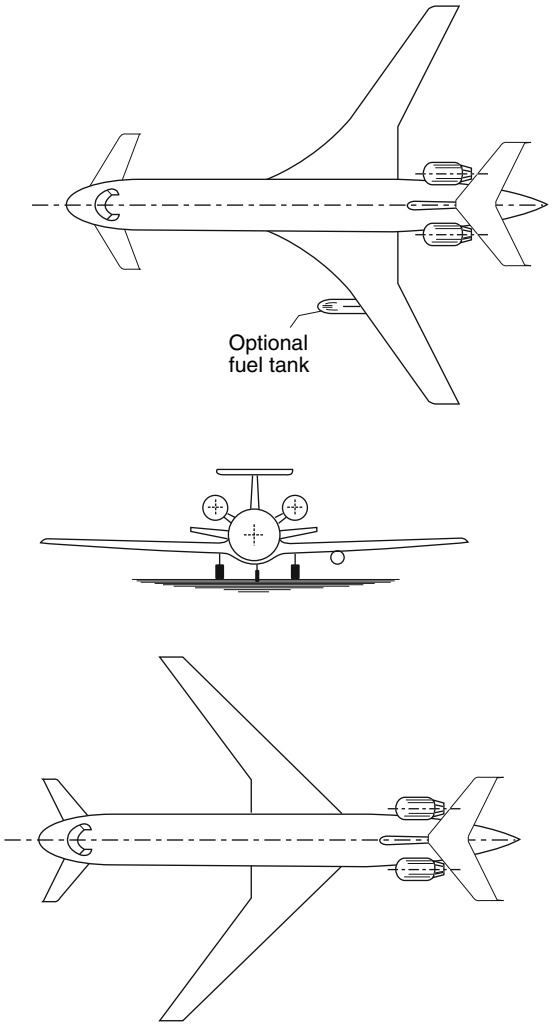


Figure 5.1-5 Three-surface layout.

It is anticipated that the swept forward configuration will be more suited to the development of laminar flow but may be heavy due to the need to avoid structural divergence. The bodyside wing chord will need to be sufficient to permit laminar flow systems to be installed. This is easier to arrange on the swept back layout. The increased internal wing volume created by the larger root chord will also provide increased wing fuel tankage. This together with the better flap efficiency of the swept back wing makes it the preferred choice of layout. The rear mounted engines will reduce cabin noise and visual intrusion although increase aircraft structural mass.

This layout is a strong contender for the preferred layout of our aircraft as the technical risks involved are low yet the configuration is distinctive. There may be a slight problem in positioning the forward passenger door due to the canard location but this should be solvable.

5.1.5.4 Blended Body Layout (Figure 5.1-6)

There has been a lot of interest in this configuration for the new supercapacity (550–1000 seats) aircraft. It is not novel. Several previous aircraft designs (mostly military) have adopted the layout. Aerodynamically, this layout is very efficient and lends itself to the installation of laminar flow control systems. For a long-range aircraft this is a major advantage, as fuel consumption will be reduced. The large internal volume of the wing should provide sufficient fuel tankage.

The main disadvantages relate to the difficulty of providing cabin windows and ensuring passenger evacuation in the case of an accident. Some innovation will be necessary to overcome these problems and satisfy airworthiness authorities. Airlines may be cautious of making this ‘step into the unknown’ due to the uncertainty of passenger

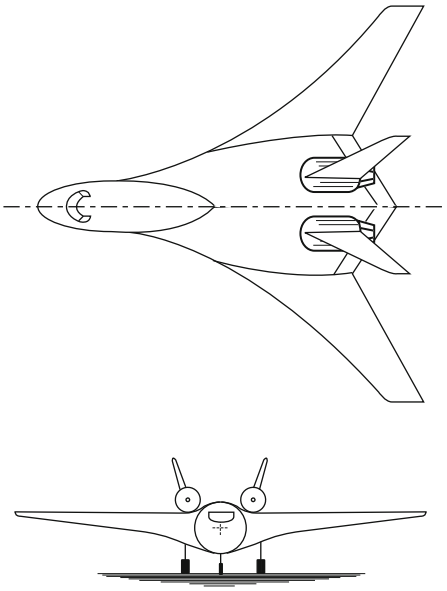


Figure 5.1-6 Blended body layout.

acceptance. A further problem, inherent in this layout, is the difficulty of stretching the integrated aircraft structure during programme development.

5.1.5.5 Configuration Selection

From a narrow commercial viewpoint, the conventional layout should be chosen, as it is a low-risk, low-cost option. However, there are doubts regarding the adequate provision of fuel tankage and the lack of a new 'aesthetic' for the service. Of the conventional layouts, the best is version B (rear fuselage mounted engines). This makes the passenger cabin less noisy, which would be seen as an advantage for an executive-class aircraft operating long endurance flights.

Although the conventional design is the natural choice, we will select one of the more radical configurations. In this way, it should be possible to compare, in more detail, the strengths and weaknesses of the design relative to the alternative (competitive) strategy of using a modified version of an existing aircraft. This comparative study could form the conclusion to the study.

Of the three novel configurations described above, the most radical is the braced wing layout. This option presents a larger commercial risk; therefore, it will not be pursued. The blended body aircraft is potentially a strong contender as the integrated structure/aerodynamic concept provides a technically efficient layout. The generous internal volume of the aircraft will suit development potential and offer adequate fuel tankage. The main difficulty is the lack of understanding of the internal structural framework. The integration of the wing air loading and the fuselage pressurisation loads is not easy to envisage. For a larger aircraft, this problem is eased due to the internal space available. For our smaller aircraft, such separation of load paths may not be feasible. Even if the structural problems could be solved, the integration of structure and aerodynamic designs would make it difficult to stretch the design into a family of layouts (e.g. Airbus A318, 319, 320, 321).

The three-surface configuration has been successfully used on other aircraft and has been shown to offer performance advantage over the conventional layout. The saving in fuel during cruise will reduce the tankage requirement. As mentioned previously, the canard will make the front fuselage design more complicated but this difficulty can be overcome by detail design. If necessary, a shortage of fuel volume may be avoided if external tanks are added to the wing or if fuselage tankage is allowed.

Based on the arguments above the three-surface layout will be selected for our baseline configuration.

5.1.6 Initial Sizing and Layout

At this point in the design process, we can begin to realise the aircraft geometrical configuration. It is necessary first to make an estimate of the aircraft mass. Using this value, the engine and wing sizes can be determined. The fuselage shape is determined from the internal layout and tail requirements. With the main components individually defined, it is then possible to produce the first scale drawing of the aircraft. Crude estimates from similar aircraft types are necessary to complete the layout (e.g. tail and landing gear).

5.1.6.1 Mass Estimation

From section 5.1.4 the following mass parameters were suggested:

Empty mass fraction = 0.52
Fuel mass fraction = 0.35

The mass estimations below are shown in kg only (conversion factor: 1 kg = 2.205 lb).

The payload is specified as 80 business-class passengers and their baggage. For this type of 'premium-ticket' operation the mass allowance per passenger (including baggage) will be larger than normal. We will allow 120 kg per passenger. The flight rules will dictate at least two pilots. Airlines will want to provide high-class service in the cabin so we will assume four cabin attendants are required. It is common practice[1] to allow 100 kg for each flight crew and 80 kg for each cabin attendant.

Hence, the payload is estimated to be

$$M_{pay} = (80 \times 120) + (2 \times 100) + (4 \times 80) = 10\,120 \text{ kg}$$

To cover incidental flight services, allow an extra 5 kg per passenger. This adds 400 kg to the payload mass, making the aircraft 'useful mass' = 10 520 kg.

Using the equation with the values above gives

$$\text{MTOM} = 10\,520/(1 - 0.52 - 0.35) = 80\,923 \text{ kg } (178\,435 \text{ lb})$$

The initial mass statement is:

	kg	% M_{TO}
Operational empty	42 080	52
Extra services	400	13
Crew (2 + 4)	520	
Passengers	9 600	
Fuel	28 323	35
M_{TO} =	80 923 (178 435 lb)	100

5.1.6.2 Engine Size and Selection

The literature survey (section 5.1.4) indicated a thrust to weight ratio of 0.32 was appropriate.

Hence:

Engine total take-off thrust = 0.32 × 80 923 × 9.81 = 254 kN (57 100 lb)

With two engines this equates to 127 kN per engine (28 550 lb)

A choice of engines from different manufacturers is always the preferred commercial position for the airframe manufacturer. This ensures that the engine price and availability is more competitive. It also provides the potential airline customer with more bargaining power when selecting the aircraft/engine purchase.

There are several available engines that would suit our requirement. All of them are currently used on civil aircraft operations; therefore, considerable experience is available. The engines below are typical options:

- CFM56-5B as used on the A320
- CM56-5C as used on the A340
- IAE-V2533 as used on the MD90 family
- IAE-V2528 as used on the A321.

The details* below are representative of these engines:

Bypass ratio	5.5
Thrust ISA-sea-level static	31 000 to 34 000 lb (138 to 151 kN)
Typical cruise thrust (max.)	5840 to 6910 lb (26 to 31 kN)
Cruise specific fuel consumption	0.594 to 0.567
Length	102 in (2.6 m)
Diameter	68.3 to 72.3 in (1.7 to 1.8 m)
Engine dry weight (mass)	5250 lb (2381 kg)

*Note: engine manufacturers commonly quote values in Imperial units.

These details will be enough for initial performance and layout purposes, but as the design progresses it will be necessary to periodically review the choice of engines to be used on the aircraft.

5.1.6.3 Wing Geometry

The recommended wing loading is 450 kg/sq. m; hence:

Wing gross area (S) = 80 923/450 = 180 sq. m; (1935 sq. ft)

Selecting a high aspect ratio (AR) will lower induced drag in cruise and save fuel. A value of 10 is to be used. The choice of aspect ratio will need to be reviewed in a trade-off study later in the design process. Using the wing area and aspect ratio we can determine:

Wing span (b) = $(\mathrm{AR} \times S)^{0.5}$ = 42.4 m (135 ft)

Mean chord (c_m) = (b/AR) = 4.24 m (13.5 ft)

Selecting a taper ratio of 0.3 gives (approximately)

Tip chord = 2.0 m (6.6 ft)

Centre line chord = 6.52 m (21.4 ft)

Using published data on critical Mach number analysis, at our preferred cruise speed of M0.85, a sweepback angle of 30° will allow a maximum wing thickness of 15 per cent without incurring wave drag penalties. This wing thickness will allow space for the proposed laminar flow control system and offer extra fuel tankage. An unswept inboard trailing edge will make the flap more effective and provide space for the main undercarriage, retraction mechanism. This will add extra area so the wing chords will be changed from the values calculated above to 1.8 m and 6.0 m (5.9 and 19.7 ft) respectively to retain approximately 180 sq. m (1935 sq. ft) area.

The basic wing geometry is shown in Figure 5.1-7. This also shows the location of the wing fuel tanks and the position of the mean aerodynamic chord (MAC). As an initial assumption the longitudinal position of the wing on the fuselage will be arranged to line up the position of the MAC quarter-chord with the aircraft centre of gravity.

Fuel volume considerations

At this point in the design process it is necessary to determine the size of the required fuel volume estimated in the initial mass estimation and then to compare this to the available space in the wing. This involves transforming the fuel mass into a volume. Fuel volume/capacity is often quoted in terms of 'gallons'. This must be converted into linear units (cubic metres or cubic inches) to relate the size to the aircraft geometry. To do this conversion it is necessary to understand the various systems of units used. The calculation is shown in detail below because it is seldom to be found in other aircraft design textbooks although it is always a significant consideration.

In SI units – one litre is defined as the volume required to hold one kilogram of water. It is further defined as 1 litre = 1000 cubic centimetres (i.e. 0.001 cubic metres).

Hence, 1000 kg of water occupies 1.0 m^3

In USA – one US gallon equates to the volume required to hold 8.33 lb of water. For water at 62.43 lb/ft^3 a US gallon therefore corresponds to a volume of 231 cubic inches.

Hence, 1000 lb of water occupies 120 US gallons (= 16.02 ft^3)

Warning: In the UK the definition of the gallon is different to the USA gallon!

In the UK, the Imperial gallon is used to measure liquid capacity. This is defined as the volume required to hold 10 lb of water. This makes the Imp gal. = 277.42 cubic inches.

Hence 1000 lb of water occupies 100 Imp. gallons (= 16.02 ft^3)

Note: the density of a liquid (water) does not change with the system of units!

Here are some useful conversion factors:

1 US gal = 0.833 Imp. gal	1 Imp. gal = 1.2 US gal
1 US gal = 3.79 litres	1 Imp. gal = 4.55 litres
1 cubic foot = 28.32 litres	
1 cubic metre = 1000 litres	1 cubic metre = 35.3 cubic feet

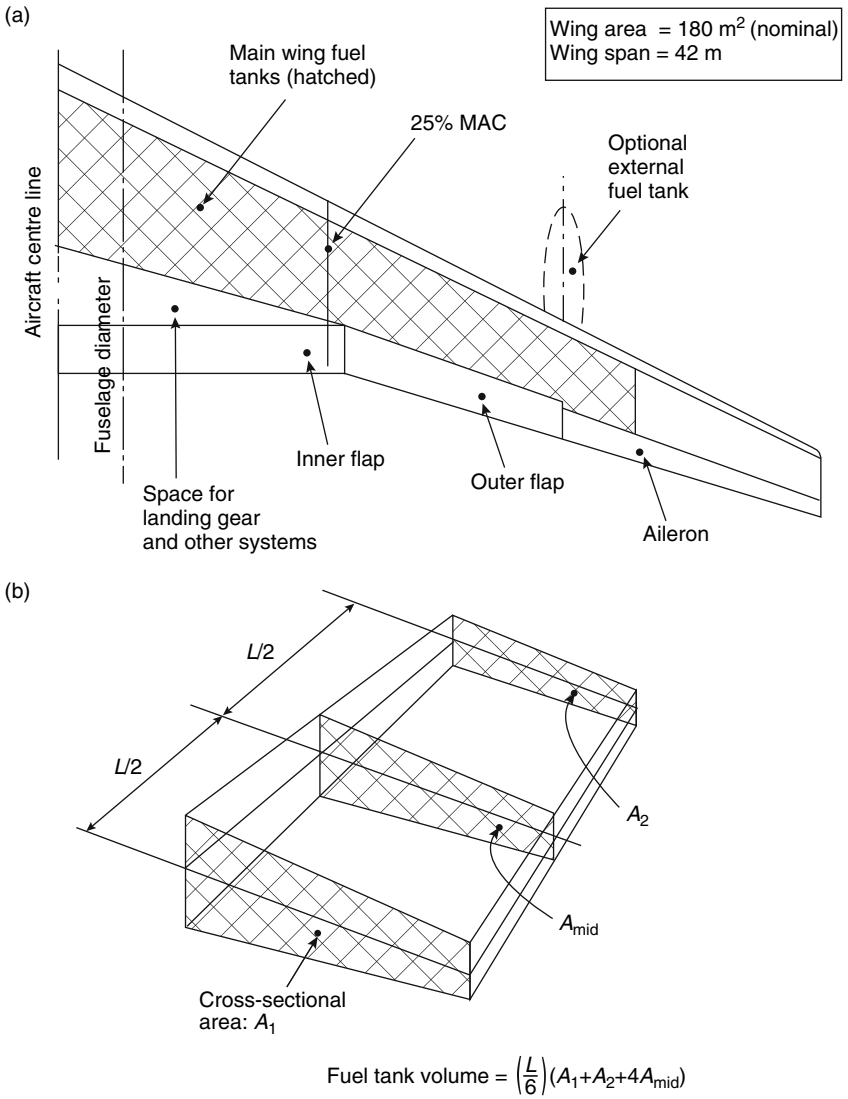


Figure 5.1-7 (a) Initial wing planform geometry. (b) Fuel tank volume.

Specific gravity is the unit that relates the density of a liquid to water. For aviation fuel, specific gravity varies with the type of fuel (e.g. JP1, 3, 4, 5, 6, or kerosene) between values in the range 0.82 to 0.76. For civil aircraft fuel, a value of 0.77 can be assumed; hence:

1000 kg of fuel occupies 1.3 m^3
1000 kg of fuel occupies 1300 litres
1000 lb of fuel occupies 155.8 US gallons
1000 lb of fuel occupies 129.9 Imp. gallons
1000 lb of fuel occupies 20.8 ft^3

Estimation of wing fuel volume

To determine the usable capacity of a fuel tank it is possible to calculate the external volume and then reduce this value to account for internal obstructions caused by structural and system components within the tank. Typically, reduce the available internal volume to 85 per cent for integral tanks and to 65 per cent for bladder or 'bag tanks'. Note: these factors do not account for landing gear or other significant intrusions into the available space. Such factors must be considered separately when deciding the overall location of fuel tanks.

For the aircraft in this project, we can consider the fuel to be held in wing tanks on each side of the aircraft fuselage as shown in Figure 5.1-7. Each tank occupies the space between the leading edge and trailing edge high-lift structure and associated mechanisms. The tanks will be of the integral type. The space ahead of the ailerons will not be used

Table 5.1-3

Section	A/C centre	Mid.	Outer
Wing chord (c)	8.5	5.0	3.0
Wing thickness (t/c)	0.15	0.13	0.11
Thickness factor (k)	0.6	0.65	0.65
Thickness (T)	0.765	0.422	0.214
Widths (W)	4.2	2.6	1.7
Cross-sect. areas (A)	3.06	1.52	0.32
Tank length (L) = 16.0	Tank max. volume (profile) = 21.2 m^3		

for fuel tankage, as the wing section here is too thin. Also, the space in front of the inboard flap is not used for fuel volume, as this is likely to be where the main landing gear will be stowed. The generalised geometry of the fuel tank is shown in Figure 5.1-7a.

The cross-section areas of the spanwise ends (A_1 and A_2) and mid-span (A_{mid}) sections of each tank are determined by multiplying the average tank thickness (chordwise) (T) by the distance between the front and rear spars (W). The cross-sectional areas of each end of a tank are added and then multiplied by half the spanwise distance between the ends (L) to give the profile volume. This volume is then multiplied by the appropriate factor for the type of tank. Hence:

$$\text{Average thickness } T = k \cdot (t/c) \cdot c$$

where $k =$ factor to relate the average tank depth to the max. wing profile depth. The value depends on the shape of the section profile and the allowance made for structure. Typical values lie between 0.8 and 0.5.

$(t/c) =$ wing section profile thickness ratio (this will vary from thicker values near the bodyside to thinner values towards the tip).

$c =$ the local wing chord length

$$\text{Cross-sectional area } (A) = T \cdot W$$

where W is the width of the tank

$$\text{Tank profile volume} = (L/6)[A_1 + A_2 + 4A_{mid}]$$

where A_1 and A_2 are the cross-sectional areas of the ends of the tank

A_{mid} is the cross-sectional area at the mid-length position

L is the length of the tank

The tank measurements (in metres) are quoted in Table 5.1-3. Total tank volume (both sides), including an 85 per cent factor for integral tankage:

Available tank volume $= 2 \times 21.2 \times 0.85 = 36.1$ m^3

Required fuel mass (from section 5.1.6.1) = 28 323 kg (62 450 lb)

Using the volume conversion shown above, for typical aviation fuel:

Required tank volume $= (28\,323/1000) \cdot 1.3 = 36.82$ m^3 (590 ft^3)

Hence, within the accuracy of the calculation, the required fuel can be accommodated in the wing profile tanks. Extra fuel volume will be useful to extend the range of the aircraft for reduced payload operations. This could be provided by the optional external wing tanks but these would add extra drag.

Table 5.1-4

Class	Seats abreast	Seat width	Cabin internal width
Executive	4	0.70 m	$(4 \times 0.7) + 0.6 = 3.4$ m
Tourist	5	0.56 m	$(5 \times 0.56) + 0.6 = 3.4$ m
Charter	6	0.47 m	$(6 \times 0.47) + 0.58 = 3.4$ m

5.1.6.4 Fuselage Geometry

For most aircraft, the fuselage layout can be considered in isolation to the wing and other control surfaces. The internal space requirements, set by the aircraft specification, are used to fix the central section of the fuselage. For civil aircraft, this shape is governed by the passenger cabin layout.

The fuselage width is set by the number of seats abreast, the seat width and the aisle width. The depth is set to accommodate the cargo containers below the floor and the headroom above the aisle. A circular section is preferred for an efficient structural pressure shell. This requirement may impose constraints on the preferred width and depth sizes. Although this aircraft is designed principally as an executive aircraft, we must make sure that the size is suitable for any other variants that we may want to consider as part of the aircraft 'family'. For an aircraft of 80+ capacity, the conventional seating (mixed class) would be five abreast for economy and four abreast for business, with a single aisle. For our executive layout, four abreast would be sensible. As the aircraft mission is long range, it is necessary to provide a high comfort level. A typical maximum first-class seat is 0.7 m (27.5 in) wide. Providing a generous 0.6 m (24 in) aisle would make the cabin width 3.4 m (136 in). Adding 0.1 m (4 in) each side for structure makes the fuselage outside diameter 3.6 m (142 in). This width would allow five abreast 'tourist' seating with a seat width of 0.56 m (22 in). This is currently regarded as a very generous provision for this class. At six abreast the 'tourist/charter' seat width is 0.47 m (18.5 in). This is narrow for normal tourist provision but generous for the charter operation. The fuselage layout options are as shown in Table 5.1-4.

Adding 0.2 m for the pressure cabin structure makes:

Total fuselage external diameter equal to 3.60 m (11.8 ft or 142 in)

The fuselage cross-section must also be considered in relation to the cargo pallet sizes to be accommodated below the cabin floor. This may require the fuselage profile to be altered to suit the geometry of standard containers. For example, the Boeing 757 fuselage section is 10 in deeper than the circular cabin shape. It is too early in the design process to consider such details but this aspect must be carefully studied later.

The length of the cabin is determined by the seat pitch. This varies as the class. Typical values are: executive class is 1.0 to 1.1 m (40 to 43 in), tourist is 0.8 to 0.9 m (31 to 35 in), and charter is 0.7 to 0.8 m (28 to 31 in). Using the longest executive seat pitch with four abreast seating requires a cabin length of 22 m (72 ft). With this length of cabin, the number of tourist passengers that can be accommodated is 140 and 120 for the short and longer pitches respectively. A similar calculation for the charter layout would provide 192 or 140 passengers respectively. It may not be

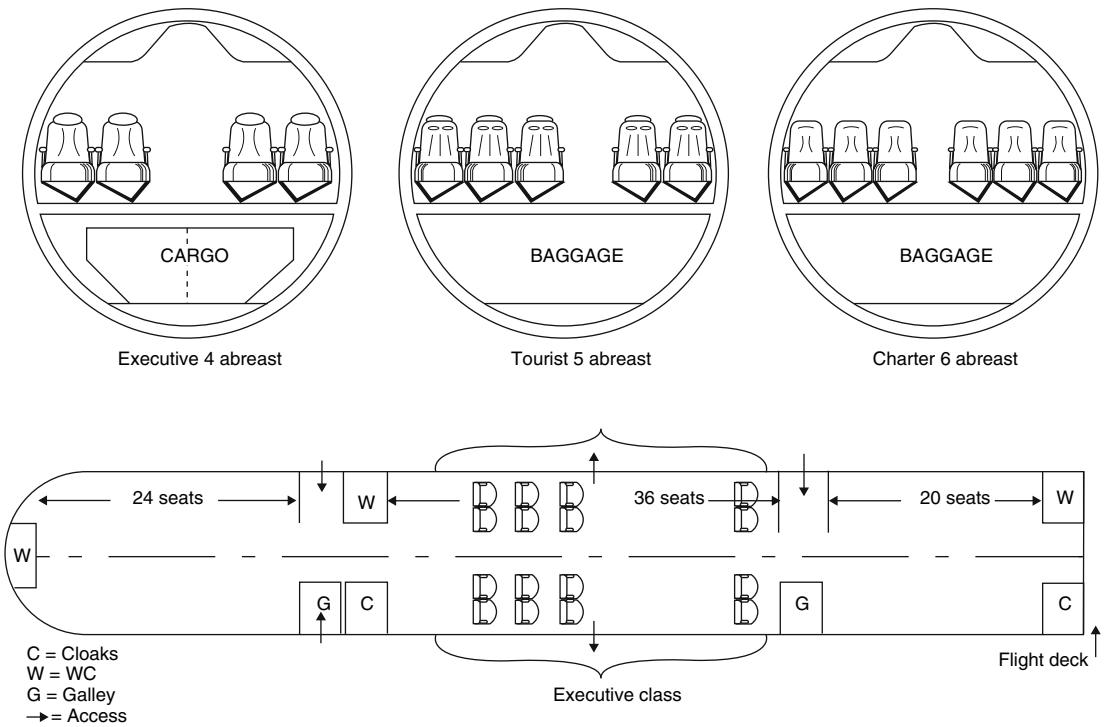


Figure 5.1-8 Fuselage layout options.

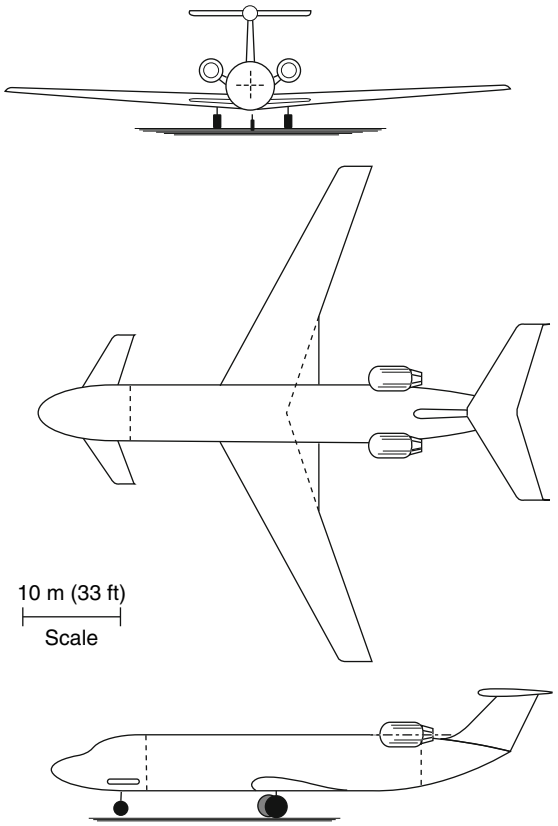


Figure 5.1-9 Initial 'baseline' layout.

possible for technical reasons (e.g. provision of emergency and other services) to accommodate the larger capacities calculated here. Nevertheless, the 22 m cabin length seems to offer a good starting point for the initial layout.

It is desirable to split the cabin into at least two separate sections. This makes the in-flight servicing easier and allows more options for the airline to segregate different classes. For the exclusive executive layout, this division will allow a quieter environment within the cabin. A service module (catering or toilets) is positioned at this location. External service doors and hatches are positioned here and these can act as emergency exits. The provision of service modules and the 'wasted' space adjacent to the doors will add about 4 metres (13 feet) to the cabin length.

The fuselage length is the sum of the cabin and the front and rear profile shaping. The front accommodates the flight deck and the rear provides attachment for the engines (in our case) and the tail surfaces. From an analysis of similar aircraft, the non-cabin length is about 15 metres.

$$\text{Hence, the total fuselage length is } (22 + 4 + 15) = 41 \text{ m } (134 \text{ ft})$$

The resulting fuselage layout and geometry are shown in Figure 5.1-8.

5.1.6.5 Initial 'Baseline Aircraft' General Arrangement Drawing

With details of the engine, wing and fuselage available, it is now possible to produce the first drawing of the aircraft. The control surface sizes are estimated from area and tail volume coefficients of other similar aircraft. The aircraft general arrangement (GA) drawing is shown in Figure 5.1-9. With the sizes of the major components of the aircraft available from the GA, it is possible to make the initial technical assessments.

5.1.7 Initial Estimates

Before we can estimate aircraft performance, we must more accurately determine the aircraft mass, drag, lift and engine characteristics. Detailed calculations are not shown in the sections below as they follow conventional methods. Where appropriate, reference is given to the methods used.

5.1.7.1 Mass and Balance Analysis

Each of the aircraft component masses will be estimated separately and then summed. For presentational convenience, the values below are quoted in kg only (conversion 1 kg = 2.205 lb).

Wing structure

Using the estimation formula in reference 1, with:

MTOM = 81 000 kg N_{ult} = 3.75
Wing area = 180 sq. m Wing aspect ratio = 10
Wing taper ratio = 0.3
Wing av. thickness = 15% Wing LE sweep = 30°
$R = (M_{wing} + M_{fuel}) = (8100 + 32\ 400)$ kg
Hence, the wing is calculated at : M_{wing} = 11 209 kg

This is 13.8 per cent of MTOM. This is uncharacteristically high for this type of aircraft. The formula is based on old and existing aircraft types. The average value of aspect ratio for the source aircraft is about 6. The much higher aspect ratio of our design (10) seems to have caused a large increase in wing mass. In addition, the formula was based on traditional metallic construction whereas our design will incorporate substantial composite structure. For these reasons, we will apply a reduction factor of 25 per cent to the estimated mass:

$M_{wing} = 11\ 209 \times 0.75 = 8407$ kg (10.4% MTOM)

As we have used the wing gross area, we will assume that this mass includes the flap weight.

Tail structure

With little knowledge of the tail design at this time, we will assume a representative percentage. We know that the extra control surface (canard) will add some weight so we will use a slightly higher percentage than normal for this type of aircraft (2.5 per cent). As we will be constructing these surfaces in composite materials, we will apply a technology reduction factor of 25 per cent.

$M_{tail} = 0.025 \times 81\ 000 \times 0.75 = 1519$ kg (1.9% MTOM)

Fuselage structure

The mass of the body will be estimated using a formula for body[1] with the parameter values shown below:

MTOM = 81 000 kg; O/A length = 40.0 m;
max. diameter = 3.75 m; V_D = 255 m/s

Gives:

$M_{body} = 9232$ kg

Increasing by 8 per cent for pressurisation, 4 per cent for tail engine location and reducing by 10 per cent for modern materials and construction gives:

$M_{body} = 9306$ kg (11.5 MTOM)

Nacelle structure

Based on an engine thrust of 254 kN:

$M_{nacelles} = 1729$ kg (2.1% MOTOM)

Landing gear

We will assume this to be 4.45 per cent MTOM:

$M_{landing\ gear} = 3604$ kg

Surface controls

Using a typical value of 0.4 × (MTOM)$^{0.684}$:

$M_{s/controls} = 911$ kg

Aircraft structure

Summing the above components gives the aircraft structural mass:

$$M_{\text{structure}} = 25\ 475 \text{ kg}$$

This is 31.5 per cent MTOM which is representative of this class of aircraft.

Propulsion system

Using the quoted engine dry weight of 5250 lb (each) and a system multiplying factor of 1.43 gives

$$M_{\text{propulsion}} = 6810 \text{ kg } (8.4\% \text{ MTOM})$$

Fixed equipment

A typical value for this type of aircraft is 8 per cent but as we will be providing more cabin services we will increase this to 10 per cent MTOM:

$$M_{\text{fix/equip}} = 8100 \text{ kg}$$

Aircraft empty (basic) mass

Summing the structure, propulsion and fixed equipment masses gives

$$M_{\text{empty}} = 40\ 385 \text{ kg } (49.9\% \text{ MTOM})$$

Operational empty mass (OEM)

Adding the flight crew ($2 \times 100 = 200$ kg), cabin crew ($4 \times 70 = 280$ kg), cabin service and water (@21.5 kg/pass. = 1724 kg) to the aircraft basic mass gives

$$M_{\text{OEM}} = 42\ 589 \text{ kg } (52.8\% \text{ MTOM})$$

This is close to the assumed value from the literature search.

Aircraft zero-fuel mass (ZFM)

This is the OEM plus the passengers ($80 \times 80 = 6400$ kg) and the passenger baggage ($40 \times 80 = 3200$ kg), giving

$$M_{\text{ZFM}} = 52\ 189 \text{ kg } (64.4\% \text{ MTOM})$$

Maximum take-off mass (MTOM)

In the analysis above, the MTOM has been assumed to be 81 000 kg.

Fuel mass

The aircraft zero-fuel mass (ZFM) and the assumed MTOM define the available fuel mass:

$$M_{\text{fuel}} = \text{MTOM} - \text{ZFM}$$

Hence,

$$M_{\text{fuel}} = 81\ 000 - 52\ 189 = 28\ 811 \text{ kg } (35.6\% \text{ MTOM})$$

This is less than previously assumed so it will be necessary to recalculate the fuel mass ratio using a more detailed method. The Breguet range equation can be used if assumptions are made for the aircraft (L/D) ratio and the engine fuel consumption (c).

$$\text{Range} = (V/c)(L/D)\log_e(M_1/M_2)$$

where V = cruise speed = $M.85$* = 255 m/s = 485 kt
c = assumed engine fuel consumption = 0.55 N/N/hr
(L/D) assumed to be = 17 in cruise
M_1 = start mass = MTOM = 81 000 kg
M_2 = end mass = ZFM = 52 189 kg

*the cruise speed is set to avoid incurring significant drag rise. Typically, a 20 point drag count (one drag count = 0.0001) rise sets this speed.

With the speed in knots, this gives

Range = 6589 nm

Although this may seem close to the specified range of 7000 nm, it is necessary to account for the fuel allowances. Using the formula shown below,[1] the required design range can be used to calculate the equivalent still-air-range (ESAR). This includes the fuel reserves (diversion and hold) and other contingency fuel:

ESAR = 568 + 1.06 design range

Hence, the required ESAR (for our specified design range of 7000 nm) = 7988 nm.

Reversing this process with the 6589 nm range calculated above would only give a design range of 5927 nm. This shows that there is a substantial shortfall in the design range. The original assumption of 0.35 for the fuel fraction seems to be in error for our design. This is a major error as the aircraft is not viable at an MTOM of 81 000 kg. We can use the Breguet equation above to determine a viable fuel ratio for the 7988 nm ESAR:

$$7988 = (485/0.55)\ 17(\log_e(M_1/M_2))$$
$$(M_1/M_2) = 1.704$$
$$M_2 = M_1 - M_{\text{fuel}}$$
$$(M_{\text{fuel}}/\text{MTOM}) = 1 - (1/1.704) = 0.413$$

This is a big change to the value used in the initial MTOM prediction. We will use the aircraft empty mass ratio of 0.495 determined in the component mass evaluation above, in a new estimation of MTOM:

MTOM = 110 520/(1 − 0.495 − 0.413) = 114 348 kg (25 214 lb)

Note: the denominator in the expression above is only 0.092. This makes the evaluation very unstable. For example, if the empty mass and fuel mass ratios are incorrect by only +/ − 1 per cent, the MTOM would change to 146 111 and 93 928 kg respectively. These values are 28 per cent more and 18 per cent less than the predicted value. This illustrates the inappropriate use of the initial MTOM prediction method when the denominator is small. However, as we do not have another prediction, we will have to use the 114 348 value and, as quickly as possible, validate it with a detailed component mass prediction.

As we have still not evaluated the aircraft performance, we will need to use the thrust and wing loading values (0.32 and 450) determined in the literature survey. The new value of MTOM will force a change in the engine thrust and wing area:

Engine thrust (total) = 359 kN (80 710 lb)
Wing area (gross) = 254 sq. m (2730 sq. ft)

As other alterations are likely to follow, changes to the engine selection caused by the above will not be considered at this point in the design process.

The values above, together with the resulting heavier MTOM, will change the component mass predictions made earlier. Using the same methods, the aircraft mass statement (kg) is calculated as listed below:

Wing structure	= 13 224	(11.6% MTOM)
Tail structure	= 2859	(2.5% MTOM)
Body structure	= 10 278	(9.0% MTOM)
Nacelle structure	= 2441	(2.1% MTOM)
Landing gear	= 5088	(4.45% MTOM)
Surf. controls	= 1153	(1.0% MTOM)
STRUCTURE	= 35 043	(30.6% MTOM)
Propulsion	= 9625	
Fixed equip.	= 11 435	
A/C EMPTY	= 56 103	(49.1% MTOM)
Operational items	= 1724	
Crew	= 480	
OEM	= 58 307	(51.0% MTOM)
Passengers	= 6400	
Baggage	= 3200	
ZFM	= 67 907	(59.4% MTOM)
Fuel	= 46 441	(40.6% MTOM)
MTOM	= 114 348 (252 137 lb)	(100% MTOM)

Note: the fuel ratio is still slightly under the requirement. The calculation should be done again to obtain the correct ratio.

Applying the Breguet range equation with values determined above ($M_1 = 114\,348$ and $M_2 = 67\,068$ kg) gives a range of 7812 nm. (A spreadsheet method with an iterative calculation function is very useful in this type of work.) As we have still made some gross assumptions in the calculations above (e.g. if the aircraft L/D ratio is 18 instead of 17 the range would increase to 8272 nm), we will continue the design process using the 114 348 MTOM value.

Before moving on to the aerodynamic calculations, it is necessary to redraw the aircraft with larger wings, control surfaces and engines. The fuselage shape will not change. The overall aircraft layout will be similar to that shown later in Figure 5.1-11. Assuming that the internal wing volume increases as the cube of the linear dimensions, the wing will be able to hold 52 668 kg (116 134 lb). This will be large enough to hold the extra fuel mass of the bigger aircraft.

5.1.7.2 Aerodynamic Estimations

Conventional methods for the estimation of aircraft drag can be used at this stage in the design process. As it is assumed that, with careful detail design, the aircraft can fly at speeds below the critical Mach number, substantial additions due to wave drag can be ignored. Therefore, only zero-lift and induced drag estimations are required.

Parasitic drag is estimated for each of the main component parts of the aircraft and then summed to provide the 'whole aircraft' drag coefficient. The component drag areas are normalised to the aircraft reference area (normally the wing gross area).

$$\text{Component parasitic drag coefficient, } C_{\text{Do}} = C_{\text{f}} F Q [S_{\text{wet}}/S_{\text{ref}}]$$

where C_{f} = component skin friction coefficient. This is a function of local Reynolds number and Mach number
F = component form (shape) factor which is a function of the geometry
Q = a multiplying factor (between 1.0 and 1.3) to account for local interference effects caused by the component
S_{wet} = component wetted area
S_{ref} = aircraft drag coefficient reference area (normally the wing gross area)

Aircraft not in the 'clean' condition (e.g. with landing gear and/or flaps lowered, with external stores or fuel tanks) will also be affected by extra drag (ΔC_{Do}) from these items. The extra drag values will be estimated from past experience. Several textbooks (e.g. references 1 to 5) and reports provide data that can be used:

$$\text{Whole aircraft parasitic drag, } C_{\text{Do}} = \sum [\text{component } C_{\text{Do}}] + \sum [\Delta C_{\text{Do}}]$$

From the previous analysis the reference area will be 254 sq. m (2730 sq. ft).

Cruise (at 35 000 ft and M0.85)

The component drag estimations for the aircraft in this clean configuration are shown in Table 5.1-5.

From reference 1, induced drag coefficient,

$$C_{\text{Di}} = (C_1/C_2/\pi A) C_{\text{L}}^2 + (0.0004 + 0.15 C_{\text{Do}}) C_{\text{L}}^2$$

where (C_1 and C_2) are wing geometry factors (close to unity) and (A) is the wing aspect ratio.

For our design the equation above gives

$$C_{\text{Di}} = 0.035 C_{\text{L}}^2$$

Table 5.1-5

Component	R. No.*	C_{f}	F	Q	S_{wet}	(ΔC_{Do})
Wing	3.32	0.00234	1.50	1.50	432.0	0.00593
H controls	1.86	0.00255	1.31	1.2	59.7	0.00094
V control	2.99	0.00237	1.32	1.2	33.7	0.00050
Fuselage	2.65	0.00175	1.07	1.0	437.4	0.00321
Nacelles (2off)	3.42	0.00231	1.5	1.0	84.6	0.00116
Secondary items						0.00192
					$\sum$ (aircraft C_{Do})	0.01376

R. No. = Reynolds number ($\times 10^{-7}$)

5.1.7.3 Initial Performance Estimates

Cruise

Hence, at the start of cruise:

$$C_D = 0.0137 + 0.035C_L^2 \quad \text{and} \quad C_L = 0.339,$$

Making,

$$C_D = 0.01\,774$$

Therefore, at the start of cruise,

$$\text{Aircraft drag} = 54.3 \text{ kN}$$

Assuming, at this point, the aircraft mass is (0.98 MTOM), then L/D ratio = 19.1
Engine lapse rate to cruise altitude = 0.197 (based on published data[1])
Hence, available engine thrust = $0.197 \times 359 = 70.7$ kN

This shows that the engine cruise setting could be 77 per cent of the take-off rating.

At the end of the cruise phase, assuming that aircraft mass is (0.65 MTOM) the aircraft C_L reduces to 0.225 if the cruise height remains constant. This reduces the aircraft L/D ratio to 14.5. This would increase fuel use. To avoid this penalty the aircraft could increase altitude progressively as fuel mass is reduced to increase C_L. This is called the 'cruise-climb' or 'drift-up' technique during which the aircraft is flown at constant lift coefficient. At the end of cruise, the aircraft would need to have progressively climbed up to a height of 43 600 ft. To reach such an altitude may not be feasible if the engine thrust has reduced (due to engine lapse rate) below that required to meet the cruise/climb drag.

Cruise/climb

At the initial cruise height, the aircraft must be able to climb up to the next flight level with a climb rate of at least 300 fpm (1.524 m/s).

This will require an extra thrust of 6758 N.

Adding this to the cruise drag gives 61.1 kN. This is still below the available thrust at this height (approximately 86 per cent of the equivalent take-off thrust rating).

Performing a reverse analysis shows that an aircraft (T/W) ratio of 0.276 would be adequate to meet the cruise/climb requirement.

Landing

The two-dimensional (sectional) maximum lift coefficient for the clean wing is calculated at 1.88. The finite wing geometry and sweep reduce this value to 1.46. Adding simple (cheap) trailing edge flaps ($\Delta C_{L\max} = 0.749$) and leading edge device ($\Delta C_{L\max} = 0.198$) produces a landing max. lift coefficient for the wing of 2.41.

At this stage in the design process, it is sufficient to estimate the landing distance using an empirical function. Howe[3] provides as simplified formula that can be used to estimate the FAR factored landing distance. The approach lift coefficient (C_{Lapp}) is a function of the approach speed. This is defined in the airworthiness regulations as 1.3 times the stall speed in the landing configuration. Hence C_{Lapp} is (2.4/1.69) = 1.42. Assuming the landing mass is (0.8 MTOM), the approach speed is estimated as 64 m/s (124 kt). This equates to a landing distance of

$$\text{FAR landing distance} = 1579 \text{ m } (5177 \text{ ft})$$

This is less than the design requirement of 1800 m.

Take-off

Reducing the flap angle for take-off decreases the max. lift coefficient to 2.11.

As for the landing calculation, it is acceptable at this stage to use an empirical function to determine take-off distance (TOD). For sea level ISA conditions, reference 3 gives a simplified formula for the FAR factored TOD. Assuming lift-off speed is 1.15 stall speed, the lift coefficient at lift-off will be $(2.11/1.15^2) = 1.59$, with $(T/W) = 0.32$ and $(W/S) = 450 \times 9.81 = 4414$ N/sq. m, the following values are calculated:

Ground run = 1292.6 m, Rotation distance = 316.1 m, Climb distance = 81.6 m, FAR TOD = 1690 m (5541 ft)

This easily meets the previously specified 1800 m design requirement.

Second segment climb with one engine inoperative (OEI)

For the second segment calculation the drag estimation follows the same procedure as described above but in this case the Reynolds number and Mach number are smaller. The undercarriage is retracted and therefore does not add extra drag but the flaps are still in the take-off position and will need to be accounted for in the drag estimation. The failed engine will add windmilling drag and the side-slip (and/or bank angle) of the aircraft will also add extra drag.

Using published methods to determine flap drag[3] and other extra drag items[1]:

$$C_{\mathrm{L}} = 1.59, \qquad C_{\mathrm{DO}} = 0.0152, \qquad C_{\mathrm{DI}} = 0.0376,$$
$$\Delta C_{\mathrm{Dflaps}} = 0.015, \quad \Delta C_{\mathrm{Dwdmill}} = 0.0033, \quad \Delta C_{\mathrm{Dtrim}} = 0.0008$$

These values determine an aircraft drag = 116.1 kN
Thrust available (one engine), at speed V_2 = 161.5 kN
This provides for a climb gradient (OEI) = 0.0405
This is better than the airworthiness requirement of 0.024
To achieve this requirement would demand only a thrust to weight ratio of 0.254

Later in the design process, it will be necessary to determine the aircraft balanced field length (i.e. with one engine failing during the take-off run).

5.1.7.4 Constraint Analysis

The four performance estimates above have indicated that the original choice of aircraft design parameters (T/W, W/S) may not be well matched to the design requirements as each of the design constraints was easily exceeded. The assumed thrust and wing loadings were selected from data on existing aircraft in the literature survey. It seems that as our design specification is novel, this process is too crude for our aircraft. As we now have better knowledge of our aircraft geometry, it is possible to conduct a more sensitive constraint analysis. The methods described above will be used to determine the constraint boundaries on a T/W and W/S graph. The results are shown in Figure 5.1-10.

Moving the design point to the right and downwards makes the aircraft more efficient. The constraint graph shows that it would be possible to select a design point at $T/W = 0.3$ and $W/S = 500$ kg/sq. m (102.5 sq. ft). Recalculating the aircraft mass using the same method as above and with these new values gives

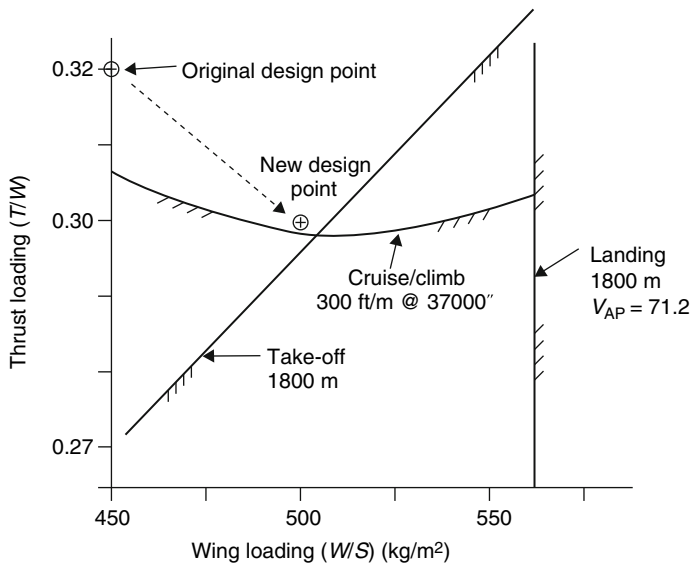


Figure 5.1-10 Constraint diagram.

Wing structure	= 11 387
Tail structures	= 2025
Body structure	= 10 050
Nacelle structure	= 2161
Landing gear	= 5088
Surface controls	= 1109
	STRUCTURE MASS = 31 538 (29.2%)
Propulsion mass	= 8520
Fixed equipment	= 10 800
	AIRCRAFT EMPTY MASS = 50 858 (47.1%)
	OPERTN EMPTY MASS = 52 862 (49.0%)
	ZERO FUEL MASS = 62 462 (57.8%)
Fuel mass = 45 538 kg (42%)	
	MAX. MASS (MTOM) = 108 000 (100%) (23 814 lb)

Using this mass and our new thrust and wing loading ratios gives

- Total engine thrust (static sea level) = 317.8 kN (71 450 lb)
- Gross wing area (reference area) = 216 sq. m (2322 sq. ft)

Assuming the wing tank dimensions are proportional to the wing linear size, the new wing area could accommodate 41 460 kg (91 400 lb) of fuel. This is less than predicted above (by 9 per cent). As we have made several assumptions and have not made a detailed analysis of the geometry and performance, we will delay the effect of this on the design of the wing until later in the design process.

5.1.7.5 Revised Performance Estimates

Range

With the cruise speed of 250 m/s (485 kt), assumed SFC of 0.55 force/force/hr, aircraft cruise L/D ratio of 17, initial mass (M_1) = MTOM (108 000 kg), and final mass (M_2) = ZFM (62 462 kg) gives

Range = 8209 nm

This is slightly longer than the previously estimated ESAR of 7988 nm but is within our calculation accuracy. The fuel ratio in the new design is 42.2 per cent whereas only 41.3 per cent is required; therefore, we have about 900 kg slack in the zero fuel estimation.

Cruise

With the new mass and geometry, the drag polar (start of cruise, 35 000 ft @ M0.85) is calculated as

$$C_D = 0.0148 + 0.0352C_L^2$$

At the start of cruise, the lift coefficient is 0.40; hence, $C_D = 0.0204$.

This equates to a drag = 53.1 kN (11 938 lb), and hence a cruise $L/D = 19.5$

The engine lapse rate at cruise is 0.197. Therefore, the available thrust at the cruise condition = 0.197 × 317.8 = 62.6 kN (14 073 lb)

This gives an engine setting in cruise of 85 per cent of the equivalent take-off rating

Cruise climb

Adding a climb rate of 300 fpm at the start of cruise makes the required thrust at the start of cruise = 59.4 kN (13 354 lb). This is 95 per cent of max. take-off thrust rating.

Landing

The approach speed is 64.5 m/s (125 kt)
This seems reasonable for regional airport operations
The landing distance is calculated as 1594 m (5225 ft)
This is well below the 1800 m design requirement

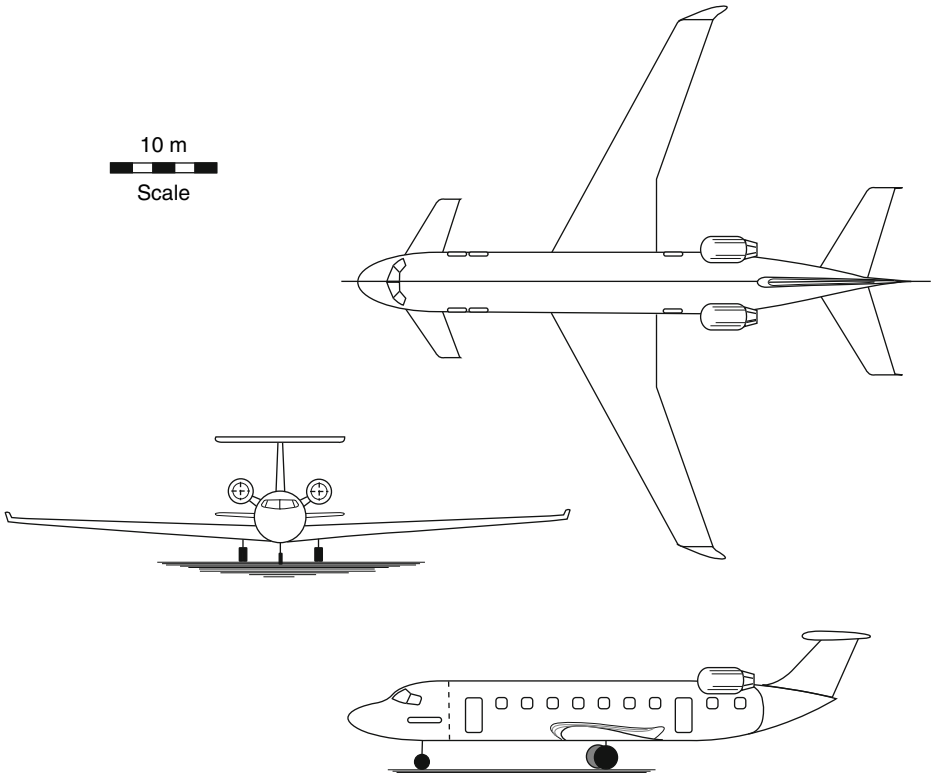


Figure 5.1-11 Refined baseline layout.

Take-off
The TOD is 1790 m (5869 ft)
The balanced field length is 1722 m (5647 ft)
These satisfy the design requirement of 1800 m
The second segment climb gradient (OEI) = 0.033
This satisfies the airworthiness requirement of 0.024

All of the design requirements have been achieved with the new aircraft geometry.
It is now possible to draw the refined GA of our aircraft, Figure 5.1-11.

5.1.7.6 Cost Estimations

Using the methods described in reference 1:

For an aircraft OEM = 52 862 kg, the aircraft purchase price will be \$42M (1995)
Assuming an inflation rate of 4 per cent per year
This brings the 2005 aircraft price = \$62M

For engines of about 40 000 lb TO thrust, the price would be \$4.0M (1995)

For two engines (2005 prices)	= \$12M
Airframe cost = \$62M − \$12M	= \$50M (i.e. aircraft priceless engines)
Assume 10% spares for airframe	= \$5.0M
Assume 30% for engine spares	= \$3.6M
Total investment	= \$70.6M

Assuming depreciation to 10 per cent over 20 years

Annual depreciation = (0.9 × 70.6)/20	= \$3.18 M
Assume interest on investment cost of 3.5% per yr	= \$2.47 M
Assume insurance 0.5% per yr of investment	= \$0.35 M
Total standing charges per year	= \$6.00 M

For the cruise range of 7000 nm at 485 kt, the flight time will be 14.4 hr
Add 0.75 hours to account for airport ground operations = 15.15 hr

Total block time = 15.15 hr

Some cost methods use this time in the calculation of DOC. Others use the flight time only. We will use the flight time in the calculations below.

Assume aircraft utilisation of 4200 hr per year (typical for long-range operations)

Standing charges per flying hour = \$1429

Crew costs (1995) per hr = 2 × 360 for flight crew + 4 × 90 for cabin crew
= \$1080 = \$1594 per hr (2005)

Landing and navigation charges per flight = 1.5 cents/kg MTOM = \$1620 per flight
Ground handling charge = \$3220 per flight
Total airport charges = \$4840 per flight = \$336 per flight hr
From the mass and range calculations: fuel used for ESAR (8209 nm) = 45 538 kg
Estimated fuel used for the 7000 nm design range = 40 300 kg
Assuming that little fuel is burnt in the ground,
Fuel used per flight hour = 40 300/14.4 = 2798 kg
Fuel volume = 2798/800 = 3.5 sq. m = 3500 litres = 3500/3.785 = 924 US gal
Assuming the price of fuel is 90 cents per gal,

Fuel cost = \$832 per flight hour

As maintenance costs are too difficult to assess at this time in the design process, we will assume them to account for 15 per cent of the total operating cost.

Total operating cost \$ per flight hour

Standing charges	= 1429
Crew cost	= 1594
Airport charges	= 336
Fuel cost	= 832
Maintenance costs	= (15%) (739)
	= \$4930 per flight hour

Hence DOC, Total stage cost = 4616 × 14.4 = \$70 996
Aircraft mile cost = 66 477/7000 = \$10.14
Seat mile cost (100% load factor) = 12.68 cents

Operators who lease the aircraft use 'cash DOC' to determine flight cost. They add the lease charges to their indirect costs as they are committed to this expense regardless of the aircraft utilisation. Cash DOC is calculated in the same way as above but without the standing charges. As aircraft maintenance is unaffected by the 'accountancy' method used to determine DOC, the cost is assumed to be the same as used above.

Cash DOC, Operational cost = \$3504 per hr
Total stage cost = \$50 451
Aircraft mile cost = \$7.21
Seat mile cost = 9.01 cents

Assuming that the ticket price (LHR–Tokyo) is \$4500 for the executive-class fare:
Revenue per flight (assuming 65 per cent load factor) = \$234 000
This compares favourably with the direct operating stage cost of \$70 996
Even allowing for a 100 per cent indirect operating cost (IOC) factor added to DOC, the operation would be viable

The SMCs calculated above are substantially larger than those quoted for high-capacity mixed-class services in which about 75 per cent of the seats are assigned to economy-class travellers. The revenue from such customers is significantly lower than from the executive class as they will be charged only about 20 per cent of the higher price fare. Without a detailed breakdown of the financial and accounting practices of an airline, it is impossible to determine the earning potential of the new service compared with the existing operation. However, the revenue assessment shown above is encouraging enough to continue with the project.

5.1.8 Trade-Off Studies

There are many different types of trade-off studies that could be undertaken at this stage in the design process. These range from simple sensitivity studies on the effect of a single parameter or design assumption, to extensive multi-variable optimisation methods. The studies shown below include trade-off plots that are used to determine the best choice of aircraft geometry. Wing loading and wing aspect ratio are chosen as the main trade-off parameters. These are regarded as the most significant design parameters for the short field, long-range requirements of the aircraft operation.

The studies shown in this section are presented as typical examples of the type of work appropriate at this stage of aircraft development. Many other combinations of aircraft parameters could have been selected and, in a full project analysis, would have been performed.

5.1.8.1 Alternative Roles and Layout

As mentioned in section 5.1.6.4 (fuselage layout), for all aircraft design studies it is necessary to consider the suitability of the aircraft to meet other operational roles. Although the principal objective of the project is to produce an efficient large business exclusive aircraft, we must also consider other mixed-class variants. In this way, a family of aircraft can be envisaged. This will increase the number of aircraft produced and reduce the design and development overhead per aircraft. Recognising this requirement, the fuselage diameter was designed to be suitable not only for the four abreast executive class seating but also for five and six abreast layouts of higher capacity options. The cabin length of 22 metres plus 4 metres for services and egress space is a fixed parameter and will control the layout and capacity of alternative roles. Within this length, various combinations of passenger layouts can be arranged. The position of doors and service modules (toilets, cupboards and galleys) is fixed but these can be used to provide natural dividers between classes.

From the previous fuselage layout drawing (Figure 5.1-8), the rear cabin is 6.5 metres, centre cabin 9.0 metres and front cabin 6.5 metres long. Using seat pitches of 1.1, 0.85, and 0.75 metres for executive/business, economy and charter classes respectively results in the layouts shown in Table 5.1-6.

For civil aircraft, it is common practice to stretch the fuselage in a later development phase. Typically, this may increase the payload by 35 per cent. Using this value (approximately), the single-class, economy version would grow to 160 passengers. At the 0.85 metre seat pitch, this would equate to a lengthening of the fuselage by 6.8 metres. To maintain aircraft balance a 2.8 m plug would be placed in the rear cabin and a 4.0 m plug forward of the wing joint. In this version the capacity of the aircraft would increase to the values shown below:

A Executive (single class) = 104 seats
B Mixed class = 141 seats (105 econ. and 36 exec.)
C Economy (single class) = 160 seats
D Charter (single class) = 204 seats

The extra capacity would require more passenger service modules and extra emergency exits to be arranged in the new cabin. This would reduce the space available for seating and slightly reduce the capacities shown above. Alternatively, the fuselage stretch would need to be increased by about a further 1.0 to 1.5 metres (40 to 60 in). Figure 5.1-12 shows some of the layout options described above.

Non-civil (military) versions of the aircraft could also be envisaged. With only 0.7 seat pitch for troop carrying a total of 186 soldiers could be carried in the original aircraft and 246 in the stretched version. The large volume cabin (for a small aircraft), the long endurance and the short field capabilities would be suitable for reconnaissance and electronic surveillance roles. In such operations, the reduced payload mass could allow extra fuselage fuel tanks to be carried to extend the aircraft duration. The high-speed, long-range performance could be useful for military transport

Table 5.1-6

		Rear cabin	Centre cabin	Front cabin	Total seats
A	Executive	24	32	24	80
B	Mixed*	35 econ.	50 econ.	24 exec.	85/24 = 109
C	Economy	35	50	35	120
D	Charter**	48	72	48	(168) = 150

*This provides 22 per cent business occupancy.
**The maximum capacity is reduced by about 10 per cent to account for extra spacing at emergency exits.

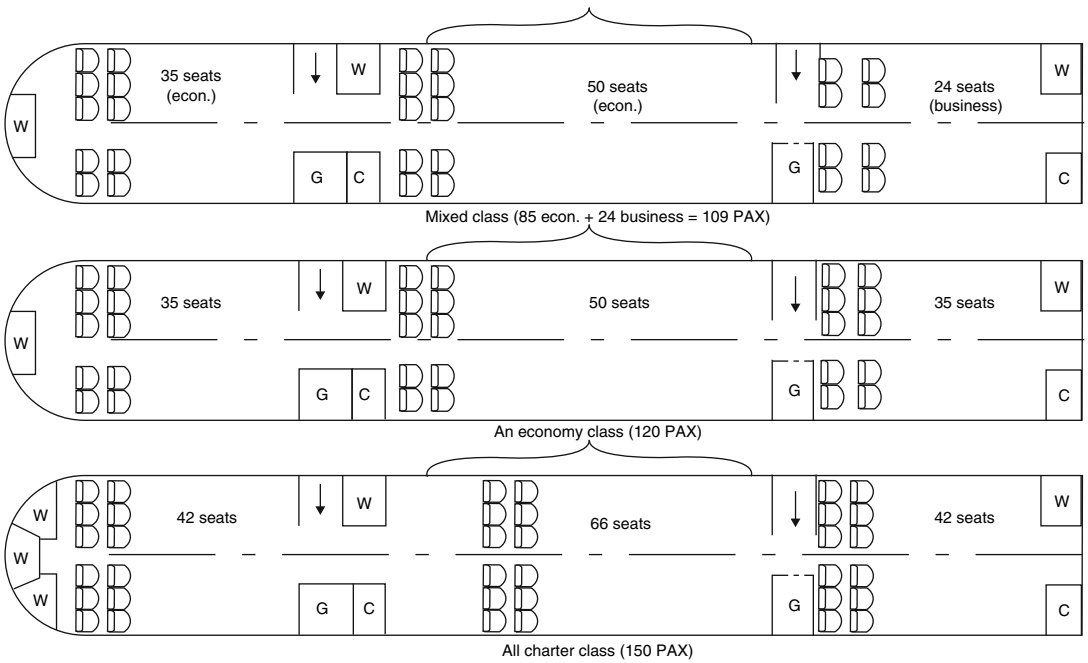


Figure 5.1-12 Fuselage development options (see also Figure 5.1-8 for all executive (baseline) layout).

command. Using this aircraft would avoid diplomatic complications caused by the need to refuel in foreign countries in conflict scenarios.

Many other versions of the aircraft may be envisaged (e.g. freighter/cargo, corporate jet, and communication platform) but these would not significantly affect the design of the current aircraft configuration.

5.1.8.2 Payload/Range Studies

For any aircraft design, it is uncommon to consider just the capability of the aircraft at the design point. Trading fuel for payload with the aircraft kept at the max. design mass results in a payload range diagram (shown in Figure 5.1-13).

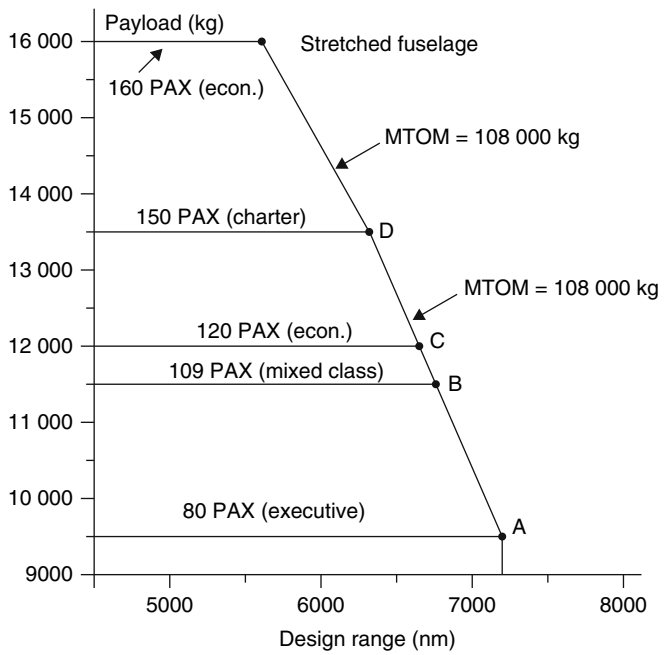


Figure 5.1-13 Payload/range diagram (developments).

Point A (the design point) shows the aircraft capable of flying 80 executive-class passengers over a 7000 nm range. Points B and C relate to the alternative cabin layouts described in the section above. At point B the payload is 11 380 kg. This means that 1780 kg of fuel is sacrificed for payload. At point C, 2400 kg of fuel is lost. Assuming the aerodynamic and engine efficiencies remain unchanged, the available range in these two cases reduces to 6811 and 6675 nm respectively. The reduction of about 530 nm in range for a 50 per cent increase in passenger number is a result of the low value of (M_{pay}/M_{TO}) on this design. Stretching the design to accommodate 160 passengers as described above would therefore be relatively straightforward on this design. This development is also shown on the payload range diagram. Even after allowing for an increase in structure and system mass of 1000 kg, the range in this configuration only reduces to 5625 nm.

As the aircraft is seen to be relatively insensitive to changes in payload, it is of interest to determine the effect of passenger load factor. Commercial aircraft do not always operate at the full payload condition. For this type of operation, an average load factor of 70 per cent is common. With less payload, the aircraft could increase fuel load (providing that space is available to accept the extra fuel volume). At 70 per cent passenger load factor with extra fuel, the Breguet equation gives an increase in range of 668 nm. If extra space is not available, the aircraft at 70 per cent load factor and with normal fuel load would be able to fly a stage length of about 7500 nm. The sensitivity of the range calculation to passenger load factor raises the question of the choice of the realistic design payload. Designing for the 0.7 × 80 passenger load would significantly reduce the aircraft max. take-off mass and fuel load. This would considerably reduce the stage cost and aircraft price.

The payload/range study has shown that the loading conditions around the design point must be carefully considered. As the effect of range and associated fuel is uncharacteristically sensitive on this aircraft, it is important to reconsider the original design specification to account for this aspect.

5.1.8.3 Field Performance Studies

The evaluation of field performance calculated earlier (section 5.1.7.3) was concerned with the aircraft at the design condition only. The predictions can be recalculated for variations in the aircraft parameters. In the earlier work, the take-off performance was seen to be well matched to the 1800 m requirement. However, it is also necessary to understand the sensitivity of the calculation to changes in the main, aircraft design parameters (e.g. thrust and wing loadings). A carpet plot can be constructed to show these effects (Figure 5.1-14).

Note: none of the study points achieves the 1400 m field length originally considered (section 5.1.3.3). The aircraft thrust loading is shown most influential in reducing TOD. To investigate this further, a second trade-off study has been conducted. Keeping the wing loading constant, the wing max. lift coefficient and thrust loading have been varied. The results are shown in Figure 5.1-15.

The carpet plot shows that increasing the thrust loading to 0.32 would allow a reduction in take-off lift coefficient to 2.05. This would reduce wing structure complexity and thereby wing mass. Obviously, an increase in engine thrust would also involve a corresponding increase in propulsion group mass. A more detailed study would need to be done later in the design process to draw a firm conclusion to this trade-off.

In the revised aircraft layout, the landing performance was shown to be well within the 1800 m design constraint. This suggests that changes could be considered to the aircraft. For a fixed wing area the two parameters that have an

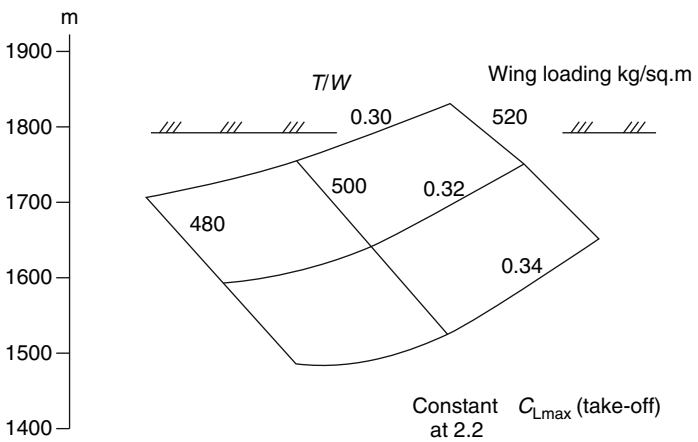


Figure 5.1-14 Take-off distance study (*T/W* and *W/S*).

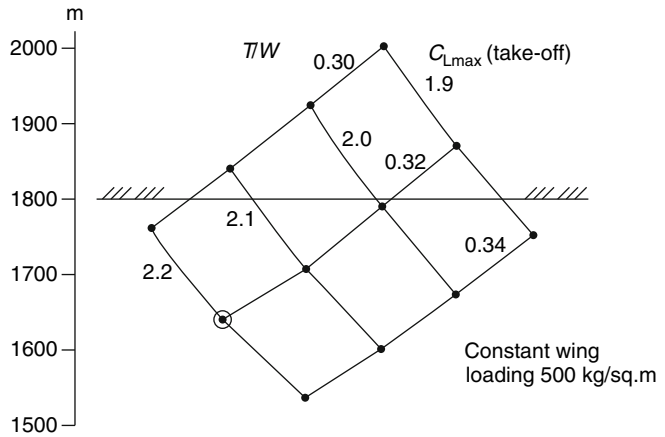


Figure 5.1-15 Take-off distance study (T/W and C_{Lmax}).

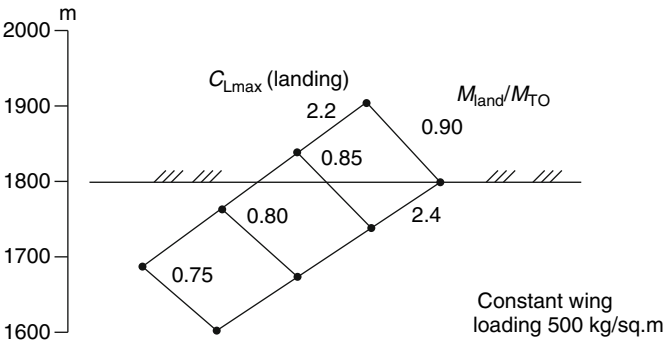


Figure 5.1-16 Landing distance study.

effect on the landing performance are aircraft max. lift coefficient (in the landing configuration) and the landing mass ratio ($M_{landing}/M_{TO}$). The trade-off study results are shown in Figure 5.1-16.

This shows that the aircraft is capable of landing in the 1800 m field at 90 per cent MTOM. The max. landing lift coefficient could be reduced to 2.2 and still allow a 82 per cent MTOM landing mass. This lift coefficient is the same as previously used for the take-off condition; therefore, a further set of trade-off studies should be done to select the best combination of lift coefficients for take-off and landing. This would fix the flap type and deflection angles to give the optimum design combination. To do this, more detailed aerodynamic analysis is required than is available at this stage in the development of the aircraft.

5.1.8.4 Wing Geometry Studies

To conduct a full and accurate analysis of the wing parameters (e.g. area and aspect ratio) would involve a full, multivariate optimisation method. As most of the design parameters are interconnected, this would be a complex process. At this early stage in the design process and with limited resource and time available, it is not possible to undertake such a comprehensive study. Some simplifying assumptions are necessary to enable a sensitivity study to be done. For example, we may assume that the engine parameters are kept constant and that the less sensitive mass components (e.g. surface controls, systems, etc.) are held constant or considered to be directly proportional to MTOM. With such assumptions, the results of the study can only be used to indicate the sensitivity of variations to the design and the direction of possible changes to the existing configuration.

A study has been completed around the current aircraft configuration to investigate the effect of changes to wing area and aspect ratio on some of the aerodynamic and mass parameters. A series of carpet plots (Figures 5.1-18 to 5.1-23) illustrate the results of the study. To allow comparison to the earlier work, the study used wing loading to represent area variations. The resulting wing area values are plotted in the carpet plots. Wing loading (kg/sq. m) and aspect ratio values selected for the study are shown below:

Wing loading	400, 450, 500, 550
Aspect ratio	8, 10, 12, 14

To appreciate the geometrical implications of these changes, the extreme layouts (400/14 and 550/8) together with the existing baseline configuration (500/10) are illustrated in Figure 5.1-17. (Note: the drawings of these aircraft are illustrative only as they have not been balanced.)

As expected, the study shows that increasing the size of the wing (i.e. reducing wing loading) and/or increasing aspect ratio increases the wing mass. Figure 5.1-18 illustrates these effects clearly and provides quantitative data of the mass changes around the design point (500/10). The increasing slope of the aspect ratio lines shows the progressive mass penalty, especially for the larger area wings.

Wing mass, although important, represents only a component of aircraft mass. The combined effect on aircraft empty mass is illustrated in Figure 5.1-19. Although a similar pattern is seen on this plot, the changes represent a smaller proportion (about a quarter of the previous percentage values). For example, moving from the design point to point 550/8 is shown to reduce wing mass by about 35 per cent but the empty mass is reduced by 8 per cent. Note: the wing loading of 550 was shown to violate the original take-off constraint (Figure 5.1-10). Making this move would require the take-off and possibly the climb performance to be reconsidered.

Making the wing smaller and increasing aspect ratio has a significant effect on both parasitic and induced drag. Both will be reduced. Figure 5.1-20, which plots aircraft lift/drag ratio, shows how the aerodynamic efficiency of the aircraft is improved. Note that the design point shows a value higher than that originally assumed (i.e. $L/D = 17$). Over the range of geometrical changes investigated the L/D ratio varies between 16 and 21. This is a significant variation that shows the sensitivity of choice of wing geometry.

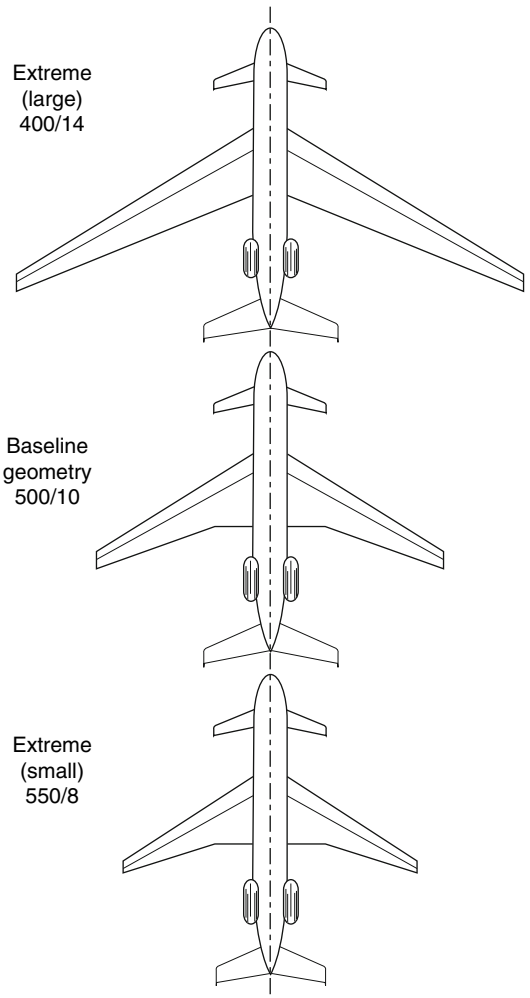


Figure 5.1-17 Geometrical variations.

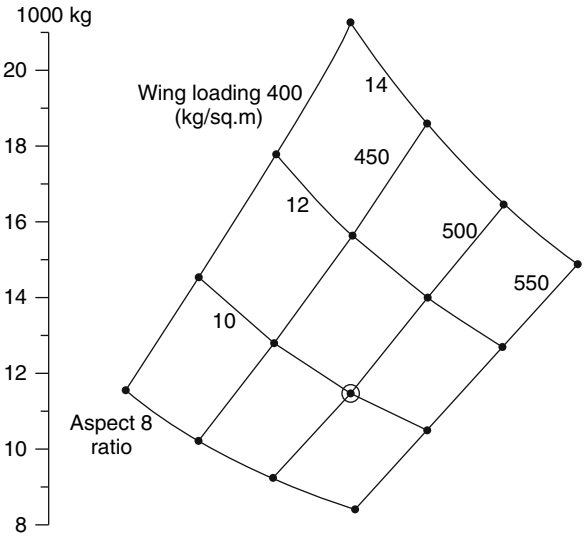


Figure 5.1-18 Trade-off study: wing mass.

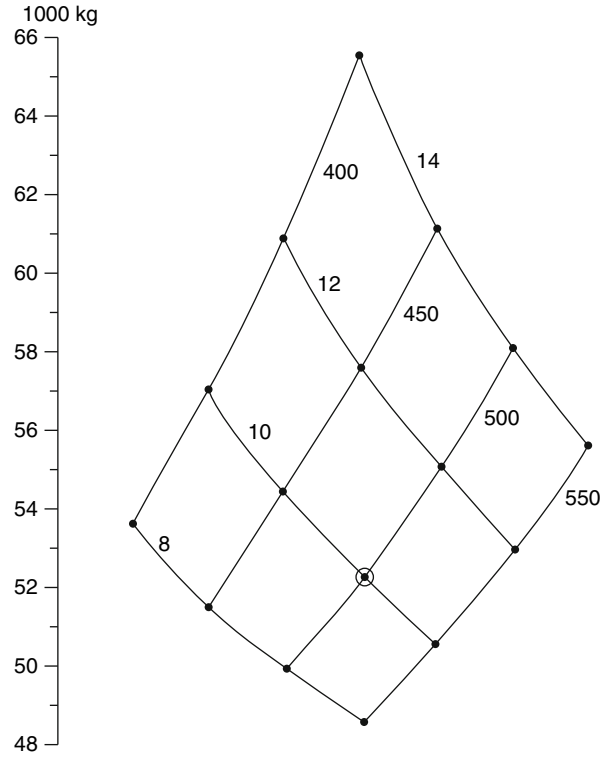


Figure 5.1-19 Trade-off study: aircraft empty mass.

The aircraft L/D ratio and max. take-off mass (discussed below) are important parameters in the calculation of the required fuel mass to fly the 7000 nm stage length. Assuming that the cruise speed and engine specific fuel consumption remain unchanged from their previous values, the resulting fuel mass calculations are shown in Figure 5.1-21. At each of the wing loading lines the 'optimum' aspect ratio value moves progressively from about 9 for the large wing to 14 for the smallest wing. At the design wing loading of 500, there appears to be a small advantage to increasing aspect ratio from the design value of 10 to 12. Extending to 14 is not seen to be worthwhile.

Aircraft take-off mass (MTOM) is dependent on both structure mass and fuel mass. The studies above have shown that the wing geometrical changes may increase structure mass but then reduce fuel mass. The combined effect is shown in the MTOM carpet plots (Figure 5.1-22). Due to the reduced fuel mass, the significance of the structure

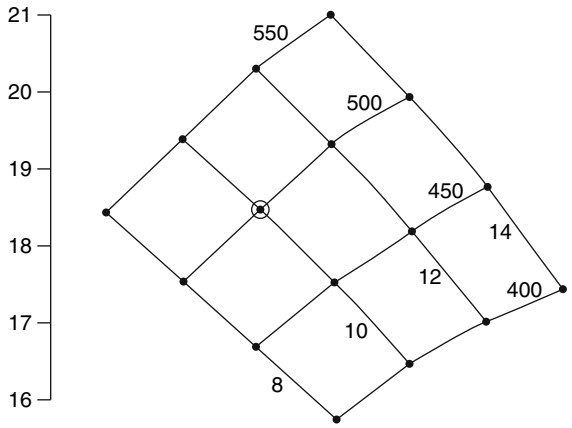


Figure 5.1-20 Trade-off study: cruise *L/D* ratio.

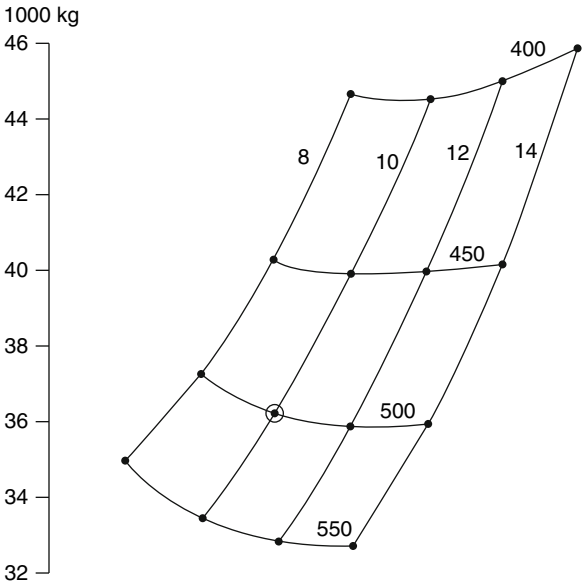


Figure 5.1-21 Trade-off study: stage fuel mass.

mass changes is eroded but the overall pattern remains similar to the empty mass plots discussed earlier. A move from the design point to a wing loading of 550 kg/m^2 (112.6 lb/ft^2) and aspect ratio of 8 (i.e. 550/8) would reduce MTOM by about 5 per cent. This is a significant reduction and is worth investigating further if the economic studies described below confirm this advantage.

Wing area is a function of wing loading and MTOM. To show the dimensional effects of the changes in these parameters the absolute values for wing area have been plotted (Figure 5.1-23). Note the significance of aspect ratio on the larger wings and the relatively low sensitivity for small wings.

5.1.8.5 Economic Analysis

The results from the studies above can be used, together with operational data, to assess the economic viability and sensitivity to the aircraft geometrical changes. The aircraft price is related to the aircraft empty mass and engine size. The cost of fuel is proportional to fuel mass. Other operational costs are related to aircraft take-off mass. Hence, changes to the aircraft configuration will affect both aircraft selling price and operating costs. For civil aircraft designs, these two cost parameters are often selected as the principal design drivers (optimising criteria). Although the aircraft configuration may not be selected at the optimum configuration for these parameters, the design team will need to know what penalty they are incurring for designs of different configuration.

All of the cost calculations have been normalised to year 2005 dollars by applying an inflation index based on consumer prices. Several separate cost studies have been performed as described below.

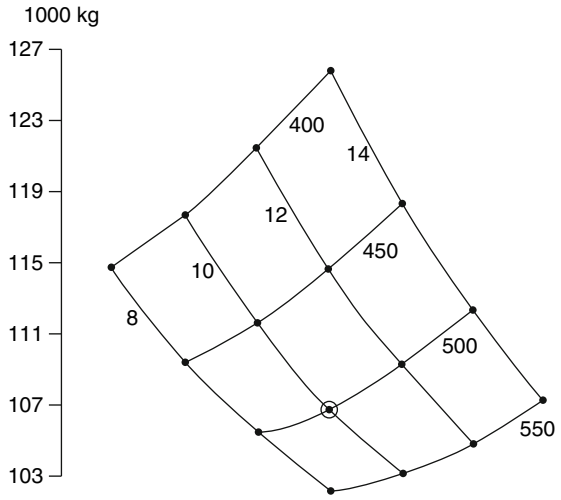


Figure 5.1-22 Trade-off study: aircraft max. TO mass.

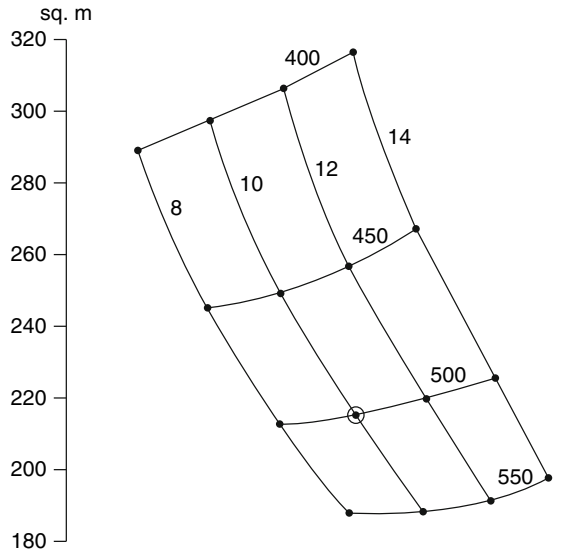


Figure 5.1-23 Trade-off study: wing area.

Aircraft price

Aircraft price is one component in the evaluation of total investment. This includes the cost of airframe and engine spares. For this aircraft, the total investment is about 12 per cent higher than the aircraft price.

Figure 5.1-24 shows the variation of aircraft price for the geometrical changes considered previously. At the design point the price is estimated to be $70.5 m. The carpet plot shows that this price would fall by about 5 per cent if the aspect ratio was reduced to 8, and by about 9 per cent if the configuration was moved to point 550/8. The effect of reducing wing loading progressively increases aircraft price (e.g. reducing wing loading to 400 kg/sq. m increases the price by 9 per cent). Similarly, increasing wing aspect ratio increases price (e.g. moving from 10 to 14 increases the price by over 10 per cent). Without consideration of other operating costs, the main conclusion of this study is to move the design point to lower values of both wing loading and aspect ratio.

DOC per flight

There are two fundamentally different methods of estimating aircraft DOC. The traditional method includes the depreciation costs of owning the aircraft. On this aircraft, this would be about 33 per cent of the total DOC. If the aircraft operator leases the aircraft, the annual cost of the aircraft is regarded as a capital expenditure. This would be considered as an indirect aircraft operating cost. In this case, the aircraft standing charges (depreciation, interest and

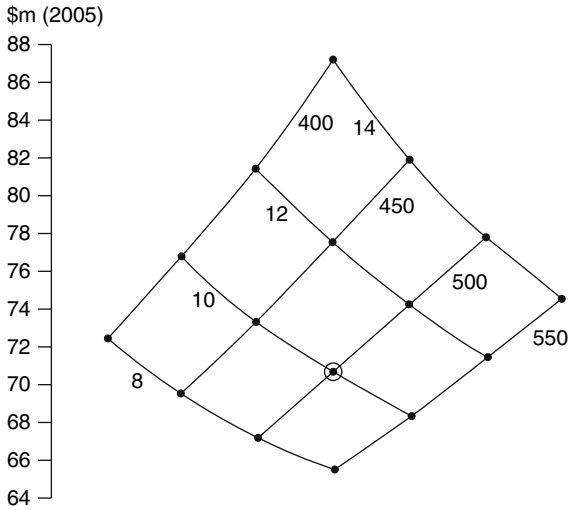


Figure 5.1-24 Trade-off study: aircraft price.

insurance) are not included in the calculation and the resulting cost parameter is termed 'Cash DOC'. It is important to calculate both of the DOC methods. The results of the DOC calculations are shown in Figures 5.1-25 and 5.1-26.

The DOC per flight at the design point (500/10) is $72 740. This figure would be reduced by 3 per cent if the design was moved to point 550/8 and still satisfy the technical design requirements. Increasing wing area and/or aspect ratio from the design point is not seen to be advantageous. At the design point the cash DOC is estimated to be $49 470. In this case, curves are seen to be flatter than for the full DOC values. This results in optimum points for aspect ratio. At the design point, the existing value of aspect ratio is seen to be optimum. Moving to the higher wing loading (550), if feasible, would reduce cash DOC by about 2 per cent.

It is of interest to note that the design conclusions from the two DOC methods are different. This implies that the design strategy to be adopted is conditional on the accounting practices used by the operator. This is a good example of the need for the designers to understand the total operating and business environment in order to select the best aircraft configuration.

Seat mile cost

The cost of flying the specified stage (design range) is dependent on the payload. In the DOC calculations above, the aircraft has been assumed to be operating at full payload. This is conventional practice as it allows the maximum seats

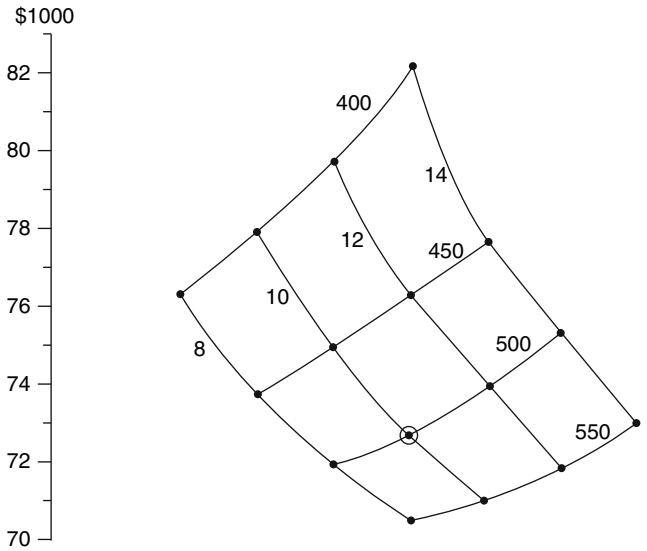


Figure 5.1-25 Trade-off study: DOC per flight.

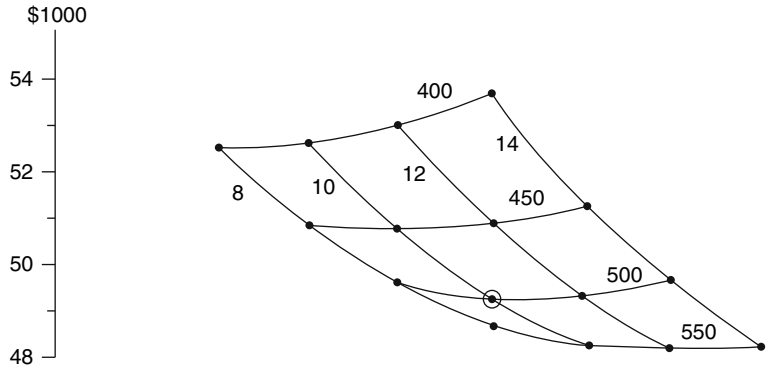


Figure 5.1-26 Trade-off study: cash DOC per flight.

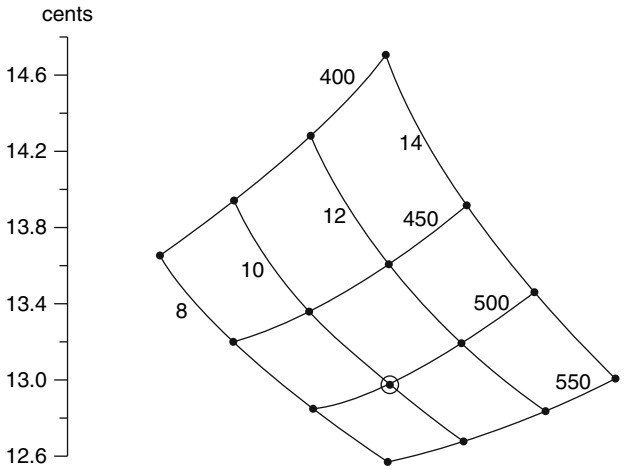


Figure 5.1-27 Trade-off study: seat mile cost (SMC).

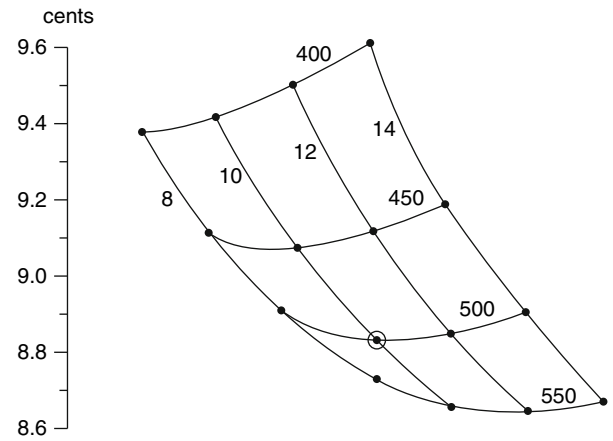


Figure 5.1-28 Trade-off study: cash SMC.

to be used in the evaluation of seat mile costs (SMCs). Flying at max MTOM, the DOC per flight does not vary with passenger numbers. The SMC shown in Figures 5.1-27 and 5.1-28 (for DOC and Cash DOC respectively) is evaluated for the 80-seat executive version of the aircraft.

Other versions have been evaluated at the design point and are listed in Table 5.1-7.

Table 5.1-7

Layout	Passengers	SMC	Cash SMC
Executive only	80	12.99	8.83
Mixed class	107	9.74	6.62
Economy only	120	8.70	5.91
Charter only	150	6.85	4.68

Note the powerful effect of passenger numbers in reducing SMC and the substantial reduction in the cash SMC method. When using values from other aircraft, it is important to know the basis on which cost data has been calculated.

The SMC and cash SMC carpet plots show a similar pattern to the DOC figures. The SMC curves show an advantage to reducing aspect ratio for all values of wing loading but this advantage reduces with higher wing loadings. For the design point, a reduction of aspect ratio to 8 would be recommended. For the cash SMC calculation, the best aspect ratio varies from 8 at low loading to 12 at the high loading. For the design point, the value of 10 is seen to be about 'optimum'.

5.1.9 Initial 'Type Specification'

At the end of the initial concept stage it is important to record all of the known details of the current design. This document forms the initial draft of the aircraft type specification. As the design evolves over subsequent stages, this document will be amended and enlarged until it forms the definitive description of the final configuration. The initial draft will form the input data for the next stage of the design process. The sections below are typical of a professional aircraft specification.

5.1.9.1 General Aircraft Description

This aircraft is designed for exclusive, business/executive, long-range routes from regional airports. Although apparently conventional in configuration, it incorporates several advanced technology features. These include natural laminar flow control, composite material and construction, enhanced passenger cabin services and comfort standards, and three-surface control and stability. The single aisle cabin layout is arranged to accommodate four abreast seating for the baseline executive configuration. In other configurations it will provide five abreast economy-class seating and six abreast charter operations. In these versions the increased passenger numbers reduce the range capability of the aircraft. In the baseline executive layout, space for 80 sleeperettes at 1.1 m (44 in) pitch is available. For the other layouts, 120 economy seats at 0.8 m (32 in) pitch or 150 charter seats at 0.7 m (28 in) pitch are feasible. Toilet, galley and wardrobe provision is adjusted to suit the layout using fixed service facilities in the cabin. Emergency exits and other safety provisions meet FAR/JAR requirements. Underfloor cargo and baggage holds are positioned fore and aft of the wing/fuselage junction structure.

To reduce engine operating noise intrusion into the cabin, during the long endurance flights, the engines are positioned at the rear of the fuselage, behind the cabin pressure bulkhead. Several existing and some proposed new engine developments are suitable to power the aircraft. This provides commercial competitiveness and flexibility to the potential airline customers. All the engines are modern, medium-bypass (typically 6.0) turbofans offering efficient fuel economy.

The modern, aerodynamically efficient, high aspect ratio wing layout provides good cruise efficiency. A lift to drag ratio in cruise of 19 is partly achieved due to the aerodynamic section profiling and the provision of natural laminar flow. Leading and trailing edge, high-lift devices provide the short field performance required for operation from regional airfields.

The three-surface (canard, main wing and tail) layout offers a reduction to trim drag in cruise and improved ride comfort. Integrated flight control systems with fly-by-wire actuation to multi-redundant electric/hydraulic controllers provide high levels of reliability and safety.

Aircraft manufacture combines established high-strength metallic materials with new composite construction techniques. The combination of conventional and novel structural and manufacturing practices offers reduced structural weight with confidence.

5.1.9.2 Aircraft Geometry

Principal dimensions		
Overall length	= 43.0 m,	141 ft
Overall height	= 13.0 m,	42.6 ft
Wing span (total)	= 48.0 m,	157 ft
Main wing		
Gross (ref.) area	= 216 sq. m,	2322 sq. ft
Aspect ratio	= 10	
Sweepback (LE)	= 22°	
Mean chord	=4.65 m,	15.25 ft
Taper ratio	= 0.3	
Thickness (mean)	= 11%	
Control surfaces		
Horizontal tail area	= 20.0 sq. m,	215 sq. ft
Vertical (fin) area	= 15.5 sq. m,	167 sq. ft
Canard area	= 7.0 sq. m,	75 sq. ft
Fuselage/cabin		
Fuselage length	= 40.0 m,	131 ft
Cabin outside dia.	= 3.6 m,	142 in
Pass. cabin length	= 22.0 m,	72 ft
Landing gear		
Wheelbase	= 18.0 m,	59 ft
Track	= 8.25 m,	27 ft

Engines (two)

Various types, static SL thrust (each) = 160 kN, 35 700 lb

5.1.9.3 Mass (Weight) and Performance Statements

Mass statement

Aircraft empty mass	50 858 kg,	112 142 lb
Aircraft operational mass	52 862 kg,	116 560 lb
Aircraft max. (design) mass	108 000 kg,	238 140 lb
Baseline (executive) version (80 PAX)		
Zero fuel mass	62 462 kg,	137 729 lb
Payload	9600 kg,	21 168 lb
Fuel load	45 538 kg,	100 411 lb
Still air range	7200 nm	
Mixed-class version (107 PAX)		
Payload	11 380 kg,	25 093 lb
Fuel	43 758 kg,	96 486 lb
Still air range	6810 nm	
All economy version (120 PAX)		
Payload	12 160 kg,	26 812 lb
Fuel	44 138 kg,	97 324 lb
Still air range	6675 nm	
Charter version (150 PAX)		
Payload	13 450 kg,	29 657 lb
Fuel	41 688 kg,	91 922 lb
Still air range	6350 nm	
Stretched version (160 economy PAX)		
Payload	16 000 kg,	35 280 lb
Still air range	5600 nm	

Performance statement

(baseline version with 80 PAX and fuel):

Cruise speed	= M0.85
Cruise altitude	= 36 000 ft
Initial cruise altitude climb rate	= 300 fpm
Take-off distance	= 1790 m (5870 ft)
Balanced field length	= 1720 m (5670 ft)
Second segment climb gradient	= 0.033
Approach speed	= 125 kt
Landing distance	= 1594 m (5225 ft)

5.1.9.4 Economic and Operational Issues

Cost statement

(baseline aircraft, 2005 US dollars)

Aircraft price	= $62.0 M
Total investment/aircraft	= $70.6 M
Standing charges/yr	= $6.0 M
Standing charges/flt hr	= $1430
DOC/fl hr	= $4930
Stage cost	= $71 000
Aircraft mile cost	= $10.14
SMC (100% PAX)	= 12.7 cents
Cash DOC/flt hr	= $3500
Total stage cash cost	= $50 450
Aircraft cash mile cost	= $7.20
Cash SMC (100%)	= 9.0 cents

Operational statement

The aircraft is capable of stretching to accommodate up to 204 charter seats. In a military role the baseline aircraft can seat 186 soldiers and in the stretched version 246 troops. In each of these versions it would be possible to fly 7000 nm (unrefuelled).

Other roles for the aircraft could include:

- Civil corporate jet
- Freighter
- Military refuelling tanker
- Communication platform
- Military surveillance aircraft
- Military supply aircraft

These versions of the aircraft have not been considered in the overall geometrical layout of the aircraft in the initial design process. A short study would be appropriate when the initial baseline study has been completed to identify any small changes to the aircraft layout to accommodate any of the above roles.

5.1.10 Study Review

This aircraft project has shown how, for a relatively simple aircraft, the design process is taken from the initial consideration of the operational requirements to the end of the concept design phase. The intervening stages have shown how the aircraft design evolves during this process. This showed that the initial configurational assumptions for thrust and wing loadings, based on data from existing aircraft, were found to be in error because of the unique operational performance of the aircraft. A more efficient aircraft layout was identified. Even the revised configuration was shown capable of improvement by the trade studies. For most aircraft projects, this iterative process is commonplace.

The economic assessment of the aircraft indicated that the project was viable and therefore worth taking into the next stage of development.

Due to time and resource restrictions in the conceptual stage, several technical aspects of the design have not been fully analysed. These include:

- The stability and control analysis of the aircraft including the assessment of the effect of the three-surface control layout.
- The aerodynamic analysis of the laminar flow control system and the associated structural and system requirements.
- The aircraft structural analysis and the realisation of the combined conventional and composite structural framework.
- The aircraft systems definition and the associated requirements for the new executive-class communication and computing facilities.
- The special requirements for aircraft servicing and handling at regional airports.
- The detailed trade-off studies applied to the field requirements (e.g. the definition of aerodynamic (flap design and deflection), propulsion (T/W), structures, systems and costs).
- The assessment of the overall market feasibility of the project.

Each of the topics in the list above involves work that is either comparable with or exceeds the work that has already been done on the project. In industry, progressing to the next stage of aircraft development would involve a 20- to 50-fold increase in technical manpower. To commit the company to this expenditure is a significant investment. A decision to proceed would only be taken after discussions with potential airline customers.

If the type of operation envisaged by this project is seen to be attractive, it will stimulate competition. This may come from airframe manufacturers who could modify existing aircraft to meet the specification. It is essential that the design team of the new aircraft anticipate this threat. They will need to conduct their own studies on the modifications to the aircraft that may be used as competitors. These are studies that require substantial effort, but in completing them, the advantages of the new design can be identified. This information will be useful to the technical sales team of the new aircraft and used to counteract the threat from the 'old-technology', 'modified' existing types.

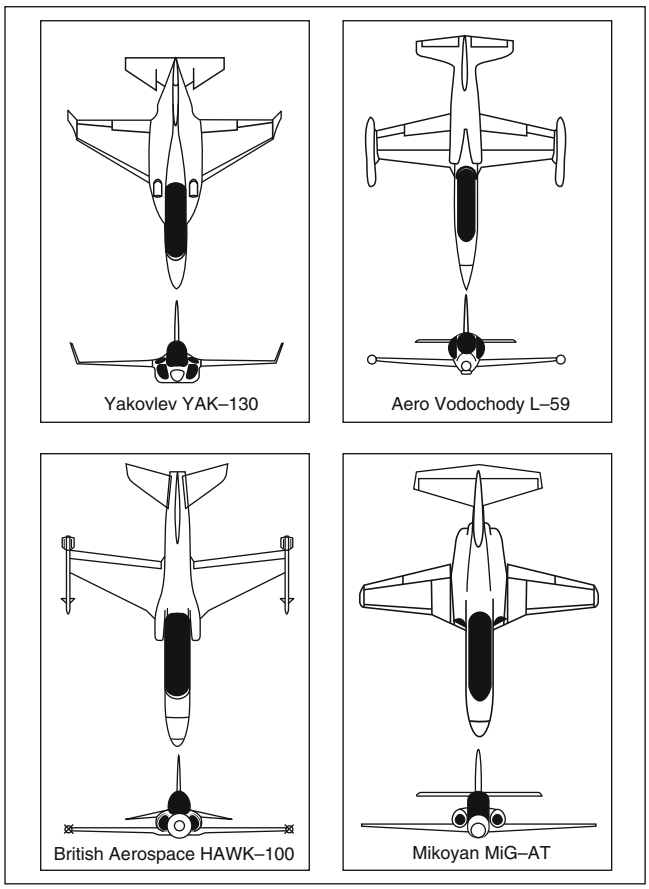
References

1. Jenkinson, L. R., Simpkin, P. and Rhodes, D., *Civil Jet Aircraft Design*, 1999, AIAA Education Series, ISBN 1-56347-350-X.
2. Mason, W. H. et al., *Low-cost Commercial Transport – Undergraduate Team Aircraft Design Competition*, 1995, Virginia Tech. AIAA 95–3917.
3. Howe, D., *Aircraft Conceptual Design Synthesis*, 2000, Prof. Eng. Pub. Ltd, ISBN 1-86058-301-6.
4. Fielding, J. P., *Introduction to Aircraft Design*, 2000, Cambridge University Press, ISBN 0-502-65722-9.
5. Hoerner, S. F., *Fluid Dynamic Drag*, published by the author, Bricktown, NJ, 1965.

Aerospace Engineering Desk Reference; ISBN 9781856175753

5.2 Military Trainer

Lloyd Jenkinson and Jim Marchman



5.2.1 Introduction

A project similar to the one described below was the subject of a EuroAVIA design workshop sponsored by British Aerospace. Undergraduate students from ten European countries worked for three weeks in separate teams to produce specifications for new training systems. The study below represents a combination of the results from this workshop and some subsequent design work done on aeronautical courses in two English universities. Acknowledgement is given to all the students who worked on these projects for their effort and enthusiasm which contributed to the study described.

In the following analysis general references are made to aircraft design textbooks.[1–5] To avoid confusion in the text, a list of current popular textbooks, useful for this project, is included in the reference section at the end of this chapter.

5.2.2 Project Brief

All countries with a national air force need an associated programme for their pilot selection and training; therefore the commercial market for military training aircraft and systems is large. Designing training aircraft is relatively straightforward as the technologies to be incorporated into the design are generally well established. Many countries have produced indigenous aircraft for training as a means of starting their own aircraft design and manufacturing industry. This has generated many different types of training aircraft in the world. For many different reasons only a few of these designs have been commercially successful in the international market. The British Aerospace Hawk (Figure 5.2-1) family of aircraft has become one of the best selling types in the world with over 700 aircraft sold. It is a tribute to the original designers that this aircraft, which was conceived over 25 years ago, is still in demand. The maturity of the Hawk design is not untypical of most of the other successful trainers. Only recently have new aircraft been produced (mainly in East European countries) but these are still unproven designs and not yet competitive with the older established products.

Since the early 1970s when the Hawk and other European trainers were designed, front-line combat aircraft operation has changed significantly. The introduction of higher speed, more agile manoeuvring, stealth, together with significant developments in aircraft and weapon systems generated a requirement for a new training system. As airframe and system development is expensive it is essential that an overall systems approach is adopted to this project.

The project brief for a new training system covers pilot training and selection from the *ab-initio* phase (assuming cadets have had 50 hours' flight training on a light propeller aircraft) to the start of the operational (lead-in) training on twin-seat variants of combat aircraft. This period covers the existing basic and advanced training phases covered by Hawk type aircraft. To represent modern fighter capabilities the new training system should also include higher flight performance and weapon system training which is not feasible on current (older) training aircraft.

The concepts to be considered are those associated with an integrated training system. This must account for the various levels of capability from the aircraft, synthetic training systems (including simulators) and other ground-based facilities. It will be necessary to define the nature of the training experiences assigned to each component of the overall training system.

The minimum design requirements for the aircraft are set out in the aircraft requirements section below but consideration should be given to the development of the training programme to include flight profiles with transonic/supersonic performance. Also, as all commercially successful training aircraft have been developed into combat derivatives, this aspect must be examined. To reduce the overall cost of the project to individual nations attention must be given to the possibility of multinational co-operative programmes. All the issues above will be influential in the choice of design requirements for the aircraft.

5.2.2.1 Aircraft Requirements

Performance

General	Atmosphere	max. ISA + 20°C to 11 km (36 065 5 ft)
		min. ISA − 20°C to 1.5 km (4920 ft)
	Flight missions – see separate tables	
	Max. operating speed,	V_{mo} = 450 kt @ SL (clean)
		V_{mo} = 180 kt @ SL (u/c and flaps down)

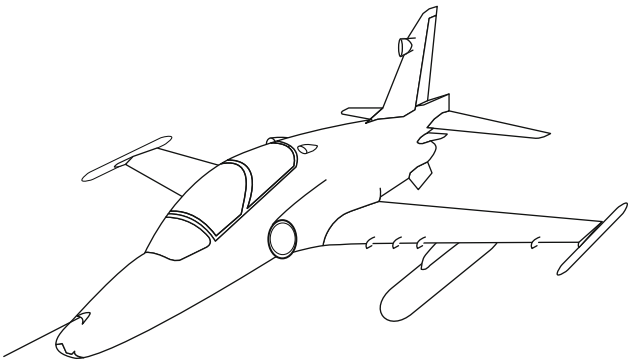


Figure 5.2-1 Hawk aircraft.

Turning	Max. sustained g @ SL = 4.0 Max. sustained g @ FL250 = 2.0 Max. sustained turn rate @ SL = 14°/s Max. instantaneous turn rate @ SL = 18°/s
Field	Approach speed = 100 kt (SL/ISA) TO and landing ground runs = 610 m (2000 ft) Cross-wind capability = 25 kt (30 kt desirable) Canopy open to 40 kt Nose wheel steering
Miscellaneous	Service ceiling > 12.2 km (40 000 ft) Climb – 7 min SL to FL250 (note: one flight level, FL = 100 ft) Descent – 5 min FL250 to FL20 (15° max. nose down) Ferry range = 1000 nm (2000 nm (with ext. tanks)) Inverted flight = 60 s

Structural

- Flight envelope $n_1 = +7$, $n_3 = -3$
- Max. design speed M0.8
- $V_D > 500$ kt CAS
- Utilisation = 500 h/year
- Fatigue life = 30 yr

Operational

- Hard points = 2 @ 500 lb (227 kg) plus 2 @ 1000 lb (453 kg), all wet
- Consideration for fully armed derivatives
- Consideration for gun pod installation
- Provision for air-to-air refuelling

Cockpit

- Aircrew size – max. male 95 per cent, min. female 50 per cent
- Ejection – zero/zero
- All weather plus night operations
- Cockpit temperature, 15–25°C
- Oxygen system

Systems

- Avionics to match current/near future standards
- Consideration given to fly-by-wire FCS
- Consideration given to digital engine control
- Glass cockpit
- Compatibility to third and fourth generation fast jet systems where feasible

5.2.2.2 Mission Profiles

Mission profiles used in the design of the aircraft are to be defined by the design team but they must not have less capability than described below:

1. **Basic** This is to represent early stages of the flight training. Two sorties are to be flown without intermediate refuelling or other servicing.

Phase	*Description*	*Height*	*Time (min)*
1	Start, taxi	SL	4
2	Take-off	FL20	1
3	Max. climb	FL250	7
4	Cruise to training area	FL250	6
5	High-speed decent	FL20	5
6	General handling (Buffet control, etc.)	FL20	10
7	Max. climb	FL250	6
8	Manoeuvres (Turns, spin, etc.)	FL250	4
9	Cruise to base	FL250	5
10	Descent	FL20	5
11	Recover to base*	FL 20	3
	– 100 nm fuel + 5% reserve or		
	– 5 circuits + 10% reserves		
12	Landing, taxi, shutdown	SL	4
		Mission elapsed time	**60**

*(reserve fuel is only applicable to the second sortie)

2. **Advanced** This mission is typical of fighter handling at the advanced training stage.

Phase	*Description*	*Height*	*Time (min)*
1	Start, taxi	SL	4
2	Take-off	FL20	1
3	Max. climb	FL250	7
4	Cruise to training area	FL250	6
5	Weapon training	FL250	10
6	Aerobatics and high *g*	FL50	10
7	Low-level flying	250 ft	10
8	Climb to cruise	FL250	7
9	Cruise to base	FL250	6
10	High-speed descent	FL20	5
11	Recover to base*	FL20	6
	– 100 nm fuel + 5% reserve or		
	– 5 circuits + 10% reserves		
12	Landing, taxi, shutdown	SL	4
		Mission elapsed time	**76**

(*reserve fuel is only applicable to the second sortie)

Note: the times quoted in the above profiles are approximate and do not define aircraft performance requirements. (FL = flight level, 1FL = 100 ft.)

3. **Ferry** This mission is required to position aircraft at alternative bases. The ferry range is specified in section 5.2.2.1. The ferry cruise segment may be flown at best economic speed and height. Reserves at the end of the ferry mission should be equivalent to that for the basic mission profile.

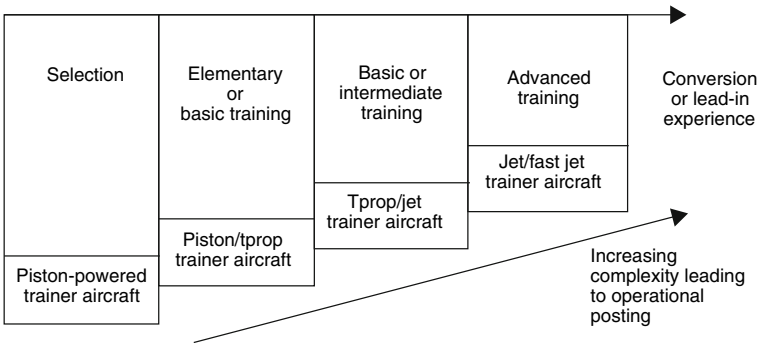


Figure 5.2-2 Air force flight training phases.

5.2.3 Problem Definition

The main difficulty with this project lies is the broad spectrum of training activities that are expected to be addressed by the system. To cover all flight training from post- *ab-initio* to pre-lead-in will include the basic, intermediate and advanced training phases (Figure 5.2-2). In most air forces this involves the use of at least two different types of aircraft (e.g. a basic trainer like the Tucano and an advanced trainer like the Hawk). There will be about 90 hours of training in the selection and elementary phases. To reduce flight costs most of this will be done on modified light aircraft with a single piston/propeller engine and semi-aerobatic capability (e.g. Bulldog, Firefly). Such aircraft have a limited top speed of about 130 kt. The next phase (basic training) lasts for about 120 hours, using faster turboprop or light turbojet trainers (e.g. Tucano, L39). This includes visual flying experience (climbs, descents, turns, stall and spin) together with some aerobatics navigation training, instrument flying and formation flying. The advanced training phase is about 100 hours' duration and takes the pilot up to the point of transfer to an operational conversion unit (OCU). This phase will involve using an advanced turbojet trainer (e.g. Hawk) to provide experience at higher speeds (530 kt) and higher '*g*' manoeuvres. The programme will include air warfare, manoeuvrability, ground attack, weapon training and flight control integration. The operational conversion unit will use two-seat derivatives of fast jets and provide the experience for lead-in to operational type flying.

To devise a training system for both basic and advanced phases based on a single aircraft type will present commercial opportunities to the manufacturer together with overall cost and operational advantages to the air force. If innovation can be harnessed to produce a system to meet all the through-training requirements it would offer substantial advantages over all existing training aircraft and current projects which offer less capability. This is obviously a difficult task but the key to the successful solution to this problem lies in the careful exploitation of new technologies that have been used in other aeronautical applications.

Designing a new training system that introduces, develops and relies on innovation carries a commercial risk associated with the unpredictability of the technology. Although, as engineers we may have complete faith in new concepts, perhaps the principal drawback in using a novel, high-tech system lies in the conservative nature of our proposed customers (i.e. training organisations). Any new system must possess the ability to gradually evolve new features even if this means a temporary partial degrading of the overall concept in the early stages.

With the above considerations in mind we (the designers) are required to produce a technically advanced system to meet the defined training requirements yet exhibit sufficient capability to avoid initial scepticism from established customers. The system must show technical and economic advantages over existing equipment and possess the possibility to develop alternative combat aircraft variants based on the trainer airframe, engine and systems.

5.2.4 Information Retrieval

Researching trade journals (e.g. the annual military aircraft reviews in aviation magazines, like *Flight International* and *Aviation Week*) provides data on existing and recently proposed training aircraft. Clearly the market is saturated with training aircraft of various types. The list below shows aircraft that are available to potential customers.

Aeromacchi (MB339/S211A/S260)	Italy
Aeromacchi/Alenia/Embraer/Aerospatial (AMXT)	International
Aero Vodachody (L39/L59/L139/L159B)	Czechoslovakia
AIDC (AT-TC-3A/B)	Taiwan
Avionne (JARA/G-4M)	Romania
BAE Hawk	UK
Boeing (T2/T28/T43)	USA
Boeing/McD (T45 Goshawk)	USA
Bombardier/Shorts (Tucano)	UK
CASA (C101DD)	Spain
Cessna (T37)	USA
Daewoo (KTX-1)	South Korea
Dassault-Breguet/Dornier, Alpha Jet	International
Denel (MB326M)	South Africa
Fuji (T3/T5)	Japan
HAL (Kiran 1A/MK2)	India
Israel Aircraft (TC2/TC7)	Israel
Kawasaki (T4)	Japan
Lockheed Aircraft (IA63)	Argentina
MAPO (MIG-AT)	Russia
Mitsubishi (T2)	Japan
NAMC (K8)	China
Northrop-Grumman (T38)	USA
PAC	Pakistan
Polskie (WSK PZL M-93V, I-22)	Poland
Raytheon-Beechcraft (T1/JPATS)	USA
Rhein-Flugg. (Fantrainer)	Germany
SAAB (SK 60W)	Sweden
Samsung/Lockheed (KTX-2)	South Korea
Socata(TB30/TB31)	France
UTVA (Soko G4)	Serbia
Yakovlev(Yak 130)	Russia
(Yak/Aeromacci) (Y130)	International

The list above is a 'mixed-bag' of aircraft including propeller types, derivatives of existing non-training aircraft, and some purely national projects. It is necessary to review the collection to select aircraft that we feel are more appropriate to this project. The following aircraft are regarded as significant:

1. B.Ae. Hawk (Mk60/100): this is one of the most successful training aircraft in the world with more than 700 produced and sold internationally.
2. L139/159: are 'westernised' versions of the very successful earlier Czech training aircraft (L39/59) which were used by air forces throughout the old Eastern Bloc. When fully developed it may present a serious competitor in future trainer markets.
3. MB339: is a derivative of the very successful Italian trainer (MB326). It has been extensively modernised with upgraded avionics and a modern cockpit.
4. MiG-AT: compared with the above aircraft this is a completely new design by the highly competent Russian manufacturer. It is in competition with other aircraft for the expected 1000+ order for the Russian air force and their allies. It presents a serious competitor to this project.
5. Yak/AEM 130: this is a new subsonic trainer from a Russian/Italian consortium. It will compete with the MiG-AT for the Russian air force order and could be a considerable challenge to the Hawk in future years.
6. KTX-2: is a new supersonic (M1.4) trainer from a South Korean manufacturer (in association with Lockheed Martin). It is expected to be sold in direct competition with all new trainer developments and with other light combat aircraft.
7. AMX-T: this is a trainer development of the original AMX attack aircraft. It is produced by an international consortium and will be a strong contender in future advanced trainer aircraft markets.

5.2.4.1 Technical Analysis

Details of the aircraft in the list above have been used in the graphs described below to identify a suitable starting point for the design. Decisions on selected values to be used in the project are influenced by this data. To reduce format confusion the graphs are plotted in SI units only.

Empty mass data (conversion: 1 kg = 2.205 lb)

Figure 5.2-3 shows the empty mass plotted against maximum take-off mass for jet trainers. The graph also shows the constant 'empty mass ratio' radials. These radials can be seen to bracket 0.75 to 0.45. Our **selected value of 0.6** lies between the higher values for the Russian aircraft and the Italian MB338 but above those for the L159, Hawk and Alpha Jet.

Wing loading (conversion: 1 kg/sq. m = 0.205 lb/sq. ft)

Figure 5.2-4 is a graph of the maximum take-off mass versus wing reference area for existing aircraft. The wing loading radials bracket 500 to 200 kg/m^2. Our **selected value is 350 kg/m^2**. Most of the specimen aircraft have higher wing loading but our specified low approach speed requirement will dictate a lower wing loading.

Aspect ratio

Figure 5.2-5 plots the wing aspect ratios for the trainer aircraft. Most seem to lie in the region of 5 to 6. A value of **5 will be used** as an initial guide to the wing planform geometry. In subsequent phases of the design process, it will be necessary to conduct detailed 'trade-off studies' to establish the technical 'best' choice of wing aspect ratio. At this stage in the development of the aircraft it is impossible to do such studies as sufficient details of the aircraft are unknown.

Thrust loading (conversion: 1N = 0.225 lb)

Figure 5.2-6 shows the installed thrust versus maximum aircraft take-off mass. The radials show that modern aircraft lie along the 40 per cent SLS thrust line. As might be expected the manoeuvre and performance of these aircraft are similar. The new supersonic aircraft are above this line and older aircraft substantially below. This reflects the requirement for improved performance for newer aircraft. We **will select a 40 per cent based on the SSL thrust rating**.

5.2.4.2 Aircraft Configurations

Looking in detail at the configuration of aircraft in the candidate list confirms the impression that most of the existing trainers are conventional in layout. They all have twin, tandem cockpits with ejector seats and large bubble canopies. Apart from the latest Russian designs they are single-engined with fuselage side intakes. The slower aircraft have thick (12 per cent) relatively straight wings. Some of the later designs have thinner swept wings to match the faster (supersonic) top speeds. The wing/fuselage position is mostly low set but with some at mid-fuselage. The Alpha Jet has a shoulder wing position. Tail position for all aircraft except the MiG is conventional with the tailplane set on the aft fuselage with the fin slightly ahead to give protection for post-stall control. The MiG originally had a 'T' tail but this was later changed to a mid-fin location.

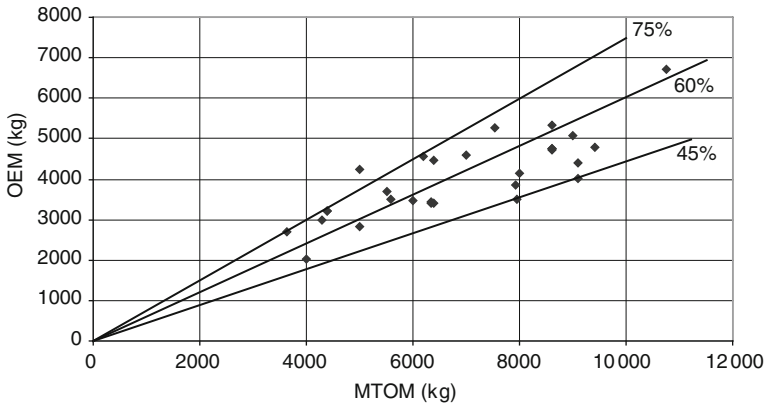


Figure 5.2-3 Survey of empty mass ratio.

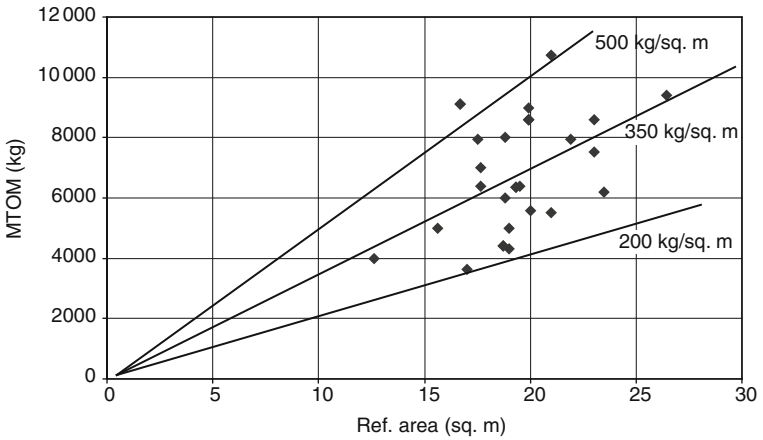


Figure 5.2-4 Survey of wing loading (kg/m^2).

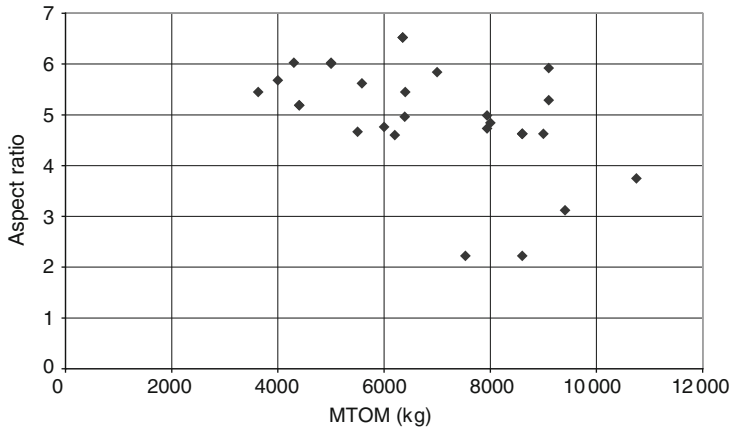


Figure 5.2-5 Survey of wing aspect ratio.

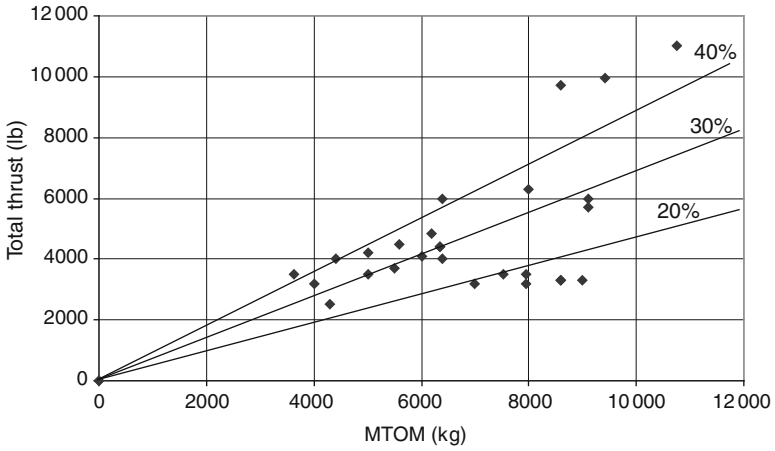


Figure 5.2-6 Survey of thrust/weight ratio (lb/kg).

5.2.4.3 Engine Data

Engines suitable for trainer aircraft lie in the 9 to 29 kN (2000 to 6500 lb) thrust range. The engines shown in Table 5.2-1 are available.

Figure 5.2-7 shows the engine weight (mass) versus SSL thrust data.

Table 5.2-1

Engine (Manufacturer)	Used on	Thrust SLS (kN/lb)	SFC (@ SLS) (–/hr)	Eng. mass (kg/lb)
FJ44-2A (Williams/RR)	–	10.20/2300	0.45	203/447
JT15D (P&W Can)	Citation	13.54/3045	0.55	284/627
Larzac 04-C20 (TM/Snec.)	MiG-AT	14.12/3175	0.74	302/666
J85-21 (Gen. Elec.)	F5/T38	15.57/3500	1.00	310/684
Viper 680 (Rolls Royce)	MB339	19.30/4339	0.98	–
PW 545A (P&W Canada)	Citation	19.79/4450	0.44	347/765
DV-25 (PS/Russian)	Yak 130	21.58/4852	0.60	450/992
TFE 731-60 (Allied-Sig.)	Citation	24.86/5590	0.42	421/929
CFE 738 (GE/ASE)	Falcon 2000	24.90/5600	0.37	601/1325
PW 306A (P&W Canada)	DO 328	25.35/5700	0.39	473/1043
Adour 871 (Rolls Royce)	Hawk/T45	26.81/6028	0.78	602/1328
F124-100 (Allied-Signal)	–	28.02/6300	0.81	499/1100
AE 3007C (AEC)	UAV	28.89/6495	0.33	717/1581

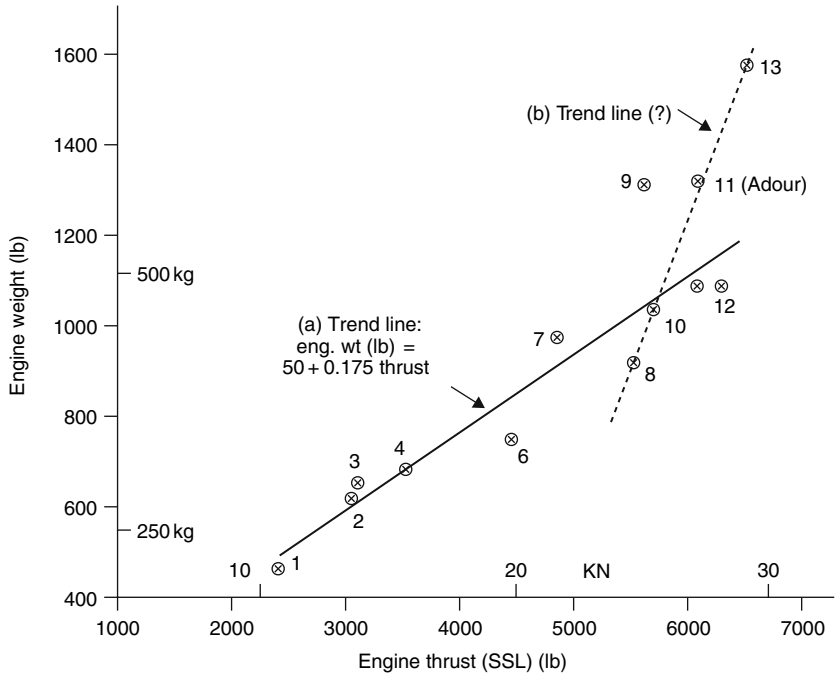


Figure 5.2-7 Survey of engine weight versus SSL thrust.

5.2.5 Design Concepts

To provide a stimulus for the design of the aircraft it has been decided that a radical (novel) solution to the problem should be investigated. This consists of specifying a total training system to cover all the required phases. It comprises an advanced simulator, a single-seat aircraft (see later comment), ground-based instructor console(s) and a modern communication and data linking facility (Figure 5.2-8). Removing the instructor from the training aircraft is regarded as feasible with the adoption of new technologies that have been proven in other applications. Experience from flight test data links and recording gives assurance that technically the systems are available and feasible on which to develop a remote instructor system. Without a second seat the aircraft will be simpler, lighter and cheaper and flying solo the pilot will be in a more realistic operational environment. A further advantage lies in the development of the aircraft into a combat derivative. It will obviously be essential to carefully design the communications link and the instructor module to ensure reliable and safe operation. Modern electronic and video equipment should be capable of providing the necessary confidence. This decision was later reviewed.

Development of a two-seat simplified version of the aircraft will be possible by sacrificing some of the payload and performance capability. This may provide a means of avoiding some of the apprehension centred on the use of the

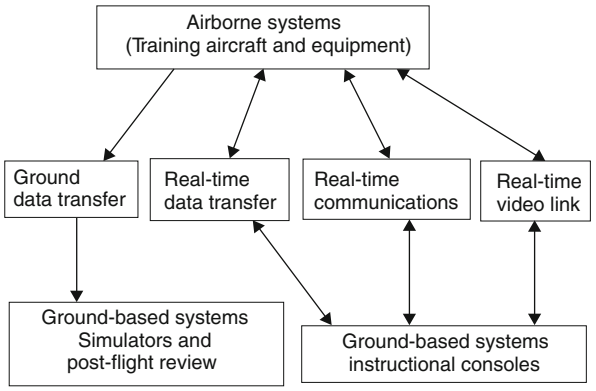


Figure 5.2-8 Proposed total training system diagram.

system in the basic and early parts of the intermediate training phases. The two-seat version represents a relatively straightforward development of the aircraft.

The single-seat aircraft strategy makes it possible to set the design point for the aircraft at the upper end of the advanced training spectrum. This will guide the definition of the critical performance, payload and systems specification. As previously mentioned setting this specification will also provide a better baseline for the development of the combat aircraft derivative.

5.2.6 Initial Sizing

Using the assumed values for empty mass ratio (0.6), wing loading (350 kg/m^2 (72 lb/sq. ft)), aspect ratio (5), thrust loading (0.40) and the specified useful load (pilot + operational equipment + 3000 lb weapon load, totalling an assumed 3308 lb (1500 kg)), it is possible to make estimates of the initial mass and sizes for the aircraft.

$$M_{TO} = \frac{M_{UL}}{1 - (M_E/M_{TO}) - (M_F/M_{TO})^*}$$

*Using a value of 0.15 for the fuel fraction (from Hawk data, this will need to be verified later) and substituting the known and assumed values gives:

$$\frac{1500}{1 - 0.6 - 0.15} = 6000 \text{ kg } (13\,230 \text{ lb})$$

With this aircraft mass the assumed wing loading gives:

Wing reference area (S) = (6000/350) = 17.14m^2 (184 sq. ft)
Using an aspect ratio of 5 sets of wing span (b) = $(5 \times 17.14)^{-0.5}$ = 9.26 m (30.4 ft)
This sets the mean chord (c_{mean}) = (9.26/5) = 1.85 m (5.9 ft)

Assuming a wing taper ratio of 0.25 sets the approximate values for the centre line chord of 3.0m (9.8 ft) and tip chord of 0.75 m (2.5 ft)

In drawing the aircraft we will round off the measurements to give a span of 9.0 m (29.5 ft). This results in the slightly larger wing area of 16.88 m^2 (181sq. ft)

The selected thrust loading of 0.4 in association with the estimated aircraft mass gives a required static sea level thrust of $(0.4 \times 6000 \times 9.81)$ = 23.54 kN (5300 1b).

The choice of engines to provide this thrust rests between:

- the old and slightly overpowered Adour engine used in the Hawk,
- a more modern higher bypass engine from Allied Signal/P&W Canada/GE (used on business jets),
- a slightly underpowered Russian engine as used on the Yak 130, or two TM/Snecma engines as specified for the MiG-AT.

This presents a somewhat difficult choice as:

- the Hawk engine is thirsty,
- the higher bypass engines are larger diameter and lose thrust at altitude (a major disadvantage for the proposed combat variant),
- the Russian manufactured engines may not appeal to established Western customers,
- installing two engines will complicate the systems and cockpit (but would be representative of modern fighter configurations).

After careful consideration it has been decided to use the Adour engine as it is well respected by established customers, is reliable and will add confidence to our novel training system. The Adour 861 provides 5700 lb of thrust so it would be possible to derate the engine for the main specification (this would extend engine life). This strategy would make extra thrust available for the faster trainer and combat variants. With maximum thrust available, the thrust loading would be increased to 45 per cent.

The initial sizing above has provided sufficient data to consider in more detail the initial aircraft layout.

5.2.6.1 Initial Baseline Layout

With our understanding of the configurational options used on existing trainers together with representative sizes of components from our initial sizing, it is possible to consider the detailed layout of the aircraft (Figure 5.2-9).

The following decisions on the aircraft geometry have been taken:

- The aircraft will be of conventional layout with rear fuselage-mounted tail surfaces.
- A single RR Adour 861 engine will be mounted in the aft fuselage, length 2.0 m (80 in), width 0.76 m (30 in), height 1.04 m (42 in), intake diameter 0.7 m (28 in), nozzle diameter 0.6 m (24 in).
- Single seat cockpit with ejector seat and enclosed canopy/windscreen providing the required vision capability.
- The fuselage aft of the cockpit will be suitably configured to permit an easy modification to accommodate a twin-tandem layout for the basic trainer variant.
- The wing will be mid-to-high fuselage (shoulder) mounted to provide generous ground clearance for underwing stores. It will also be manufactured as a single piece (tip to tip) structure which will be mounted above the upper fuselage longerons. This will avoid complicated wing-to-fuselage structural joints.

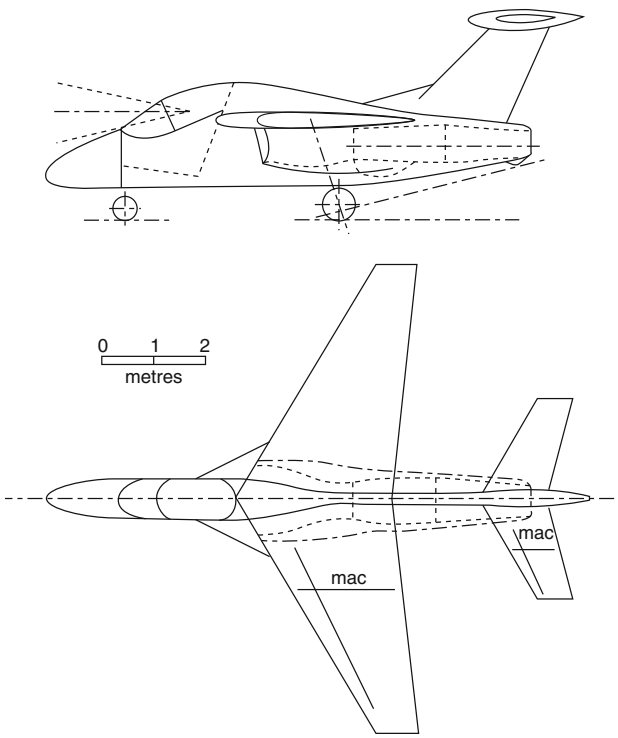


Figure 5.2-9 Initial aircraft layout.

- The wing will be trapezoidal in planform with at least 30° leading edge sweepback and a thin (10 per cent) section (to provide for future higher-speed variants).
- Tail area ratios will match existing aircraft data (from a review of the aircraft data file these values seem appropriate; $S_H/S = 0.255$ giving $S_H = 4.3$ m^2 (46 sq. ft) and $S_V/S = 0.185$ giving $S_V = 3.1$ m^2 (33 sq. ft)).

(Note: an initial layout drawing of the aircraft showed that the short fuselage length makes a conventional fuselage-mounted tailplane suffer from a shortage of tail arm. Therefore a 'T-tail' arrangement has been adopted. This configuration improves both the horizontal and vertical tail effectiveness which allows a reduction in tailplane area to 22 per cent S (=3.7 m^2 (40 sq. ft)). A reduction in fin areas could also be anticipated but this was not adopted, as the proposed two-seat variant will require more fin to balance the increased fuselage nose length. The T-tail arrangement will enable the provision of a communication/video pad to be installed at the top of the fin (at the tailplane junction). This will be useful to accommodate some of the extra equipment necessary for the remote instructor facility.)

- To assist in re-energising the airflow over the rear fuselage and fin in high-alpha manoeuvres, small wing leading edge extensions will be added to the planform.
- Underwing, fuselage-side intakes will be positioned below the leading edge extensions to ensure clean airflow into the engines in high-alpha manoeuvres.
- Conventional tricycle landing gear with wheel sizes representative of existing aircraft (from the existing aircraft data file, the main and nose-wheel diameters are 0.6/0.45 m (24/18 in) respectively).

5.2.7 Initial Estimates

With a scale drawing of the baseline aircraft configuration and an understanding of the engines and systems to be used it is now possible to conduct a series of detailed calculations to estimate the aircraft mass, aerodynamic characteristics and performance.

5.2.7.1 Mass Estimates

The mass of each component of the aircraft can be calculated using methods described in aircraft design textbooks. These are generally based on geometrical and aircraft load data and are often derived from analysis of existing aircraft configurations. Suitable adjustments need to be made in those areas where the proposed design is significantly different from past designs. In our case there are two such considerations:

- much more composite material will be used than in the predominately aluminium alloy aircraft built previously and,
- for this training system more sophisticated and extensive flight control and communication systems will be installed (allowance will need to be made for the reduced mass and volume of new electronic/computer systems).

A design take-off mass of 6000 kg will be assumed for determination of the aircraft structural components. Although this mass is likely to be higher than that estimated for the maximum take-off mass of the aircraft it will provide an insurance against future mass increase. If it is necessary to determine the minimum take-off mass for the aircraft, the estimation would need to be done iteratively using the calculated take-off mass as the design mass input for components mass estimations.

Detailed calculations for mass estimations have not been shown below but the input data on which the calculation was based is given (for reference). To simplify presentation, the data below is shown in SI units only. The completed mass statement is shown in dual units.

Wing: Area (S) = 16.88 m^2 Aspect ratio = 5 Wing thickness (average) = 10%
Sweepback ($c/4$) = 25° Taper ratio = 0.25 Control surface areas = 15% (S)

Conventional mass estimation = 462 kg, assuming 15 per cent reduction for composites = **392 kg**

Fuselage: Length = 9.5 m Depth = 2.0 m Width = 0.75 m

Conventional mass estimation = 576 kg, assuming 10 per cent reduction for composites = **518 kg**

Horizontal tail: Area = 22% (S) Tail span = 4.0 m Fuselage width = 0.75 m

Conventional mass estimation = 76 kg, assume 15 per cent reduction for composites = **65 kg**

Fin: Area = 18% (S), T-tail structure

Conventional mass estimation = 39 kg, assume 15 per cent reduction for composites = **34 kg**

Undercarriage: $M_{LAND} = 90\% M_{TO}$ $n_{ULT} = 3 \times 1.5 = 4.5$
Main: Length = 0.75 m → Mass = 185 kg
Nose: Length = 0.60 m → Mass = 52 kg Total = **237 kg**

No reduction on this mass for new materials

Engine: Dry = **577 kg** (from manufacturers)
Engine and fuel systems: Assume 50 per cent engine dry mass = **288 kg**
Aircraft equipment: To account for new aircraft system requirements assume 20 per cent of aircraft design mass → **1200 kg**
Fuel: This will be checked by the performance estimates, for now we will still assume 15 per cent of aircraft design mass → **900 kg**
Crew: Assume one pilot with operational equipment → 300 lb = **136 kg**
Weapon load: Specified at 3000 lb → **1360 kg**

Hence the initial mass statement for the aircraft can be compiled:

Component		**kg/lb**	**%M_{TO}**
Wing		392/664	6.9
Fuselage		518/1142	9.1
Horizontal tail		65/143	1.7
Fin		34/75	
Undercarriage		237/523	4.2
	Total structure	**1246/2747**	**21.8**
Engine (dry)		577/1272	
Engine and fuel systems		288/635	
	Total propulsion	**865/1907**	**15.2**
Equipment (total)		1200/2646	21.0
	Total aircraft empty	**3311/7300**	**58.0**
Fuel		900/1984	15.8
Crew		136/300	
Weapon load		1360/3000	
	Total Useful Load	**2396/5283**	42.0
	Maximum Take-off Mass*	**5707/12584**	100.0
	Note: aircraft design take-off mass	6000/13230	

(*This compares with 9000 to 7500 kg (20 000 to 16 000 lb) for older trainers and is competitive with the newer designs.)

At this stage all the input values used to predict component masses are somewhat tentative, therefore the use of more extensive methods are inappropriate.

The aircraft aerodynamic and performance calculations that follow will use the above mass statement to determine aircraft masses for various stages in the flight profiles:

- At take-off, the aircraft can be loaded up to a variety of different conditions depending on the operation to be performed. The most critical take-off condition will be at maximum take-off mass. This will be used in the estimation of take-off distance.
- At landing, the aircraft could again be at different mass. The most critical flight case would be to land immediately after taking off with full fuel and payload. This would be an emergency condition and therefore it is not necessary to use this as a design case (in an emergency the pilot may jettison some of the weapons and/or part of the fuel load to reduce the landing mass). The minimum landing mass would relate to zero weapon load and only a small percentage of fuel remaining (e.g. 10 per cent). This may also be a rare condition and therefore

not applicable to design calculations. The landing calculations are often performed at a landing weight of 90 per cent of the take-off weight. For the maximum take-off mass condition this relates to 5707 × 0.9 = 5136 kg (11 325 lb). Another take-off condition that may be considered is a take-off with zero weapon load (i.e. a basic training mission). Using 90 per cent of the take-off weight in this condition gives (5707 − 1360) 0.9 = 3912 kg (8626 lb). These two cases will be used in the landing distance and speed estimations.

- For the manoeuvring calculations (turns and climb) a mean flying weight will be assumed. This relates to half weapon and half fuel load (= 4577 kg/10 092 lb).
- For the ferry case the aircraft take-off condition will be without weapon load but may have carry external wing tanks.

5.2.7.2 Aerodynamic Estimates

Using the geometrical data from the initial baseline layout and masses from the above section, it is possible to make initial estimates for the aircraft lift and drag in various flight conditions.

Aircraft drag coefficients

Using standard equations from aeronautical textbooks and the data below, it is possible to estimate the drag coefficients for the aircraft in different flight conditions.
In SI units:

Take-off mass (M_{TO})	5707 kg	Wing aspect ratio (A)	5
Wing ref. area (S)	16.88 m^2	Wing LE sweep (Λ)	30°
Wing loading	3212 kN/m^2	(337 kg/m^2)	

The following equations and parameters are appropriate:

- Induced drag factor $K = (1/\pi eA)$
- Planform factor $e = 4.61((1 - 0.045A^{0.68})\ (\cos\ \Lambda)^{0.15}) - 3.1$
- Assumed skin friction coefficient $C_f = 0.0038$
- Profile drag coefficient $C_{DO} = C_f(S_w/S)$ (assumed aircraft wetted area – estimated from the initial layout drawing (total) $S_w = 78$ m^2)
- Dynamic pressure $q = 1/2\rho \cdot V^2$
  (ρ and V are the air density and aircraft speed at the flight condition under investigation)
- Aircraft lift = $C_L qS$
- Aircraft drag = $C_D qS$
- Aircraft lift/drag ratio = $L/D = C_L/C_D$

The above data was input to a spreadsheet program to determine the lift and drag parameters for a range of aircraft speed (V), operating altitude (h) and load factor (n) for the aircraft in the 'clean' (no weapons, u/c and flaps retracted) flight condition. The range of values used in the calculations is shown below:

V (m/s)	50/100/150/200/250/300/350
h (ft)	0 (SL)/25 000 (FL250)/36 000 (FL360)
n (g)	1/3/6

The spreadsheet results for sea level at $n = 1$ are shown in Figure 5.2-10. The thrust displayed in the graph is for the Adour engine (no variation of thrust with aircraft forward speed is typical for low bypass ratio engines).

The drag analysis described above assumes subsonic flow conditions but the upper value for aircraft speed is seen to be in excess of M1.0 at each altitude. Obviously aircraft drag will increase rapidly as supersonic flow is developed over the aircraft. Some allowance will need to be taken for the wave drag at higher speeds. Also, since it is intended to investigate the potential for increasing the aircraft top speed into the transonic range (e.g. M1.2) it is necessary to make suitable corrections.

Textbooks quote the Sears–Haack wave drag coefficient as:

$$C_{Dwave} = (9\pi/2)(A_{max}/L)^2(1/S)$$

where
(A_{max}) is the maximum cross-sectional area of the aircraft
(L) is the distance from the nose to the position where the area is maximum
(S) is the reference wing area for the aircraft

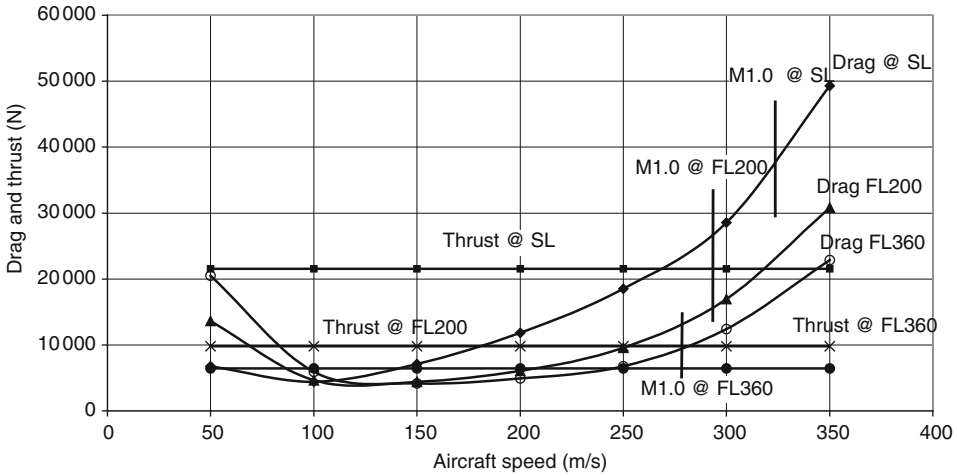


Figure 5.2-10 Aircraft drag polar.

From the aircraft drawing $A_{max} = 2.0$ m^2 (21.5 sq. ft) and l = 4.5 m (14.75 ft) are reasonable estimates. This data gives $C_{DW} = 0.14$. This value could be reduced to account for the swept wing but would need to be increased due to the poor area ruling of the design (due mainly to the effect of the canopy). At this early stage in the analysis of the aircraft it is assumed that these effects cancel leaving $C_{DO} = 0.14$ at M1.05.

Note: the increase due to wave drag at M1.05 approximately doubles the subsonic drag previously calculated. The wave drag will be progressively felt from the drag divergent Mach number, which for this wing sweep will be about M0.75. At this stage in the aircraft evaluation it will be sufficient to sketch a blend line from the subsonic drag polar to account for the values above. A suggestion from one textbook assumes half the wave drag increase to have been achieved at M1.0, therefore the blend will be from M0.75 and through the M1.0 and M1.05 points.

Aircraft lift coefficients

From an analysis of stall and approach speeds from existing aircraft it is acceptable to assume maximum lift coefficients for our aircraft of:

Landing $C_{lmax} = 2.10$

Take-off $C_{lmax} = 1.70$

5.2.7.3 Performance Estimates

The calculation of the aerodynamic coefficients described above are relatively crude and do not take into account many of the detailed factors that are known to be significant. However, they do provide 'ballpark' values which can be used to initially assess aircraft performance. From these calculations it will be possible to identify the changes that are necessary to the design to meet the specified design criteria. Later analysis can account for the more detailed aspects of the predictions.

Maximum speed

Applying the wave drag increase to the subsonic drag polars at the three altitudes (SL, FL250, FL360), shows (Figure 5.2-10) that a top speed of 320 m/s (621 kt) is possible at SL-ISA with the full rating of the Adour engine. If the engine was derated to 85 per cent of maximum thrust the speed would reduce to 305 m/s (592 kt). Both of these are well in excess of the 450 kt specified in the design brief. However, these calculations have been done with the aircraft in the 'clean' condition. If an allowance for weapon drag ($\Delta C_D = 0.01$) is made, the speeds are reduced to 280 m/s (544 kt) and 260 m/s (505 kt) respectively. These are still faster than the specified speeds.

The conclusion to this part of the performance study is that the aircraft will easily meet the specified top speed requirement.

Turn performance

Aircraft performance textbooks quote the general equations as:

- turn rate $= \chi = g(n^2 - 1)^{0.5}/V$
  where n = normal acceleration factor
  V = aircraft speed
- radius of turn $R = V/\chi$
- aircraft bank angle $\phi = \cos^{-1}(1/n)$

These equations have been solved, using a spreadsheet, for values of:

$V = (50/100/150/200/250/300/350$ m/s) and
$n = (1/2/3/4/5/6/7/8)$

(A reminder: 1 m/s is approx. 2 kt (more precisely 1.94))

Note that the turn equation above is independent of aircraft parameters. The resulting curves are shown in Figure 5.2-11. This graph describes the overall manoeuvring design space for the aircraft.

For our aircraft the boundaries to the manoeuvring space are:

- aircraft maximum speed (assumed to be M0.8),
- stall speed (clean at mean flying weight with $C_{L\max} = 1.10$),
- structural load limit of $n = 7$.

These provide the limiting lines shown on the sea-level manoeuvring diagram (Figure 5.2-12). The corner speed gives the maximum instantaneous turn rate. A value of 24°/s is predicted for our aircraft.

Determination of the sustained turn requires an analysis of the available specific excess power (SEP) for the aircraft. The zero SEP boundary provides the value for sustained turn rate. From textbooks the equation for SEP is:

$$\text{SEP} = V(T - D)/W$$

SEP will therefore vary according to the aircraft speed, engine thrust, aircraft drag and weight (i.e. M·n·g). Drag (and thrust) will vary with aircraft altitude and speed. The values determined for SEP for a range of aircraft speeds (50 to 350 m/s) and load factors (1 to 6), for both sea level and FL250, have been calculated and are plotted in Figures 5.2-13 and 5.2-14. Cross plotting the aircraft speeds at zero SEP from these graphs onto the manoeuvring diagram and joining with a smooth line provides the boundary for sustained turn rate.

- From Figure 5.2-12: maximum instantaneous turn rate = 24°/s at a flight speed of 165 m/s (the specified requirement is 18°/s).
- From Figure 5.2-13 with the results transferred to Figure 5.2-12 (point A): sustained turn rate = 17°/s at sea level at a flight speed of 140 m/s (the specified requirement is 14°/s).
- From Figure 5.2-12 with the zero SEP lines transferred from Figures 5.2-13 and 5.2-14: sustained turn g of 6 at sea level (point B) and just about 3 at (point C) and FL250 (the specified requirements are 4g and 2g).

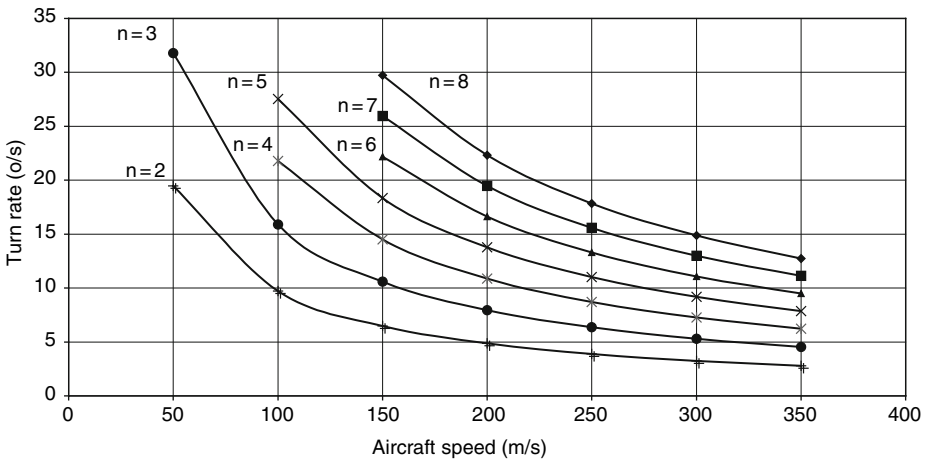


Figure 5.2-11 Manoeuvring design space.

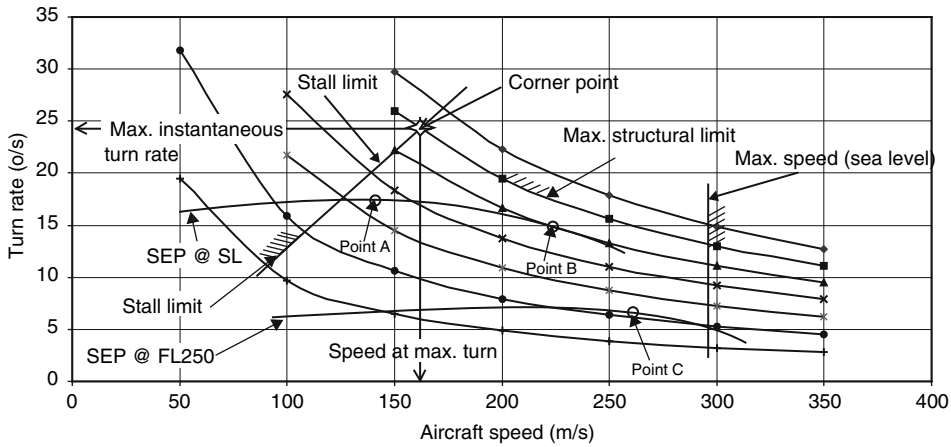


Figure 5.2-12 Aircraft turn performance.

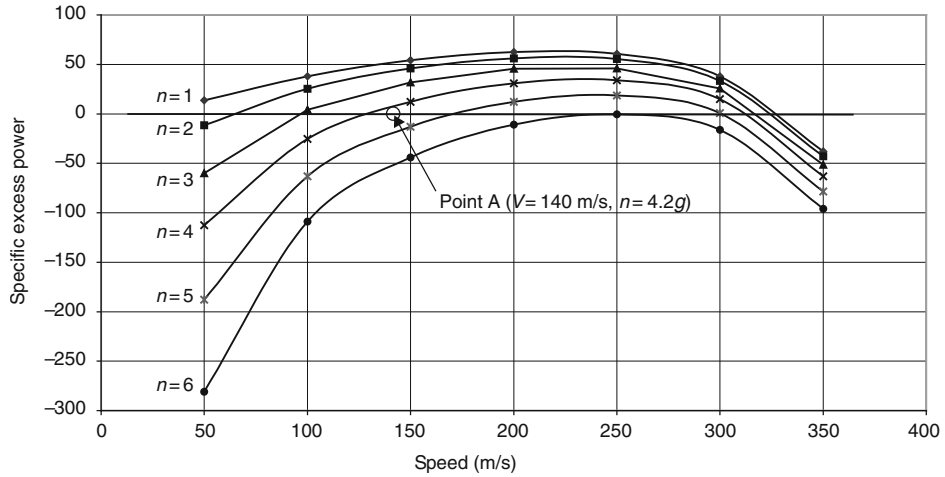


Figure 5.2-13 Specific excess power at SL.

The above calculations were done at aircraft maximum take-off mass with a clean configuration. Similar calculations were done with weapon drag added ($\Delta C_D = 0.01$) at mean flying mass (4577 kg/10 092 lb) but these were shown to be less critical than the clean-heavy case quoted above.

Note: the manoeuvre graph shows that an extra 5°/s instantaneous turn rate would be achieved if the aircraft structural limit was raised from the specified 7*g* to 8*g*.

Field performance

For take-off and landing calculations it is necessary to add the extra drag due to the extended undercarriage and flaps. The flap deflection for take-off is assumed to be 20° and for landing 40°. The following drag contributions are regarded as appropriate:

ΔC_D undercarriage = 0.0075

ΔC_D 20° flap = 0.015

ΔC_D 40° flap = 0.030

Lift coefficients are as quoted earlier (i.e. take-off 1.7 and landing 2.1).

Take-off estimation

In the early part of the conceptual design process simplified take-off calculations can be done using Nicolai's (reference 4) simplified take-off parameter (this includes *W/S*, *T/W*, $C_{l_{max}}$). As we require only the ground run to be estimated, the take-off distance attributed to the climb segment can be ignored. The calculation for total take-off

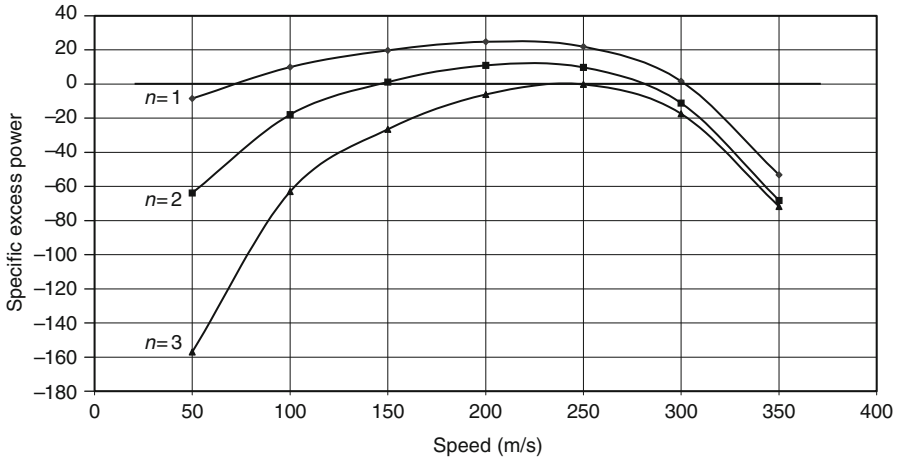


Figure 5.2-14 Specific excess power at FL250.

distance, assuming the aircraft at maximum take-off mass (5707 kg/12 584 lb) and with maximum sea level static thrust (5700 lb), gives:

$$\text{Total distance to 50 ft} = 2408 \text{ ft } (734 \text{ m})$$

Removing the distance covered in the climb gives the take-off ground run = 1856 ft (566 m) (the specified maximum take-off ground run is 2000 ft (610 m)).

Using a derated (85 per cent) engine thrust increases the ground run to 2184 ft (666 m) which is more than the required distance. In this case, to get off from a 2000 ft/610 m runway, the maximum aircraft mass would need to be limited to 5300 kg (11 685 lb).

Landing estimations

For landing, we will assume a landing mass of 90 per cent max. take-off mass. Therefore, $M_{\text{LAND}} = 5135$ kg/11 323 lb which gives an aircraft stall speed of 46.7 m/s (90 kts). The approach speed for military aircraft is set at 1.2 V_{stall} giving 108 kt which means that the specified value of 100 kt is not achieved. In the design brief the appropriate aircraft landing weight for the approach speed requirement is not specified. If we assume the low approach speed is applicable to the basic training role we can assume the weapon load to be zero. Assuming the landing weight is to be set at 90 per cent of the take-off weight gives a landing mass for the basic training role of:

$$M_{\text{LAND}} = 0.9(5707 - 1360) = 3912 \text{ kg } (8626 \text{ lb})$$

At this mass the stall speed is 40.8 m/s (79 kt) making the approach speed = $1.2 \times 79 = 95$ kt. This achieves the requirement but begs the question of the relationship of the approach speed requirement to the aircraft role.

Assuming a constant deceleration of 7 ft/s^2 (as assumed in the simplified Nicolai[4] estimation) and a touch-down speed of 1.15 V_{stall}. The landing distance is determined from:

$$\text{Landing ground run} = V_{\text{TD}}^2/(2 \times 7)$$

$$(V_{\text{TD}} = 1.15 \times 46.7 = 53.7 \text{ m/s } (176 \text{ ft/s}))$$

$$\therefore \text{Landing ground run} = 2212 \text{ ft } (675 \text{ m})$$

This is also in excess of the specified distance of 2000 ft but the calculation was done at maximum landing mass (5136 kg/11 325 lb). Performing the same calculation at the lighter landing mass assumed above (3912 kg/8626 lb), using the same aircraft assumptions gives:

$$\text{Landing ground distance} = 1690 \text{ ft } (515 \text{m})$$

This easily achieves the specified distance of 2000 ft (610 m) but again begs the question of the definition of landing weight.

The heavy and light landing mass assumptions used in the landing calculations provide a range of maximum/minimum values for the aircraft. Further discussion with the project customers would be necessary.

Mission analysis

The mission analysis allows us to estimate the fuel requirements. The project brief specifies three different mission profiles. At this stage, it is not obvious which one of these will be most critical, therefore each will be analysed to determine the required fuel. The calculations use the weight fractions suggested in Raymer's book[1] for the less significant segments of the missions.

Using the aerodynamic analysis described earlier it is possible to determine the lift/drag ratio variation with aircraft speed for the aircraft cruising at 25 000 ft. From this data a representative value of 9.0 will be assumed in the calculations. This decision will be verified later.

The specific fuel consumption of the Adour engine is 0.78 –/hr at the SSL thrust rating. This value will increase with aircraft altitude, speed and engine setting (e.g. cruise). Determining the exact extent of this rise requires engine performance data. At this stage in the design process such data is not available, a value of 0.95 has been assumed in the calculations below.

Basic training profile (two sorties)

The training profile is shown diagrammatically in Figure 5.2-15.

Analysis of the change in aircraft mass in each segment is shown in Table 5.2-2

As this is the basic training mission it is done clean (no weapon drag) and at a lower take-off mass (no weapon load). The first of the two sorties will require the following fuel:

$$\therefore M_{\text{TO1}} = 5707 - 1360 = 4347 \text{ kg (to avoid duplication remember, } 1 \text{ kg} = 2.205 \text{ lb)}$$

$$\therefore M_{\text{fuel1}} = 0.123 \times 4347 = 535 \text{ kg}$$

As there is no intermediate refuelling the second sortie will be flown at a lower take-off mass than the first:

$$M_{\text{TO2}} = 4347 - 535 + 3812 \text{ kg}$$

The same mission will be flown, therefore the same fuel fraction will apply:

$$\therefore M_{\text{fuel2}} = 0.123 \times 3812 = 469 \text{ kg}$$

$$\text{Total fuel required } = M_{\text{fuel1}} + M_{\text{fuel2}} = 535 + 469 = 1004 \text{ kg}/2234 \text{ lb}$$

This is slightly more than the fuel load of 900 kg assumed in the mass calculations.

Advanced training profile (single sortie)

As this profile is flown with full weapon load the aircraft take-off mass will be 5707 kg. To account for the weapon drag the lift/drag ratio will be reduced to 7.5.

The profile is shown diagrammatically in Figure 5.2-16 and the segment analysis is shown in Table 5.2-3.

$$\text{Fuel fraction for the total mission } = (1 - 0.835) = 0.165$$

$$\therefore \text{Fuel required } M_{\text{fuel}} = 0.165 \times 5707 = 942 \text{ kg}/2077 \text{ lb}$$

This is also slightly higher than originally assumed but less than that estimated for the basic training profile above.

Ferry

The quoted ferry range is 1000 nm. This will be flown from a maximum take-off mass of 5707 kg and at optimum speed. It is assumed from Figure 5.2-17 that V_{cruise} is 150 m/s (291 kt). At this speed the aircraft lift/drag ratio is 9 for best lift/drag ratio (clean) as shown on Figure 5.2-18. The engine specific fuel consumption (c) at this condition will be assumed to be 1.05. To estimate the fuel fraction we will use the transformed Breguet range equation:

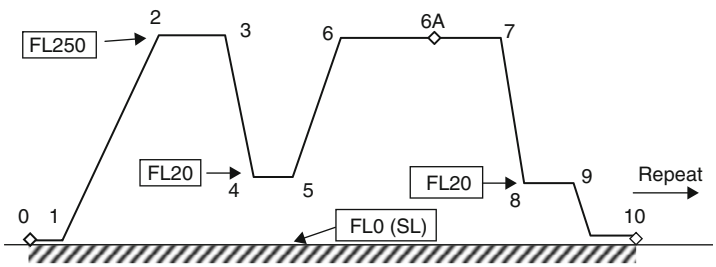


Figure 5.2-15 Mission flight profile (basic training).

Table 5.2-2

Segment	Description	Parameters	$(M_{(n+1)}/M_n)^*$
0 to 1	Take-off		0.970
1 to 2	Climb		0.985
2 to 3	Cruise (out)	6 min	0.990
3 to 4	Descent		0.995
4 to 5	General handling	10 min	0.980
5 to 6	Climb		0.985
6 to 6A	Manoeuvres	4 min	
6A to 7	Descent	5 min	0.995
7 to 8	Recover		0.995
8 to 9	Land		0.995
9 to 10	Cruise (return)		0.980
			$\sum$0.877

Fuel fraction for the above sortie $M_{fuel}/M_{TO} = (1 - 0.877) = 0.123$.
*$(M_{(n+1)}/M_n)$ is the ratio of the aircraft mass at the end of the segment relative to that at the start.

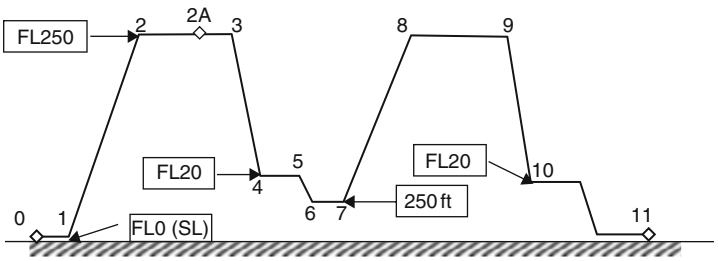


Figure 5.2-16 Mission flight profile (advanced training).

Table 5.2-3

Segment	Description	Parameters	$(M_{(n+1)}/M_n)^*$
0 to 1	Take-off		0.970
1 to 2	Climb		0.985
2 to 3	Cruise (out)	6 min	
	Weapon training	10 min	0.967
3 to 4	Descent		0.995
4 to 5	Aerobatics	10 min	0.970
5 to 6	Descent		
6 to 7	Low level	10 min	0.970
7 to 8	Climb		0.985
8 to 9	Cruise (return)	6 min	0.990
9 to 10	Descent		0.995
10 to 11	Land		0.995
			$\sum$0.835

*$(M_{(n+1)}/M_n)$ is the ratio of the aircraft mass at the end of the segment relative to that at the start.

$$(M_1/M_2) = \exp[-(R.c)/(V.(L/D))]$$
$$= 0.72$$

Multiplying the take-off, climb and landing mass fractions as used in the previous mission analysis (to account for the fuel used at the start and end of the ferry mission) to this gives the overall fuel mass fraction:

$$M_{start}/M_{end} = 0.97 \times 0.72 \times 0.995 = 0.696$$
$$\therefore \text{Fuel required} = 5707(1 - 0.696) = 1733 \text{ kg } (3821 \text{ lb})$$

This is higher than the fuel that can be carried internally (assumed to be 900 kg), therefore external tankage will be necessary. This would reduce the (L/D) ratio used above due to the increased drag from tank profile. A rough estimate of this effect shows that about another 140 kg of fuel would be required.

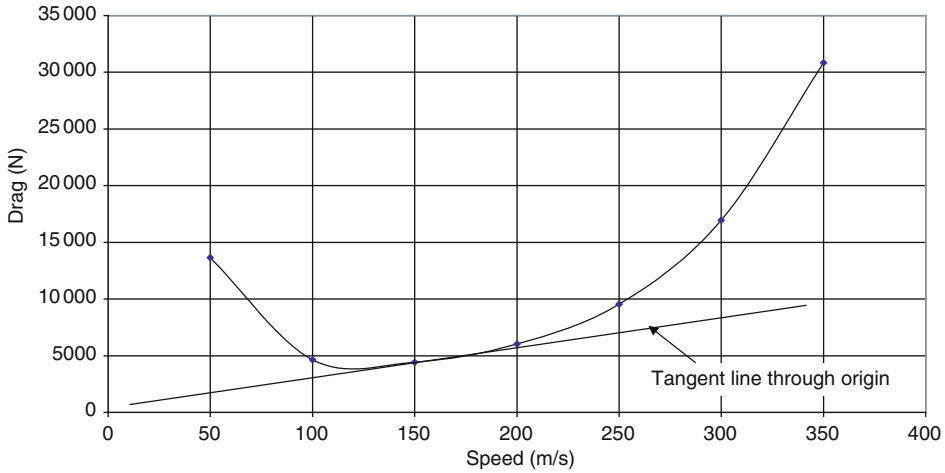


Figure 5.2-17 Ferry mission drag polar.

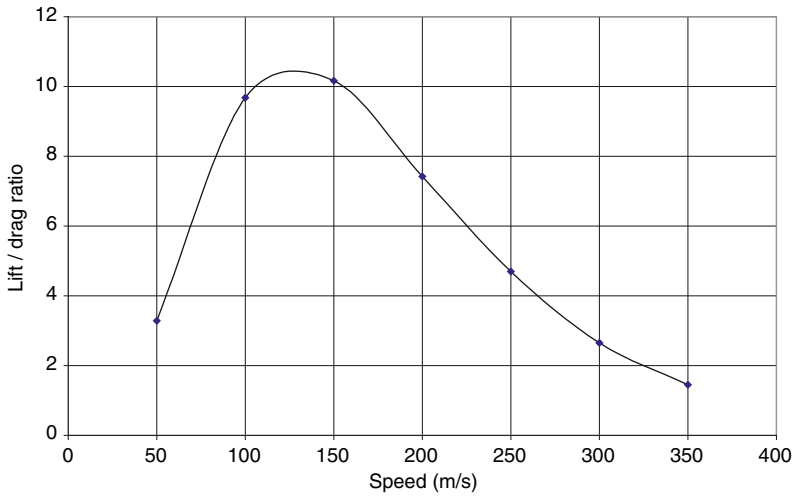


Figure 5.2-18 Aircraft cruise lift/drag ratio.

Reversing the analysis above and converting the entire payload (weapon) to fuel shows that:

$$\text{Maximum fuel mass} = 1360 + 900 = 2260 \text{ kg } (4983 \text{ lb})$$

Subtracting the take-off, climb and landing fuel fractions reduces this to 2181 kg.

Assuming 5 per cent of the external fuel mass is required for the external fuel tank structure (i.e. $1360 \times 0.05 = 68$ kg) means that only $2181 - 68 = 2113$ kg is available for fuel.

From the Breguet range equation:

$$\begin{aligned}\text{Ferry range} &= [V \times (L/D)/c]\log_e(M_0/M_1) \\ &= (291 \times 10/0.95)\log_e(5560/3447) = 1464 \text{ nm}\end{aligned}$$

(Note: this range estimation ignores any requirement for reserve fuel at the end of the flight.)

From the above calculation it appears that the requirement for 2000 nm ferry range is too difficult to meet. Keeping this requirement could seriously compromise the basic design of the aircraft. Within the accuracy of the calculations at this stage, it would be reasonable to request a reduction to this requirement down to a value of 1500 nm.

Climb and ceiling estimations

The rate of climb equation is:

$$(R \text{ of } C) = V(T - D)/W$$

Note: this is equivalent to the SEP expression.

Using the aerodynamic analysis from the earlier work it is possible to plot the rate of climb against aircraft speed for flight at SL, FL250 and FL360. Figure 5.2-19 shows the rate of climb at the mean mass condition with an allowance for an increase in drag for weapons ($\Delta C_D = 0.01$). The locus of maximum climb rate allows the data to be cross-plotted (Figure 5.2-20). This shows the service ceiling (i.e. the height at which the rate of climb falls to 100 fpm (0.5 m/s)) to be about 46 000 ft (against a specified requirement of 40 000 ft).

Figure 5.2-20 shows that the average rate of climb from sea level to FL250 is 34 m/s. This is used to predict that the time to climb to FL250 (7625 m) is 190 seconds (=3.2 min). This easily meets the specified requirement of 7 min. The same calculations were done to check the climb rate and time for maximum take-off mass with full weapon drag ($\Delta C_D = 0.017$). This showed an average rate of climb of 30 m/s, which leads to a time to climb to 25 000 ft of 4.2 min. At a similar condition but with engine derated to 85 per cent thrust gives a time of 5.8 min.

All the calculations above show that the climb and ceiling requirements are not critical.

Summary of initial performance analysis

- The initial estimate of take-off mass of 5707 kg will need to be revised to 5814 kg to account for the increase in fuel required for the basic training flight profile.
- The maximum speed even at 85 per cent thrust is 505 kt. This easily exceeds the specified requirement of 450 kt.
- All the turn performance criteria are easily met.
- Take-off ground run at 1856 ft is below the specified 2000 ft but with a derated engine of 85 per cent thrust, this increases to 2184 ft.
- The approach speed requirement of 100 kts cannot be met except by a lighter aircraft (no weapon load). In this case an approach speed of 95 kt is achieved.

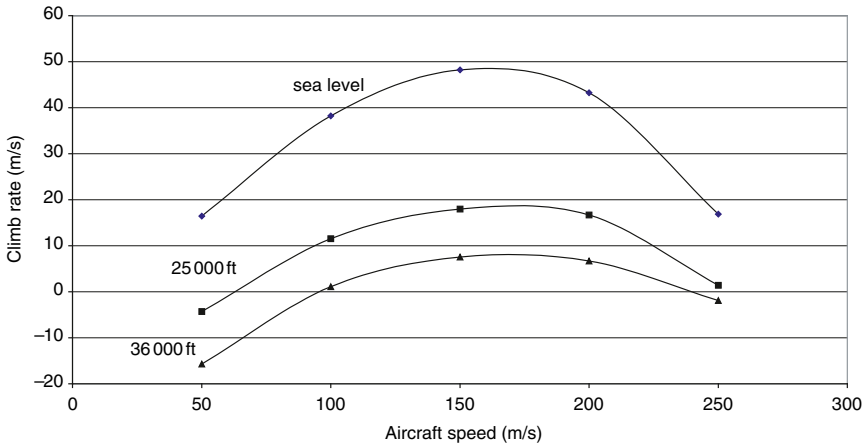


Figure 5.2-19 Aircraft rate of climb.

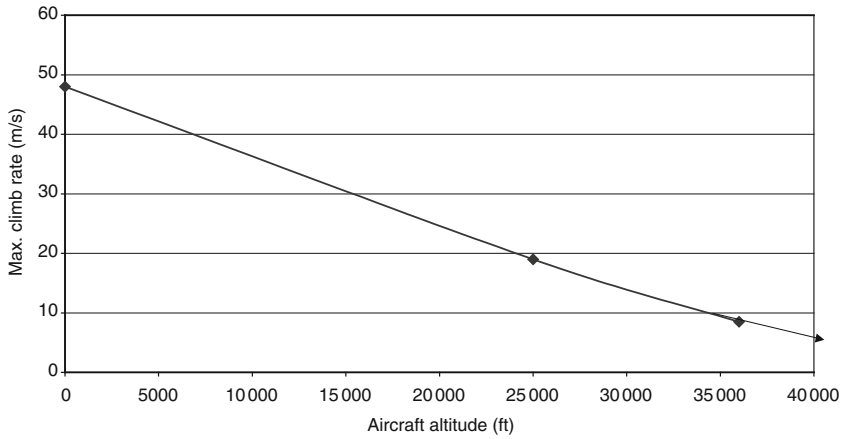


Figure 5.2-20 Aircraft climb and ceiling evaluation.

- Landing ground run at 2215 ft also exceeds the specification of 2000 ft. Only with aircraft at lighter landing mass can the specification be met.
- The ferry mission of 1000 nm cannot be met with internal fuel but can be achieved if 833 kg of fuel is carried externally. The maximum range that could be flown is estimated at 1464 nm. This is substantially less than the 2000 nm specified. It is suggested that this requirement be reviewed as in its present form it would seriously compromise the overall aircraft design.
- Climb and ceiling requirements are easily achieved.

5.2.8 Constraint Analysis

From the project brief there are six separate constraints to be considered in this analysis:

1. Take-off distance less than 2000 ft.
2. Approach speed no greater than 100 kt.
3. Landing distance less than 2000 ft.
4. Combat turn, at least *4g* at sea level.
5. Combat turn, at least *2g* at 25 000 ft.
6. Climb rate to provide for 7 min climb to 25 000 ft.

5.2.8.1 Take-off Distance

The equations to be used to determine the effect of the take-off criterion can be found in most textbooks (e.g. reference 4) as shown below:

$$(T/W) = (\text{constant})\ (W/S)(S_{\text{take-off}} \cdot C_{\text{Ltake-off}})$$

Obviously this represents a straight line on the (T/W) versus (W/S) graph. For our aircraft the lift coefficient in the take-off configuration ($C_{\text{Ltake-off}}$) is assumed to be 1.7. The value $S_{\text{take-off}}$ represents the total take-off distance (i.e. ground roll plus climb distance to 50 ft). Assuming a climb gradient from zero to 50 ft of 5° gives a ground distance covered of 571 ft. Adding this to the specified ground roll of 2000 ft gives $S_{\text{take-off}} = 2571$ ft (784 m).

The constant in the above equation is assessed from Nicholi's book[4] as 1.27 (in SI units with wing loading in kg/m^2), so

$$(T/W) = 1.27/(784 \times 1.7)(W/S) = 0.00095(W/S)$$

5.2.8.2 Approach Speed

Assuming the approach speed $V_A = 1.2\ V_{\text{stall}}$ then:

$$(W/S)_{\text{landing}} = \beta(W/S) = 0.5 \times \rho(V_A/1.2)^2 \times C_{\text{Llanding}}/g$$

V_A is specified at 100 kts (52 m/s). β is the ratio of landing mass to take-off mass. At a maximum landing weight $\beta = 0.9$. At minimum landing weight (i.e. empty aircraft plus pilot plus 10 per cent fuel = 3311 + 136 + 90 = 3537 kg) $\beta = 0.62$.

Assuming the lift coefficient in the landing configuration (C_{Llanding}) = 2.1

$$(W/S) = (0.5 \times 1.225 \times 52 \times 52 \times 2.1)/(1.2 \times 1.2 \times 9.81) = 273.6\ @\ \beta = 0.9$$
$$= 397.1\ @\ \beta = 0.62$$

Note: these constraints are constant (vertical) lines on (T/W) versus (W/S) graphs.

5.2.8.3 Landing Distance

The approximate equation to determine ground run in landing can be rewritten as shown below:

$$(W/S) = (S_{\text{landing run}} \times C_{\text{Llanding}})/(\text{constant} \times \beta)$$

The landing ground run $S_{\text{landing run}}$ is specified as 2000 ft (610 m). The lift coefficient in the landing configuration (C_{Landing}) is assumed to be 2.1 (as above). The expression will be evaluated for the two landing mass fractions used above (i.e. $\beta = 0.9$ and 0.62).

The (constant) in the expression above (in SI units with W/S in kg/m^2) is 5.0.

$$(W/S) = (610 \times 2.1)/(5.0 \times \beta) = 284.7 \text{ @ } \beta = 0.9 \text{ and } 413.2 \text{ @ } \beta = 0.62$$

Note: these are also constant vertical lines on a constraint diagram.

5.2.8.4 Fundamental Flight Analysis

The fundamental equation used in the flight cases can be found in most textbooks. In terms of sea level, take-off thrust loading the equation is:

$$(T/W)_{\text{TO}} = (\beta/\alpha)[(q/\beta)\{C_{\text{DO}}/(W/S)_{\text{TO}} + k_1(n\beta/q)^2(W/S)_{\text{TO}}\} + (1/V)(dh/dt) + (1/g)(dV/dt)]$$

where $(T/W)_{\text{TO}}$ is the take-off thrust loading
$\alpha_1 = T/T_{\text{SLS}}$
T_{SLS} = sea level static thrust (all engines)
$\beta = W/W_{\text{To}}$
C_{DO} and k_1 are coefficients in the aircraft drag equation, see below
$D = qS(C_{\text{DO}} + k_1 C_{\text{L}}^2)$
$(W/S)_{\text{TO}}$ is the take-off wing loading (N/m^2)
n is the normal acceleration factor = L/W
g = gravitational acceleration
V is the aircraft forward speed
q is the dynamic pressure = $0.5\rho V^2$
(dh/dt) = rate of climb
(dV/dt) = longitudinal acceleration

5.2.8.5 Combat Turns at SL

In this flight condition the aircraft is in 'sustained' flight with no change in height and no increase in speed therefore the last two terms in the fundamental equation are both zero.

At sea level $\alpha = 1$

Assume that the turn requirement is appropriate to the mean combat mass (i.e. aircraft empty + pilot + half fuel + half weapon load = 3311 + 136 + 450 + 680 = 4577 kg/10 092 lb)

Hence $\beta = 4577/5707 = 0.8$.

From previous analysis (in SI units) the best speed for turning at SL is about 150 m/s.

$$\therefore q = 0.5 \times 1.225 \times 150^2 = 13\,781$$

From the drag analysis done earlier (at 4577 kg with an increase in drag coefficient to represent the stores on the wing) at a speed of 150 m/s, $C_{\text{D}} = 0.03 + 0.017\, C_{\text{L}}^2$.

As specified, the aircraft is subjected to a normal acceleration $n = 4$ in the turn.

$$T/W = 13\,781\,\{(0.03/W/S) + 0.017 \times [4/13\,781]^2 \times (W/S)\}$$

5.2.8.6 Combat Turn at 25 000 ft

This is similar to the analysis above but with $\alpha = 0.557/1.225 = 0.455$.

At 25 000 ft the best speed for excess power is 200 m/s (in SI units)

$$\therefore q = 0.5 \times 0.557 \times 200^2 = 11\,140$$

With β and C_{D} values the same but with load factor $n = 2$ gives:

$$T/W = (0.8/0.445)[(11\,140/0.8)\{(0.03/(W/S) + 0.017 \times [(2 \times 0.8)/11\,140]^2 \times (W/S)\}$$

5.2.8.7 Climb Rate

This criterion assumes a non-accelerating climb, so the last term in the fundamental equation is zero but the penultimate term assumes the value relating to the specified rate of climb.

We will use an average value of climb rate of 18.15 m/s (i.e. 25 000 ft in 7min) and make the calculation at the average altitude of 12 500 ft, at a best aircraft speed of 150 m/s.

At 12 500 ft $\alpha = 0.841/1.255 = 0.686$

At 150 m/s $q = 0.5 \times 0.841 \times 150^2 = 9461$

Using the standard values for β at mean combat mass, and the drag coefficients (C_{DO} and K) previously specified, we get:

$$T/W = (0.8/0.686)[(9461/0.8)\{(0.03/(W/S) + 0.017 \times [(1 \times 0.8)/9461]^2 \times (W/S)\} + 18.15(1/150)$$

5.2.8.8 Constraint Diagram

The above equations have been evaluated for a range of wing loading values (150 to 550 kg/m^2). The resulting curves are shown in Figure 5.2-21.

The constraint diagram shows that the landing constraints (approach speed and ground run) present severe limits on wing loading.

To identify the validity of the constraints relative to other aircraft, values appropriate to specimen (competitor) aircraft that were identified earlier in the study have been plotted on the same constraint diagram Figure 5.2-21. Some interesting conclusions can be drawn from this diagram:

- The S212, T45, MiG, L159 and, to a lesser extent, the Hawk aircraft appear to fit closely to the climb constraint line. This validates this requirement.
- None of the existing aircraft satisfy the landing conditions at $M_{LAND} = 0.9\ M_{TO}$. This suggests that this requirement is too tight.
- The turn requirements do not present critical design conditions for any of the aircraft. The 25 000 ft turn criteria is seen to be the most severe. Some further detailed analysis suggests that the aircraft is capable of a 3 g turn rate at this altitude.

Warning: The constraint analysis described above is a very approximate analytical tool as it does not take into account some of the finer detail of the design (e.g. detailed changes in engine performance with speed). It can only be used in the form presented in the initial design phase. Later in the development of the layout more detailed analysis of the performance will enable the effect of the various constraints on the aircraft design to be better appreciated.

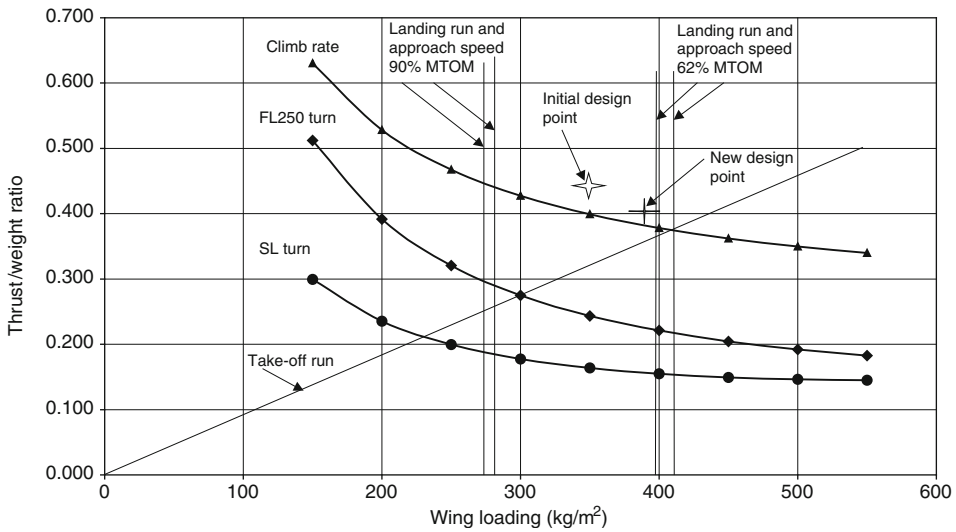


Figure 5.2-21 Aircraft constraint diagram.

However, with this consideration in mind it is possible to use the constraint diagram to direct changes to the original baseline layout as discussed below.

5.2.9 Revised Baseline Layout

The main conclusion from the constraint analysis and aircraft performance estimations is that the aircraft landing requirements are too tight and should be renegotiated with the customers. To provide evidence on the effects of the landing constraints, the revised baseline layout will ignore them. The new design can be analysed to show what landing characteristics are feasible.

With the above strategy in mind the design point for the aircraft will be moved closer to the intersection of the take-off and climb constraint lines, i.e.:

$$(T/W) = 0.38 \quad \text{and} \quad (W/S) = 390\ \text{kg/m}^2 (80\ \text{lb/sq.ft})$$

Anticipating the need to increase aircraft mass to allow more fuel to be carried, the maximum take-off mass is increased to 5850 kg (and the structural design mass increased to 6100 kg). Using the new values for (T/W) and (W/S) the new thrust and wing area become:

$$T = 0.38 \times 5850 = 4900\ \text{lb (SSL)}$$

$$S = 5850/400 = 14.65\ \text{m}^2 (136\ \text{sq.ft})$$

For an aspect ratio (AR) of 5, the new area gives a wing span (b) = 8.56 m and a mean chord = 1.71 m. For an aspect ratio of 4.5 the wing geometry becomes b = 8.12 m and mean chord = 1.80 m. Rounding these figures for convenience of the layout drawing gives:

$$c_{\text{mean}} = 1.75\ \text{m}\ (5.75\ \text{ft}) \quad \text{and} \quad b = 8.5\ \text{m}\ (28\ \text{ft})$$

$$\text{which implies, } AR = 4.86 \quad \text{and} \quad S = 14.87\ \text{sq.m}/160\ \text{sq. ft}$$

This geometry will be used in the new layout.

Also, since the tip chord on the previous layout seemed small, the taper ratio will be increased to 0.33.

$$\text{Hence } C_{\text{mean}} = (C_{\text{tip}} + C_{\text{root}})/2 = 1.75\ \text{m (assumed)}$$

$$\text{With, } (C_{\text{tip}}/C_{\text{root}}) = 0.33$$

$$\text{This gives } C_{\text{root}} = 2.63\ \text{m}/8.6\ \text{ft}, C_{\text{tip}} = 0.87\ \text{m}/2.8\text{ft}$$

5.2.9.1 Wing Fuel Volume

It is now possible to check on the internal fuel volume of the new wing geometry. Assume 15 per cent chord is occupied by trailing edge devices and 33 per cent span is taken by ailerons (assume no fuel in the wing tips ahead of the ailerons).

Although previously the wing thickness was assumed to be 10 per cent, it has now become clear that the aircraft will require substantial internal volume for fuel storage. To anticipate this, the wing thickness will be increased to 15 per cent in the expectation that supercritical wing profiles can be designed to assist in the transonic flow conditions particularly for the high-speed development aircraft.

With the above geometry (see Figure 5.2-22) and assuming 66 per cent of the enclosed volume is available for fuel, gives an internal wing fuel capacity of 0.5 m^3. A total fuel load of 1050 kg equates to a volume of 305 Imp. gal. This requires a volume of 1.385 m^3. It is therefore necessary to house some fuel in the aircraft fuselage (namely $1.385 - 0.5 = 0.885\ \text{m}^3$). This is not uncommon on this type of aircraft. The preferred place to keep the fuel is in the space behind the cockpit and between the engine air intakes. This is close to the aircraft centre of gravity, therefore fuel use will not cause a large centre of gravity movement. For our layout it would be preferable to keep the fuel tank below the wing structural platform to make the wing/fuselage joint simpler. From the original aircraft layout this fuselage space would provide a tank volume of about $1 \times 2 \times 0.5 = 1\ \text{m}^3$. This is satisfactory to meet the internal fuel requirement. Using all of this space for fuel may present a problem for the installation of aircraft systems. To anticipate the need for extra space in the fuselage to house the electronic and communication systems an extra 0.5 m will be added to the length of the fuselage. Moving the engine and intakes back to rebalance the aircraft will also provide a cleaner installation of the intake/wing junction (i.e. moving the intake behind the wing leading edge).

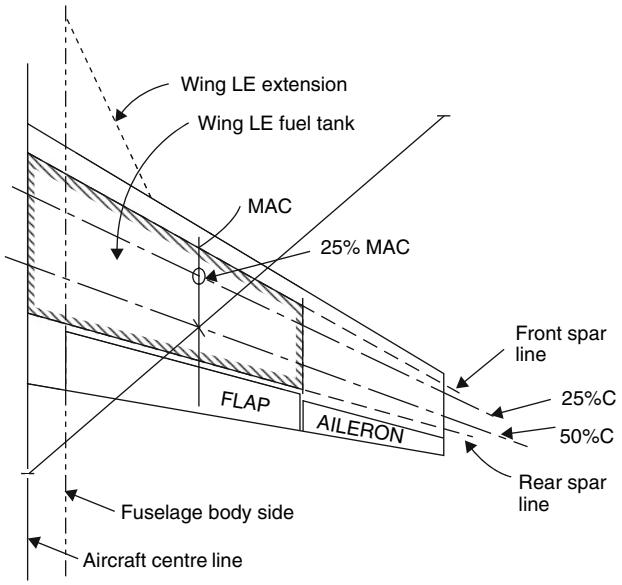


Figure 5.2-22 Revised aircraft wing planform.

Lengthening the fuselage has the effect of increasing the tail effectiveness. This may permit either a traditional low tailplane/fin arrangement, or more likely, a twin fin/tail butterfly layout. Subsequent wind tunnel tests and CFD modelling would be necessary to define the best tail arrangement. In the revised layout a butterfly tail will be shown to illustrate this option.

It is now possible to redraw the baseline layout to account for the above changes. At the same time it is possible to add more details to the geometry (Figure 5.2-23).

5.2.10 Further Work

With the new baseline aircraft drawing available and increased confidence in the aircraft layout it is possible to start a more detailed analyses of the aircraft.

We start this next stage by estimating the mass of each component using the new aircraft geometry as input data for detailed mass predictions. Such equations can be found in most aircraft design textbooks. These formulae have, in general, been derived from data of existing (therefore older) aircraft. As our aircraft will be built using materials and manufacturing methods that have been shown to provide weight savings it will be necessary to apply technology factors to reduce the mass predicted by these older aircraft related methods. The factors that are applied must correspond to the expected degree of mass reduction. Different structural components will require individual factors depending on their layout. For example, the wing structure is more likely to benefit from a change to composite material than the fuselage. The fuselage has many more structural cut-outs and detachable access panels than the wing which makes it less suitable. The mass reduction factors for composite materials may vary between 95 and 75 per cent. The lower value relates to an all-composite structure (e.g. as used for control surfaces and fin structure).

Aspects other than the choice of structural material may also influence the estimation of component mass. Such features may include the requirement for more sophistication in aircraft systems to accommodate the remote instructor concept, the requirements related to the proposal for variability in the flight control and handling qualities of the aircraft to suit basic and advanced training, and the adoption of advanced technology weapon management systems. All such issues and many more will eventually need to be carefully considered when finalising the mass of aircraft components.

When all the component mass estimations have been completed it will be possible to produce a detailed list in the form of an aircraft mass statement. Apart from identifying various aircraft load states, the list can be used to determine aircraft centre of gravity positions. As the aircraft will be used in different training scenarios (e.g. basic aircraft handling experience to full weapon training) it is necessary to determine the aircraft centre of gravity range for different overall loading conditions. With this information it will be possible to balance the aircraft and to

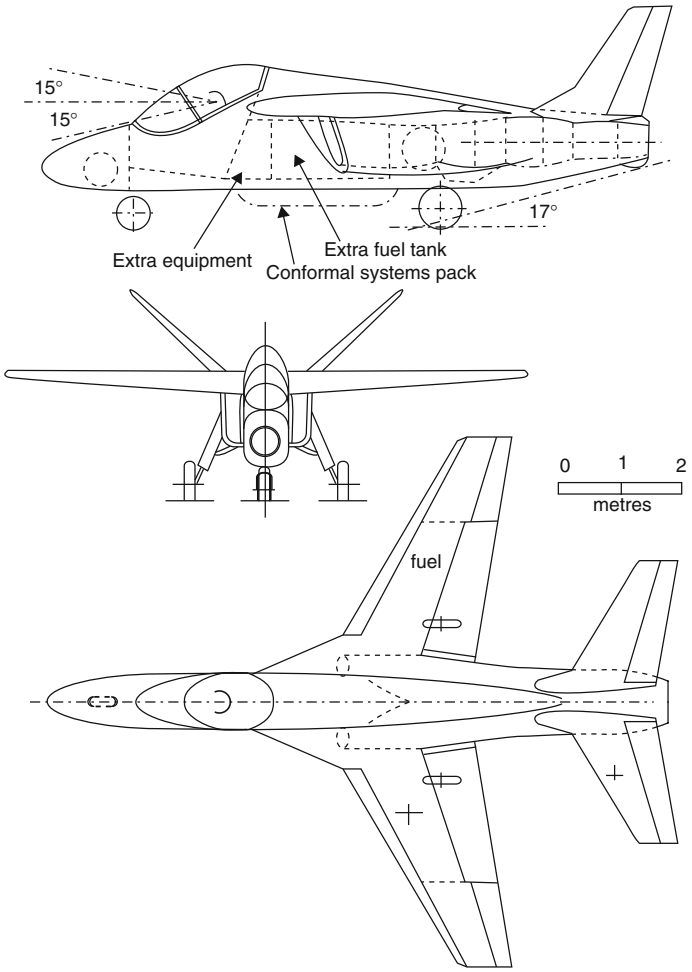


Figure 5.2-23 Revised baseline aircraft layout.

accurately position the wing longitudinally along the fuselage. Up to this point in the design process the wing has been positioned by eye (i.e. guessed).

With the wing position suitably adjusted and a knowledge of the aircraft masses and centre of gravity positions, it is now possible to check the effectiveness of the tail surfaces in providing adequate stability and control forces. Until now the tail sizes have been based on the area ratio and tail volume coefficient values derived from existing aircraft. It is now possible to analyse the control surfaces in more detail to see if they are suitably sized.

The previously crude methods used to determine the aircraft drag coefficients can now be replaced by more detailed procedures. Using the geometry and layout shown in Figure 5.2-23 it is possible to use component drag build-up techniques or panel methods to determine more accurate drag coefficients for the aircraft in different configurations (flap, undercarriage and weapon deployments). Aircraft design textbooks adequately describe how such methods can be used. Likewise, more accurate predictions can now be made for the aircraft lift coefficient at various flap settings.

Before attempting to reassess aircraft performance it is necessary to produce a more accurate prediction of engine performance. If an existing engine is to be used it may be possible to obtain such data from the engine manufacturer. If this is not feasible it will be necessary to devise data from textbooks and other reference material. It may be possible to adapt data available for a known engine of similar type (e.g. equivalent bypass and pressure ratios) by scaling the performance and sizes. Design textbooks suggest suitable relationships to allow such scaling.

More detailed aircraft performance estimations will be centred on point performance. The results will be compared to the values specified in the project brief and subsequent considerations. The crude method used previously will be replaced by flight dynamic calculations (e.g. the take-off and landing estimations will be made using step-by-step time methods).

It is also possible at this stage to use the drag and engine performance estimations to conduct parametric and trade-off studies. These will be useful to confirm or adjust the values used in the layout of the aircraft geometry (for example, the selection of wing aspect ratio, taper, sweepback and thickness).

Further detailed work on the aircraft layout will include:

- The identification and specification of the aircraft structural framework.
- The installation of various aircraft system components. This will require some additional data on the size and mass of each component in the system (e.g. Auxiliary Power Unit).
- A more detailed understanding of the engine installation. This will include the mounting arrangement and access requirements. It will also be necessary to consider the intake and nozzle geometry in more detail.
- Investigate the landing gear mountings and the required retraction geometry.
- Make a more accurate evaluation of the internal fuel tank volumes (wing and fuselage tanks).
- Detailed considerations of the layout requirements for wing control surfaces including flap geometry.

It is obvious that the above list of topics requires a great deal of extra work. All of this is necessary in order to draw the final baseline layout. It would be wasteful to do all of this work without first reviewing the project and considering the overall objectives against the predicted design. The following section outlines the nature of such a review process.

5.2.11 Study Review

There are several different ways in which a design review can be conducted. At the higher level a technique known as a SWOT (strengths, weaknesses, opportunities, threats) analysis can be used, which is also employed in this study as this will illustrate the use of this technique in a design context. However, it must be emphasised that the low- and high-level methods of review are not mutually exclusive and that in some projects it is advisable to use both.

Before starting the review it must be mentioned that the descriptions below do not constitute a complete analysis. A project of this complexity has many facets and it would be too extensive to cover all of them here. The intention is to provide a guide to the main issues that have arisen in the preceding work.

5.2.11.1 Strengths

The most obvious advantage of this project lies in the overall life cycle cost (LCC) savings that are expected from introducing a new advanced technology, training system, approach. If such savings cannot be shown it will be difficult to 'sell' the new system to established air forces. The savings will accrue from the lighter modern aircraft. The use of composites will increase the purchase cost of the aircraft based on the price per unit weight. This would also require extra stringency in inspection of the structure. More elaborate systems will also increase the aircraft first cost. However, the new concept would avoid duplicity of aircraft types in the basic to advanced phase and this will reduce life cycle costs. In addition, the aircrew will have received a higher standard of training from the advanced training system, a consequential reduction of OCR training cost.

The second most powerful advantage for the new concept lies in the ability of the aircraft to more closely match modern fast-jet performance than is currently possible with training aircraft that were originally conceived and designed in the 1970s.

Another strength of the new system is the total integration of modern flight and ground-based systems into a total system design approach. Upgraded older aircraft types are not capable of achieving this aspect of the training system.

Many more advantages could be listed for the system. How many can you identify?

5.2.11.2 Weaknesses

There are three principal weaknesses to the project as currently envisaged. To reduce these deficiencies, if at all possible, it will be necessary to devise strategies or modifications to our design.

The main and intrinsic difficulty lies in the conservative nature of all flight training organisations. This is a natural trait as they take responsibility of human life and national security. As such they will be highly sceptical of the potential advantages of conducting advanced training in a single seat aircraft with a remote instructor. For our concept, as we currently envisage it, this difficulty is insurmountable. Therefore a change of design strategy must be considered to save the credibility of the project. It will be necessary to extend the design concept to encompass a two-seat trainer throughout the full (basic to advanced) training programme. The remote instructor concept can be developed as a separate part of the aircraft/system development programme (i.e. flight testing the aircraft without the instructor present as a proof of concept). This would allow the design and validation of the ground-based instructor system and associated communication and data linking without jeopardising the success of the traditional

design. As we had already accepted that the basic training role would require the development of a two-seat variant, the new strategy will only involve an upgrade to the design to allow the full payload to be carried in this version. Initial calculations suggest that the new aircraft will be about 500 kg (1100 lb) heavier than the existing design (i.e. approximately 10 per cent increase in MTOM). At this point in the development of the project it is obvious that significant changes to the baseline aircraft would be required. Therefore, the work on the present design must be delayed until a revised baseline layout is produced.

The second weakness is associated with the risk involved in the development of new technologies on which the whole system is reliant. If the changes described above are accepted this risk to the project will be avoided. The remaining technologies used in the design can be assured by their current adoption in new aircraft projects (e.g. Eurofighter, F22 and JSF).

The third area relates to the selection of engine for the existing design. From the previous work there are two aspects that require further consideration. First, the Adour engine is shown to be too powerful for our design. The original suggestion (to derate the engine) would only seem to be sensible if the full-rated engine was to be used in future aircraft variants. For the existing trainer aircraft, incorporating an engine larger than necessary effectively adds about 100 kg to the aircraft empty mass. A second propulsion issue relates to fuel usage. Previous calculations showed that the required ferry range was not feasible without seriously penalising the aircraft MTOM. Even to accommodate the fuel required to fly the training sorties it was shown necessary to extend the fuselage to house a larger fuel tank behind the cockpit. For each of the three missions investigated it was found necessary to increase the fuel load that had been previously assumed. As the fuel requirements are directly related to the engine fuel consumption, and thereby to operational cost, it would be advantageous to use a more fuel efficient engine.

Selecting a modern higher-bypass engine with slightly less static sea-level thrust would offer a better design option than using the Adour. Although the engine will be of larger diameter and therefore increase the size of the rear fuselage, it will be lighter and use less fuel. Overall, the change will lead to a lighter and potentially cheaper aircraft.

From the engine data collected earlier (section 5.2.4.3) there are three possible engines from which to choose (specific fuel consumption (sfc) in lb/lb/hr or N/N/hr):

1. TFE 731-60 manufactured by Allied Signal and used on the Citation and Falcon business jets (SSL thrust = 5590 lb, sfc = 0.42, L = 1.83 m, dia. = 0.83 m, depth = 1.04 m, dry mass = 448 kg).
2. CFE 738 (General Electric/ASE) used on the Falcon business jet (5725 lb, sfc = 0.38 SSL, sfc = 0.64 cruise, L = 2.5 m, W = 1.09 m, depth = 1.2 m, mass = 60 1kg).
3. PW 306A (Pratt & Whitney of Canada) used on the Dornier 328 regional jet (5700 lb, sfc = 0.39, L = 2.07 m, W = 0.93 m, depth = 1.15 m, mass = 473 kg).

Aircraft manufacturers prefer to have a choice of available engines as this adds competition on price and delivery. The three engines above are all used on civil aircraft and this may further provide a cost advantage as engine manufacturers will identify an additional market for their product. This should result in a competitive commercial advantage. Approval for military applications will require some extra certification work but this extra cost will be negligible compared to that required to design and develop a completely new engine.

Selecting the PW306A engine would reduce the current dry engine mass by 130 kg (287 lb). This would also reduce the propulsion group mass, thereby reducing the aircraft empty mass. Assuming a cruise specific fuel consumption of 0.64 (as quoted for the equivalent CFE engine) reduces the fuel required to fly the 1000 nm ferry range from the previously estimated 1733 kg for the Adour engine to 1099 kg. This is close to the 900 kg (1985 lb) initially assumed for the fuel mass. The 2000 nm ferry range (assuming external tankage) would require 2328 kg of fuel. This is close to the combined fuel and weapon load (900 + 1360 = 2260 kg/4984 lb) originally specified. Therefore, it appears that by installing this type of engine it would not be necessary to request a reduction in the specified ferry range from originators of the design brief.

The design penalty for installing the higher-bypass type engine lies in the requirement for a larger rear fuselage diameter. The PW306 engine is 0.17 m (7 in) larger in diameter than the Adour. The extra fuselage mass required to house the fatter engine would be more than offset by the reduction in fuel tank weight. The higher bypass ratio engine will also suffer greater loss of thrust with altitude and speed than a pure jet engine.

For designers, the selection of an engine is always a difficult decision as many nontechnical factors may intrude into the process (e.g. political influences, offset cost and manufacturing agreements, national manufacturing preference). Without a knowledge of these influences on this project it is recommended that the PW306A engine is installed. This decision will still allow the other competitor high-bypass engines listed above to be used if commercially advantageous. Alternatively, the Adour engine could be used but this would involve a substantial reduction in aircraft range capability unless external tanks are fitted.

5.2.11.3 Opportunities

Most of the successful training aircraft were originally designed over 20 years ago. Although many have subsequently been 'modernised' they still present old technologies for structure, engines and some systems. The capability of modern fast-jets in the same period has substantially changed the nature of air warfare which has developed with these improved capabilities. This situation opens a wide gap in the effectiveness of old trainers to meet current demands. Here lies the major opportunity for a new trainer design.

Nearly all of the existing successful trainers have been developed into light combat variants for local area defence and ground attack. However, many of these aircraft are of limited capability due to the age of their systems and their inadequate performance. Our new trainer could be developed into an effective combat aircraft to compete with these existing older trainer aircraft variants.

There is therefore substantial worldwide potential for marketing a new trainer and its derivatives.

5.2.11.4 Threats

We are not alone in identifying the need for a new trainer. Two other countries have started to manufacture and develop new trainer aircraft over the past few years. These could present a serious commercial challenge to our project unless we can exploit our advanced technologies to produce a more effective and technically capable design solution.

5.2.11.5 Revised Aircraft Layout

The result of the study review has proposed significant changes to the existing baseline layout. These include:

- a two-seat cockpit,
- a change of engine,
- a requirement for less internal fuel volume.

Each of these changes will effect the aircraft mass and geometry. A revised general arrangement drawing of the new baseline layout is shown in Figure 5.2-24. Initial calculations showed that the increase in aircraft structural mass resulting from the addition of the second seat and larger diameter engine has been offset by the reduction in mass from the lighter engine and the reduced fuel requirement.

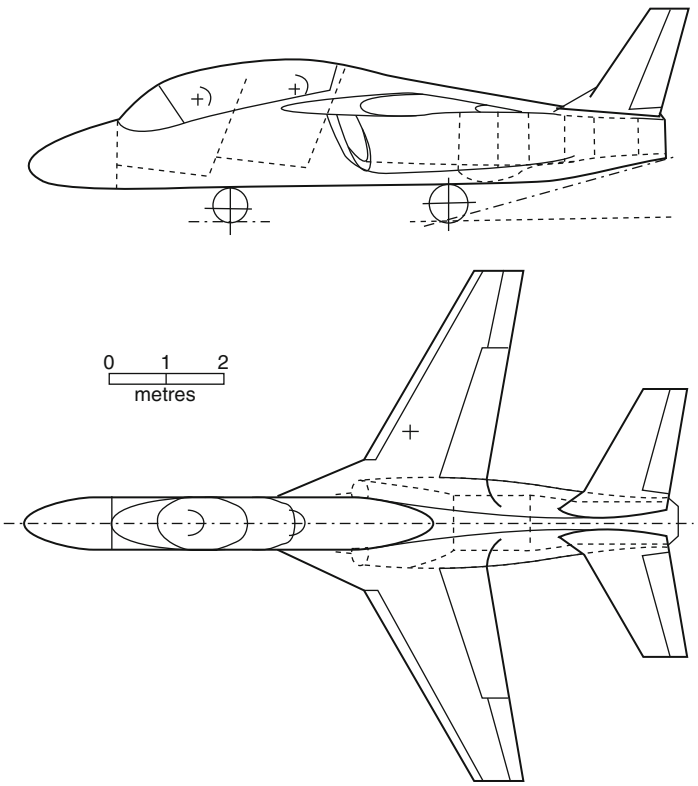


Figure 5.2-24 Post-design review layout (two seat).

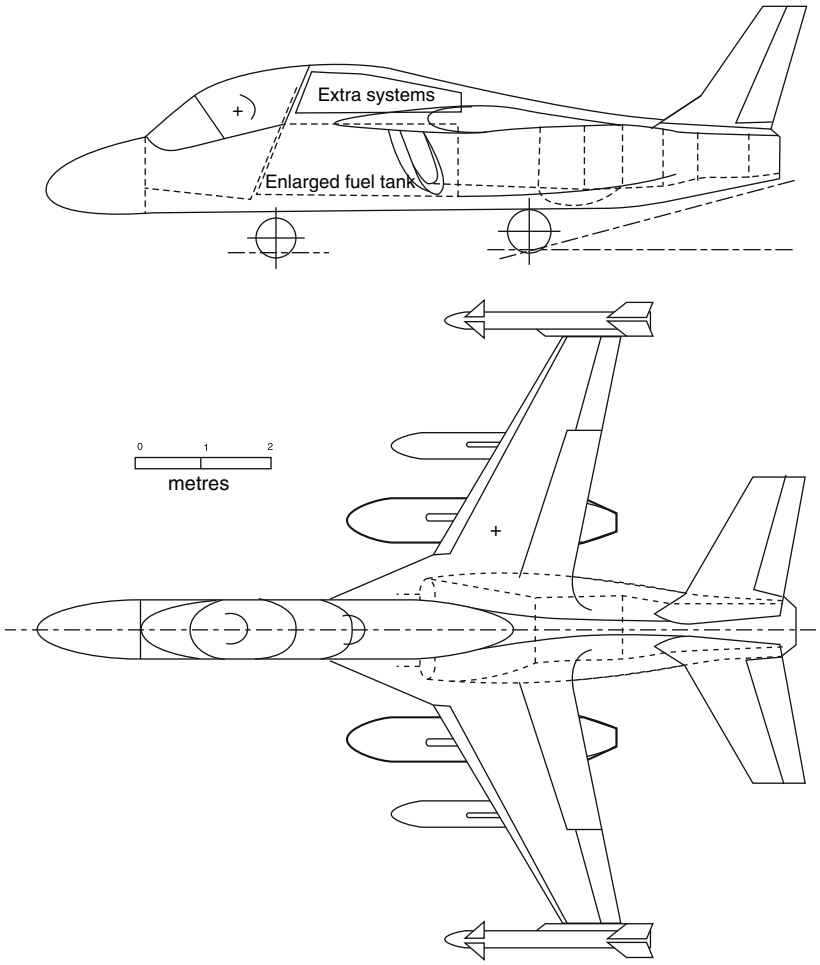


Figure 5.2-25 Single-seat aircraft variant.

The single-seat derivative of the new aircraft would benefit from either a 230 kg/507 lb increase in weapon load, or by an increase in range from the equivalent 230 kg increase in fuel load. The single-seat variant is shown in Figure 5.2-25.

The detailed analysis of the new aircraft follows the same methods as outlined earlier in this chapter. To avoid repetition these calculations have not been included in this chapter.

5.2.12 Postscript

This study has demonstrated how project design decisions may change as the aircraft is more thoroughly understood. This demonstrates the iterative nature of conceptual design. It is possible for students to continue this project into the next iterative stage using the final aircraft drawings (Figure 5.2-25) as the starting point.

References

Textbooks for military aircraft design and performance:

1. Raymer, D. P., *Aircraft Design: A Conceptual Approach*, AIAA Education Series, 1999, ISBN 1-56347-281-0.
2. Brandt, S. A. et al., *Introduction to Aeronautics: A Design Perspective*, AIAA Education Series, 1997, ISBN 1-56347-250-3.
3. McCormick, B. W., *Aerodynamics, Aeronautics and Flight Mechanics*, Wiley and Sons, 1979, ISBN 0-471-03032-5.
4. Nicolai, L. M., *Fundamentals of Aircraft Design*, METS Inc., San Jose, California 95120, USA, 1984.
5. Mattingly, J. D., *Aircraft Engine Design*, AIAA Education Series, 1987, ISBN 0-930403-23-1.

The following publication is also useful in collecting data on existing aircraft:

Aviation Week Source Book, published annually in January.

This handbook is a useful source of general aeronautical data:

AIAA Aerospace Design Engineers Guide, 1998, ISBN 1-56347-283X 1.

This page intentionally left blank

SECTION 6

Avionic Systems

This page intentionally left blank

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.1 VHF Communications

Mike Tooley and David Wyatt

Very high frequency (VHF) radio has long been the primary means of communication between aircraft and the ground. The system operates in the frequency range extending from 118 MHz to 137 MHz and supports both voice and data communication (the latter becoming increasingly important). This chapter describes the equipment used and the different modes in which it operates.

VHF communication is used for various purposes including air traffic control (ATC), approach and departure information, transmission of meteorological information, ground handling of aircraft, company communications, and also for the Aircraft Communications and Reporting System (ACARS).

6.1.1 VHF Range and Propagation

In the VHF range (30 MHz to 300 MHz) radio waves usually propagate as direct line-of-sight (LOS) waves. Sky wave propagation still occurs at the bottom end of the VHF range (up to about 50 MHz depending upon solar activity) but at the frequencies used for aircraft communication, reflection from the ionosphere is exceptionally rare.

Communication by strict LOS paths, augmented on occasions by diffraction and reflection, imposes a limit on the working range that can be obtained. It should also be evident that the range will be dependent on the height of an aircraft above the ground; the greater this is the further the range will be.

The maximum LOS distance (see Figure 6.1-1) between an aircraft and a ground station, in nautical miles (nm), is given by the relationship

$$d = 1.1\sqrt{h}$$

where h is the aircraft's altitude in feet above ground (assumed to be flat terrain).

Example 6.1.1

Determine the maximum LOS distance when an aircraft is flying at a height of (a) 2,500 feet and (b) 25,000 feet.

In (a), $h = 2{,}500$; hence,

$$d = 1.1\sqrt{2{,}500} = 1.1 \times 50 = 55 \text{ nm}$$

In (b), $h = 25{,}000$; hence,

$$d = 1.1\sqrt{25{,}000} = 1.1 \times 158 = 174 \text{ nm}$$

The actual range obtained depends not only on the LOS distance but also on several other factors, including aircraft position, transmitter power, and receiver sensitivity. However, the LOS distance usually provides a good approximation of the range that can be obtained between an aircraft and a ground station (see Table 6.1-1). The situation is slightly more complex when communication is from one aircraft to another; however, in such cases summing the two LOS distances will normally provide a guide as to the maximum range that can be expected.

Test your understanding 6.1.1

Determine the altitude of an aircraft that would provide a LOS distance to a ground station located at a distance of 125 nm.

6.1.2 DSB Modulation

Amplitude modulation is used for voice communications as well as several of the **VHF data link** (VDL) modes. The system uses **double sideband** (DSB) modulation and, because this has implications for the bandwidth of modulated

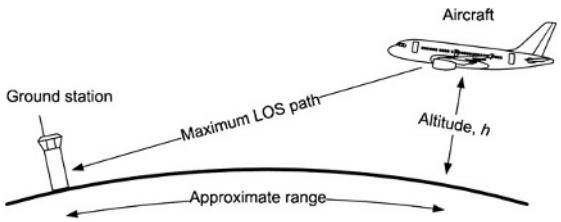


Figure 6.1-1 VHF line-of-sight range.

Table 6.1-1 Theoretical LOS range.

Altitude (feet)	Approx. LOS range (nm)
100	10
1,000	32
5,000	70
10,000	100
20,000	141

signals, it is worth spending a little time explaining how this works before we look at how the available space is divided into channels.

Figure 6.1-2 shows the frequency spectrum of an RF carrier wave at 124.575 MHz amplitude modulated by a single pure sinusoidal tone with a frequency of 1 kHz. Note how the amplitude modulated waveform comprises three separate components:

- an **RF carrier** at 124.575 MHz
- a **lower side frequency** (LSF) component at 124.574 MHz
- an **upper side frequency** (USF) component at 124.576 MHz.

Note how the LSF and USF are spaced away from the RF carrier by a frequency that is equal to that of the modulating signal (in this case 1 kHz). Note also from Figure 6.1-2 that the **bandwidth** (i.e. the range of frequencies occupied by the modulated signal) is *twice* the frequency of the modulating signal (i.e. 2 kHz).

Figure 6.1-3 shows an RF carrier modulated by a speech signal rather than a single sinusoidal tone. The **baseband** signal (i.e. the voice signal itself) typically occupies a frequency range extending from around 300 Hz to 3.4 kHz. Indeed, to improve intelligibility and reduce extraneous noise, the frequency response of the microphone and speech amplifier is invariably designed to select this particular range of frequencies and reject any audio signals that lie outside it. From Figure 6.1-3 it should be noted that the bandwidth of the RF signal is approximately 7 kHz (i.e. twice that of the highest modulating signal).

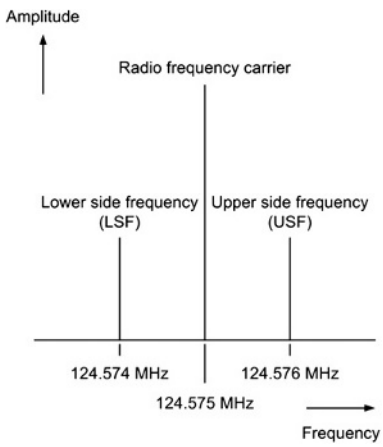


Figure 6.1-2 Frequency spectrum of an RF carrier using DSB modulation and a pure sinusoidal modulating signal.

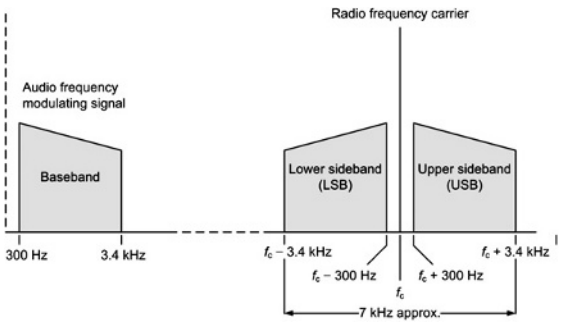


Figure 6.1-3 Frequency spectrum of a baseband voice signal (left) and the resulting DSB AM RF carrier (note that the bandwidth of the RF signal is approximately twice that of the highest modulating signal frequency).

Test your understanding 6.1.3

Determine the RF signal frequency components present in a DSB amplitude modulated carrier wave at 118.975 MHz when the modulating signal comprises pure tones at 2 kHz and 5 kHz.

6.1.3 Channel Spacing

VHF aircraft communications take place in a number of allocated channels. These channels were originally spaced at 200 kHz intervals throughout the VHF aircraft band. However, a relentless increase in air traffic coupled with the increasing use of avionic systems for data link communication has placed increasing demands on the available frequency spectrum. In response to this demand, the spacing between adjacent channels in the band 118 MHz to 137 MHz has been successively reduced so as to increase the number of channels available for VHF communication (see Table 6.1-2).

Figure 6.1-4 shows the channel spacing for the earlier 25 kHz and current European 8.33 kHz VHF systems. Note how the 8.33 kHz system of channel spacing allows three DSB AM signals to occupy the space that was previously occupied by a single signal.

The disadvantage of narrow channel spacing is that the **guard band** of unused spectrum that previously existed with the 25 kHz system is completely absent and that receivers must be designed so that they have a very high degree of **adjacent channel rejection**. Steps must also be taken to ensure that the bandwidth of the transmitted signal does not exceed the 7 kHz, or so, bandwidth required for effective voice communication. The penalty for not restricting the bandwidth is that signals from one channel may 'spill over' into the adjacent channels, causing interference and degrading communication (see Figure 6.1-7).

Test your understanding 6.1.3

How many channels at a spacing of 12.5 kHz can occupy the band extending from 118 MHz to 125 MHz?

Test your understanding 6.1.4

A total of 1520 data channels are to be accommodated in a band extending from 316 MHz to 335 MHz. What channel spacing must be used and what range of frequencies can the baseband signal have?

Table 6.1-2 Increase in the number of available VHF channels.

Date	Frequency range	Channel spacing	Number of channels
1947	118 MHz to 132 MHz	200 kHz	70
1958	118 MHz to 132 MHz	100 kHz	140
1959	118 MHz to 136 MHz	100 kHz	180
1964	118 MHz to 136 MHz	50 kHz	360
1972	118 MHz to 136 MHz	25 kHz	720
1979	118 MHz to 137 MHz	25 kHz	760
1995	118 MHz to 137 MHz	8.33 kHz	2280

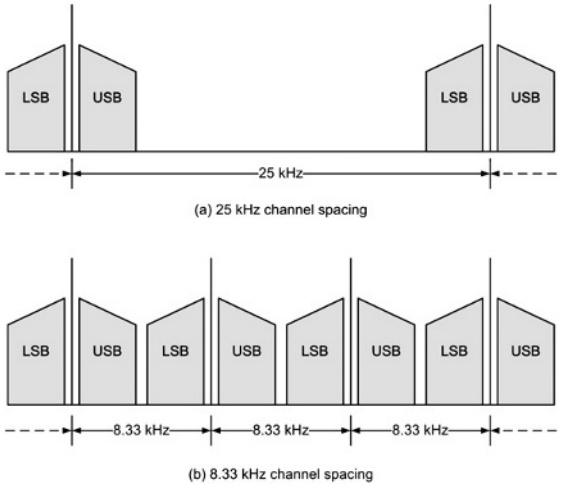


Figure 6.1-4 25 kHz and 8.33 kHz channel spacing.

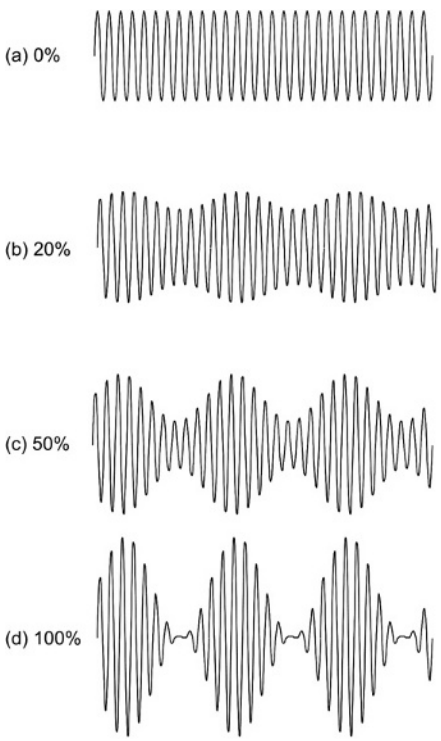


Figure 6.1-5 Different modulation depths.

6.1.4 Depth of Modulation

The depth of modulation of an RF carrier wave is usually expressed in terms of **percentage modulation,** as shown in Figure 6.1-5. Note that the level of modulation can vary between 0% (corresponding to a completely unmodulated carrier) and 100% (corresponding to a fully modulated carrier).

In practice, the intelligibility of a signal (i.e. the ability to recover information from a weak signal that may be adversely affected by noise and other disturbances) increases as the percentage modulation increases and hence there is a need to ensure that a transmitted signal is fully modulated but without the attendant risk of over-modulation (see Figure 6.1-6). The result of over-modulation is excessive bandwidth, or 'splatter', causing adjacent channel interference, as shown in Figure 6.1-7.

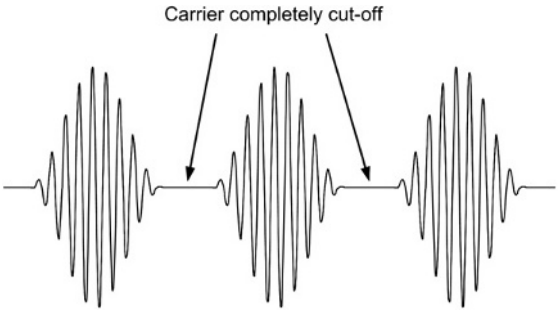


Figure 6.1-6 Over-modulation.

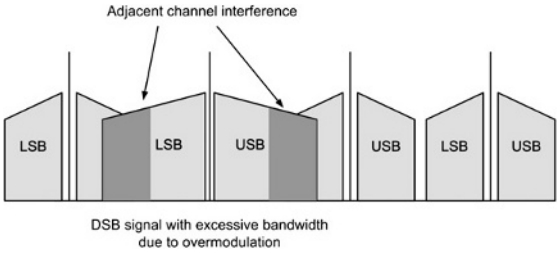


Figure 6.1-7 Adjacent channel interference caused by overmodulation.

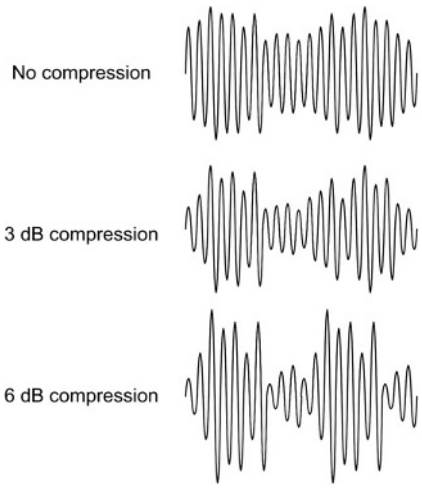


Figure 6.1-8 Modulated RF carrier showing different amounts of compression applied to the modulating signal.

6.1.5 Compression

In order to improve the intelligibility of VHF voice communications, the speech amplifier stage of an aircraft VHF radio is invariably fitted with a **compressor** stage. This stage provides high gain for low amplitude signals and reduced gain for high amplitude signals. The result is an increase in the average modulation depth (see Figure 6.1-8).

Figure 6.1-9 shows typical speech amplifier characteristics with and without compression. Note that most aircraft VHF radio equipment provides adjustment both for the level of modulation and for the amount of compression that is applied (see Figure 6.1-10).

6.1.6 Squelch

Aircraft VHF receivers invariably incorporate a system of muting the receiver audio stages in the absence of an incoming signal. This system is designed to eliminate the annoying and distracting background noise that is present when no signals are being received. Such systems are referred to as **squelch** and the threshold at which this operates is adjusted (see Figure 6.1-10) so that the squelch 'opens' for a weak signal but 'closes' when no signal is present.

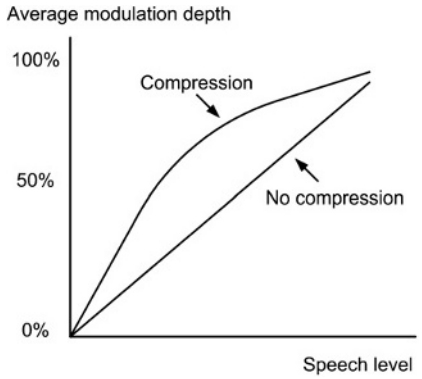


Figure 6.1-9 Effect of compression on average modulation depth.

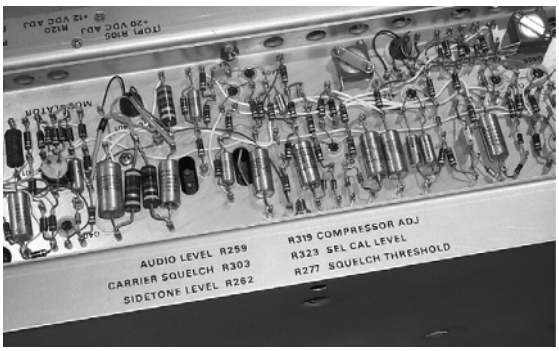


Figure 6.1-10 VHF radio adjustment points.

Two quite different squelch systems are used but the most common (and easy to implement) system responds to the amplitude of the received carrier and is known as **carrier operated squelch**. The voltage used to inhibit the receiver audio can be derived from the receiver's AGC system and fed to the squelch gate (Figure 6.1-11).

The alternative (and somewhat superior) squelch system involves sensing the noise present at the output of the receiver's detector stage and using this to develop a control signal which is dependent on the signal-to-noise ratio of the received signal rather than its amplitude. This latter technique, which not only offers better sensitivity but is also less prone to triggering from general background noise and off-channel signals, is often found in FM receivers and is referred to as **noise operated squelch**.

6.1.7 Data Modes

Modern aircraft VHF communications equipment supports both data communication as well as voice communication. The system used for the aircraft data link is known as ACARS. Currently, aircraft are equipped with three VHF radios, two of which are used for ATC voice communications and one is used for the ACARS data link (also referred to as **airline operational control** communications).

A data link terminal on board the aircraft (see Figure 6.1-12) generates **downlink** messages and processes **uplink** messages received via the VDLs. The downlink and uplink ACARS messages are encoded as plain ASCII text. In the United States, the ACARS ground stations are operated by ARINC, whilst in Europe, Asia and Latin America, the equivalent service is provided by SITA.

Initially each VHF ACARS provider was allocated a single VHF channel. However, as the use of VDLs has grown, the number of channels used in the vicinity of the busiest airports has increased to as many as four and these are often operating at full capacity.

Unfortunately, due to the pressure for additional voice channels, it has not been possible to assign a number of additional VHF channels for ACARS data link operation. As a result, several new data modes have recently been introduced that support higher data rates and make more efficient use of each 25 kHz channel currently assigned for data link purposes (Table 6.1-3).

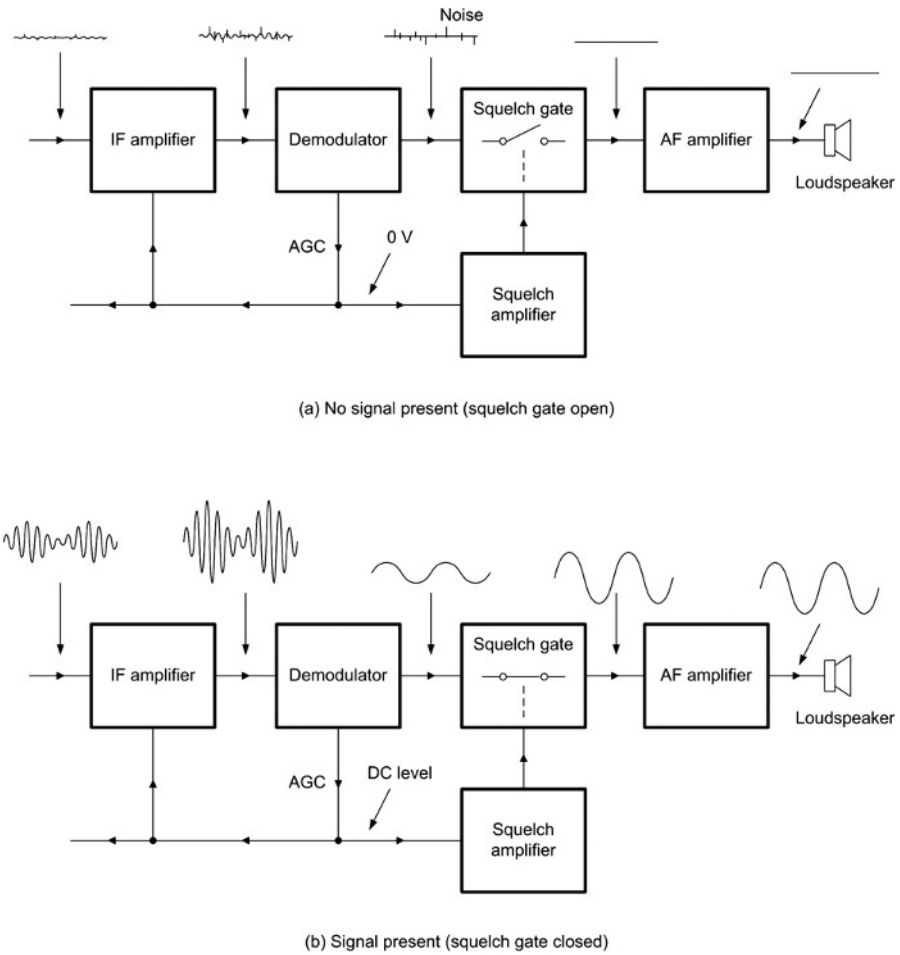


Figure 6.1-11 Action of the squelch system.

In addition, the FAA is developing a system that will permit the integration of ATC voice and data communications. This system uses digitally encoded audio rather than conventional analog voice signals.

When operating in **VDL Mode 0**, the required data link protocols are implemented in the ACARS **management unit** (see Figure 6.1-11). Data is transferred from the VHF radio to the management unit at a rate of 2400 bits per second (bps) by means of **frequency shift keying** (FSK). The FSK audio signal consists of two sinusoidal tones, one at 1.2 kHz and one at 2.4 kHz depending on whether the polarity of the information bit being transmitted is the same as that of the previous bit or is different. Note that the phase of the tones varies linearly and that there is no phase change on the transition between the two tones.

This type of modulation (in which the frequency spacing between the two audio tones is exactly half the data rate) is highly efficient in terms of bandwidth and is thus referred to as **minimum shift keying** (MSK). When data is transmitted, the MSK signal is used to modulate the amplitude of the VHF carrier (in much the same was as the voice signal). The resultant transmitted signal is then a DSB AM signal whose amplitude is modulated at 2400 bps. The RF carrier is then said to use DSB AM MSK modulation.

Table 6.1-3 Summary of voice and data modes.

Mode	Modulation	Channel spacing	Access method	Data rate	Type of traffic	Radio interface
Voice	DSB AM	25/8.33 kHz	PTT	Not applicable	Voice	Analog
Data (Mode 0)	DSB AM MSK	25 kHz	CSMA	2,400 bps	ACARS	Analog
Data (Mode A)	DSB AM MSK	25 kHz	CSMA	2,400 bps	ACARS	ARINC 429
Data (Mode 2)	D8PSK	25 kHz	CSMA	31,500 bps	ACARS and ATN	ARINC 429

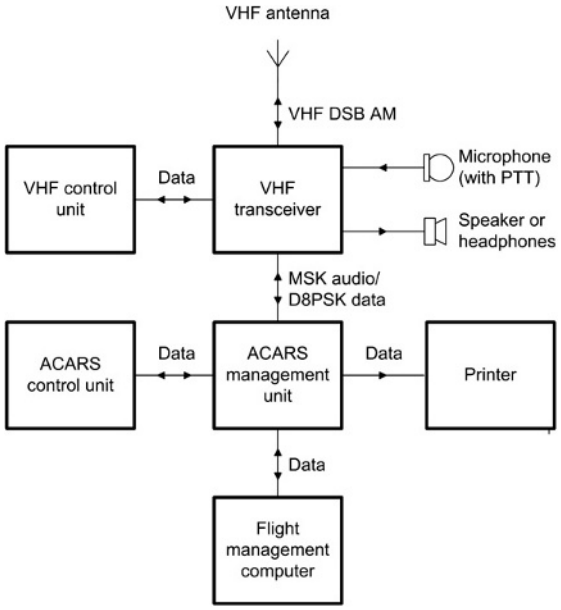


Figure 6.1-12 VHF radio data mangement

VHF carrier frequency selection and transmit/receive control is provided by the ACARS management unit working in conjunction with an ARINC 429 interface to the VHF radio (Figure 6.1-12). The channel access protocol employed is known as **carrier sense multiple access** (CSMA). It consists of listening for activity on the channel (i.e. transmissions from other users) and transmitting only when the channel is free.

Operation in **VDL Mode A** is similar to Mode 0 except uplink and downlink ACARS data packets are transferred between the VHF radio and the ACARS management unit via a transmit/receive pair of 100 kbps ARINC 429 digital interfaces rather than the analog audio interface used by Mode 0. The digital data is then used by the VHF radio to modulate the RF carrier at a rate of 2400 bps using the same DSB AM MSK modulation scheme used by VDL Mode 0.

Another difference between VDL Mode 0 and VDL Mode A is that, when using the latter, the VHF radio controls when to access the channel to transmit data using the same CSMA protocol employed by the management unit in VDL Mode 0. However, the selection of the frequency to be used is still controlled by the CMU or ATSU by means of commands issued via the same ARINC 429 interface used for data transfer. Note that, as far as the VDL ground stations are concerned, there is no difference in the air/ground VDL Mode 0 or VDL Mode A transmissions.

Operation in **VDL Mode 2** is based on an improved set of data transfer protocols and, as a result, it provides a significant increase in data capacity. VDL Mode 2 has been designed to provide for the future migration of VDL to the aeronautical telecommunications network (ATN). This network will permit more efficient and seamless delivery of data messages and data files between aircraft and the ground computer systems used by airlines and ATC facilities.

ATN will be supported by a number of air/ground networks and ground/ground networks. The air/ground and ground/ground networks will be interconnected by means of ATN **routers** that implement the required protocols and will operate in much the same way as the Internet with which you are probably already familiar.

VDL Mode 2 employs a data rate of 31,500 bits per second over the air/ground link using a single 25 kHz channel. The increased utilization of the 25 kHz channel is achieved by employing a system of modulation that is more efficient in terms of its use of bandwidth. This system is known as **differential eight phase shift keying** (D8PSK). In this system, an audio carrier signal is modulated by means of shift in phase that can take one of eight possible phases: 0°, 45°, 90°, 135°, 180°, 225°, 270° or 315°. These phase changes correspond to three bits of digital data as follows: 000, 001, 011, 010, 110, 111, 101, or 100. The D8PSK modulator uses the bits in the data message, in groups of three, to determine the carrier phase change at a rate of 10.5 kHz. Consequently, the bit rate will be three times this value, or 31.5 kbps. D8PSK modulation of the phase of the VHF carrier is accomplished using a **quadrature modulator**. Note that, in D8PSK modulation, groups of three bits are often referred to as **D8PSK symbols**.

VDL Mode 3 offers an alternative to the European solution of reducing the channel spacing to 8.33 kHz. VDL Mode 3 takes a 25 kHz frequency assignment and divides it into 120 ms frames with four 30 ms time slots (each of

```
ACARS mode: E Aircraft reg: N27015
Message label: H1 Block id: 3
Msg no: C36C
Flight id: CO0004
Message content:-
#CFBBY ATTITUDE INDICATOR
MSG 2820121 A 0051 06SEP06 CL H PL
DB FUEL QUANTITY PROCESSOR UNIT
MSG 3180141 A 0024 06SEP06 TA I 23
PL
DB DISPLAYS-2 IN LEFT AIMS
MSG 2394201 A 0005 06SEP06 ES H PL
MSG 2717018
```

Figure 6.1-13 Example of a downlink ACARS message sent from a Boeing 777 aircraft.

which constitutes a different channel). Thus Mode 3 employs **time division multiplexing** (TDM) rather than **frequency division multiplexing** (FDM) used in the European system. Note that VDL Mode 3 is the only planned VDL mode that is designed to support voice and data traffic *on the same frequency*.

6.1.8 ACARS

ACARS is a digital data link system transmitted in the VHF range (118 MHz to 136 MHz). ACARS provides a means by which aircraft operators can exchange data with an aircraft without human intervention. This makes it possible for an airline to communicate with the aircraft in their fleet in much the same way as it is possible to exchange data using a land-based digital network. ACARS uses an aircraft's unique identifier and the system has some features that are similar to those currently used for electronic mail.

The ACARS system was originally specified in the ARINC 597 standard but has been revised as ARINC 724B. A significant feature of ACARS is the ability to provide real-time data on the ground relating to aircraft performance; this has made it possible to identify and plan aircraft maintenance activities.

ACARS communications are automatically directed through a series of ground-based ARINC (Aeronautical Radio Inc.) computers to the relevant aircraft operator. The system helps to reduce the need for mundane HF and VHF voice messages and provides a system which can be logged and tracked. Typical ACARS messages are used to convey routine information such as:

- passenger loads
- departure reports
- arrival reports
- fuel data
- engine performance data.

This information can be requested by the company and retrieved from the aircraft at periodic intervals or on demand. Prior to ACARS this type of information would have been transferred via VHF voice.

ACARS uses a variety of hardware and software components including those that are installed on the ground and those that are present in the aircraft. The aircraft ACARS components include a **management unit** (see Figure 6.1-12) which deals with the reception and transmission of messages via the VHF radio transceiver, and the **control unit** which provides the crew interface and consists of a display screen and printer. The ACARS **ground network** comprises the ARINC ACARS remote transmitting/receiving stations and a network of computers and switching systems. The ACARS **command, control and management subsystem** consists of the ground-based airline operations and associated functions including operations control, maintenance and crew scheduling.

There are two types of ACARS messages: downlink messages that originate from the aircraft and **uplink** messages that originate from ground stations (see Figures 6.1-14 to 6.1-17). Frequencies used for the transmission and reception of ACARS messages are in the band extending from 129 MHz to 137 MHz (VHF) as shown in Table 6.1-4. Note that different channels are used in different parts of the world. A typical ACARS message (see Figure 6.1-14) consists of:

```
ACARS mode: 2
Aircraft reg: G-DBCC
Message label: 5U
Block id: 4
Msg no: M55A
Flight id: BD01NZ
Message content:-
01 WXRQ 01NZ/05 EGLL/EBBR .G-DBCC
/TYP 4/STA EBBR/STA EBOS/STA EBCI
```

Figure 6.1-14 Example of an ACARS message (see text).

```
ACARS mode: 2 Aircraft reg: N788UA
Message label: RA Block id: L
Msg. no: QUHD
Flight id: QWDUA~
Message content:-
WEIGHT MANIFEST
UA930 SFOLHR
SFO
ZFW    383485
TOG    559485
MAC    40.1
TRIM   02.8
PSGRS  285
```

Figure 6.1-15 Example of aircraft transmitted data (in this case, a weight manifest).

```
ACARS mode: X Aircraft reg: N199XX
Message label: H1 Block id: 7
Msg no: F00M
Flight id: GS0000
Message content:-
#CFBER FAULT/WRG [SWPA2]
INTERFACE
TCAS FAIL ADVISORY
TERRAIN 1 FAIL ADVISORY
TERRAIN 1-2 FAIL ADVISORY
THROTTLE QUADRANT 1-2 FAIL ADVISORY
22-10 221009ATA1 OC=1
TQA FAULT [ATA1]
INTERFACE
22-10 221009ATA
```

Figure 6.1-16 Example of a failure advisory message transmitted from an aircraft.

- mode identifier (e.g. 2)
- aircraft identifier (e.g. G-DBCC)
- message label (e.g. 5U—a weather request)
- block identifier (e.g. 4)
- message number (e.g. M55A)
- flight number (e.g. BD01NZ)
- message content (see Figure 6.1-14).

```
ACARS mode: R Aircraft reg: G-EUPR
Message label: 10 Block id: 8
Msg no: M06A
Flight id: BA018Z
Message content:-
FTX01.ABZKOBA
BA1304
WE NEED ENGINEERING TO DO PDC ON
NUMBER 2 IDG
CHEERS
ETL 0740 GMT
```

Figure 6.1-17 Example of a plain text message sent via ACARS.

Table 6.1-4 ACARS channels.

Frequency	ACARS service
129.125 MHz	USA and Canada (additional)
130.025 MHz	USA and Canada (secondary)
130.450 MHz	USA and Canada (additional)
131.125 MHz	USA (additional)
131.475 MHz	Japan (primary)
131.525 MHz	Europe (secondary)
131.550 MHz	USA, Canada, Australia (primary)
131.725 MHz	Europe (primary)
136.900 MHz	Europe (additional)

Test your understanding 6.1.5

Explain the need for (a) speech compression and (b) squelch in an aircraft VHF radio.

Test your understanding 6.1.6

Explain, with the aid of a block diagram, how data transfer is possible using an aircraft VHF radio.

Test your understanding 6.1.7

Explain the difference between MSK and D8PSK modulation. Why is the latter superior?

6.1.9 VHF Radio Equipment

The typical specification of a modern aircraft VHF data radio is shown in Table 6.1-5. This radio can be used with analog voice as well as data in Modes 0, A and 2. Figures 6.1-18 to 6.1-20 show typical equipment and control locations in a passenger aircraft whilst Figures 6.1-21 to 6.1-24 show internal and external views of a typical VHF radio. Finally, Figure 6.1-25 shows a typical VHF quarter-wave blade antenna fitted to an Airbus A380 aircraft.

6.1.10 Multiple Choice Questions

1. The angle between successive phase changes of a D8PSK signal is:
   (a) 45°
   (b) 90°
   (a) 180°.

Table 6.1-5 Aircraft VHF radio specifications.

Parameter	Specification
Frequency range	118.00 MHz to 136.99167 MHz
Channel spacing	8.33 kHz or 25 kHz
Operating modes	Analog voice (ARINC 716); Analog date 2400 bps AM MSK ACARS (external modem); ARINC 750 Mode A analog data 2400 bps AM MSK ACARS; Mode 2 data 31.5 kbps D8PSK
Sensitivity	2 μV for 6 dB (S+N)/N
Selectivity	6 dB max. attenuation at ±16 kHz
(25 kHz channels)	60 dB min. attenuation at ±34 kHz
Selectivity	6 dB max. attenuation at ±5.5 kHz
(8.33 kHz channels)	60 dB min. attenuation at ±14.7 kHz
Audio power output	Adjustable from less than 50 μW to 50 m into 600 Ω ± 20%
RF output power	25 W min. DSB AM operation
	18 W min. D8PSK operation
Frequency stability	±0.005%
Modulation level	0.25 V RMS input at 1 kHz will modulate the transmitter at least 90%
Speech processing	Greater than 20 dB of compression
Mean time between failure	Greater than 40,000 hours

Figure 6.1-18 Three VHF radios (on the extreme left) installed in the aircraft's avionic equipment bay.

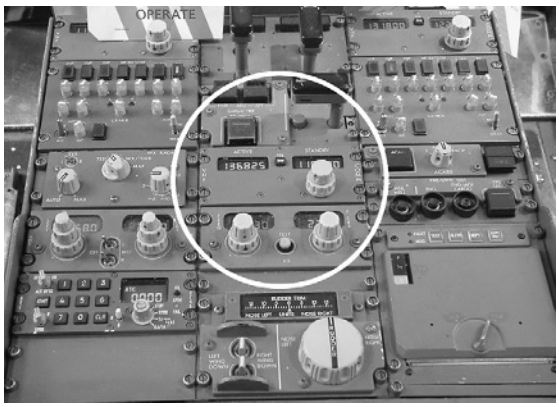


Figure 6.1-19 VHF communications frequency selection panel (immediately above the ILS panel).

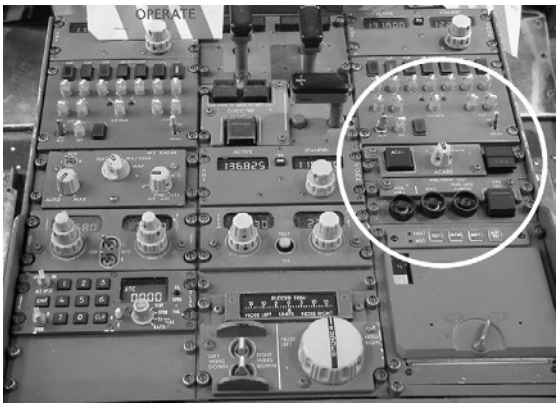


Figure 6.1-20 ACARS control panel (immediately to the right of the VHF communications frequency selection panel).

Figure 6.1-21 Aircraft VHF radio removed from its rack mounting.

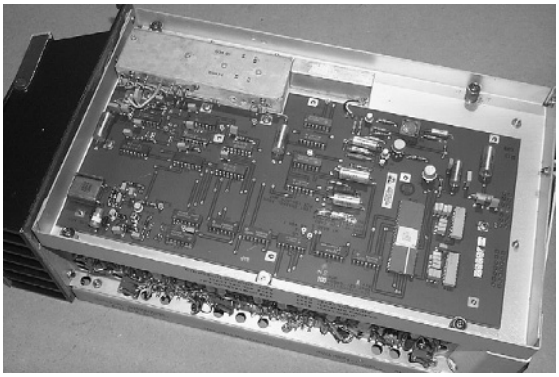


Figure 6.1-22 Digital frequency synthesiser stages of the VHF radio. The quartz crystal controlled reference oscillator is at the bottom left corner and the frequency divider chain runs from left to right with the screened VCO at the top.

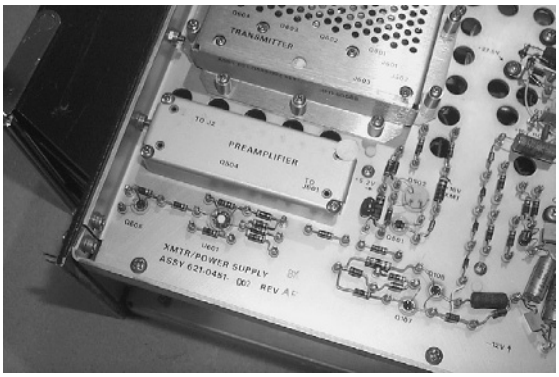


Figure 6.1-23 Screened receiver pre-amplifier and transmitter power amplifier stages (top).

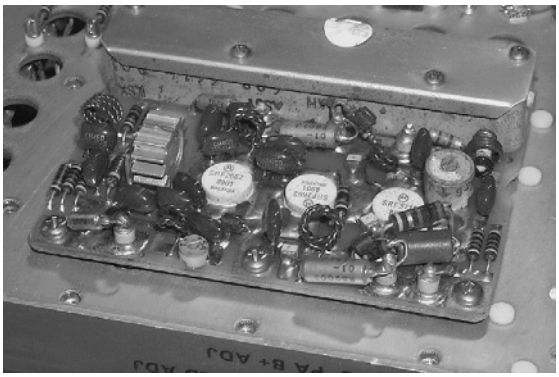


Figure 6.1-24 RF power amplifier stages with the screening removed. There are three linear power stages and one driver (left).

Figure 6.1-25 The forward quarter-wave VHF blade antenna on the Airbus A380.

2. The method of modulation currently employed for aircraft VHF voice communication is:
   (a) MSK
   (b) D8PSK
   (c) DSB AM.

3. The channel spacing currently used in Europe for aircraft VHF voice communication is:
   (a) 8.33 kHz and 25 kHz
   (b) 12.5 kHz and 25 kHz
   (c) 25 kHz and 50 kHz.

4. Which one of the following gives the approximate LOS range for an aircraft at an altitude of 15,000 feet?
   (a) 74 nm
   (b) 96 nm
   (c) 135 nm.

5. The function of the compressor stage in an aircraft VHF radio is:
   (a) to reduce the average level of modulation
   (b) to increase the average level of modulation
   (c) to produce 100% modulation at all times.

6. The function of the squelch stage in an aircraft VHF radio is:
   (a) to eliminate noise when no signal is received
   (b) to increase the sensitivity of the receiver for weak signals
   (c) to remove unwanted adjacent channel interference.

7. Large passenger aircraft normally carry:
   (a) two VHF radios
   (b) three VHF radios
   (c) four VHF radios.

8. The typical bandwidth of a DSB AM voice signal is:
   (a) 3.4 kHz
   (b) 7 kHz
   (c) 25 kHz.

9. The disadvantage of narrow channel spacing is:
   (a) the need for increased receiver sensitivity
   (b) the possibility of adjacent channel interference
   (c) large amounts of wasted space between channels.

10. The standard for ACARS is defined in:
    (a) ARINC 429
    (b) ARINC 573
    (c) ARINC 724.

11. The frequency band currently used in Europe for aircraft VHF voice communication is:
    (a) 88 MHz to 108 MHz
    (b) 108 MHz to 134 MHz
    (c) 118 MHz to 137 MHz.

12. The typical output power of an aircraft VHF radio using voice mode is:
    (a) 25 W
    (b) 150 W
    (c) 300 W.

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.2 HF Communications

Mike Tooley and David Wyatt

High frequency (HF) radio provides aircraft with an effective means of communication over long distance oceanic and trans-polar routes. In addition, global data communication has recently been made possible using strategically located HF data link (HFDL) ground stations. These provide access to ARINC and SITA airline networks. HF communication is thus no longer restricted to voice and is undergoing a resurgence of interest due to the need to find a means of long distance data communication that will augment existing VHF and SATCOM data links.

An aircraft HF radio system operates on spot frequencies within the HF spectrum. Unlike aircraft VHF radio, the spectrum is not divided into a large number of contiguous channels but aircraft allocations are interspersed with many other services, including short wave broadcasting, fixed point-to-point, marine and land-mobile, government and amateur services. This chapter describes the equipment used and the different modes in which it operates.

6.2.1 HF Range and Propagation

In the HF range (3 MHz to 30 MHz) radio waves propagate over long distances due to reflection from the ionised layers in the upper atmosphere. Due to variations in height and intensities of the ionised regions, different frequencies must be used at different times of day and night and for different paths. There is also some seasonal variation (particularly between winter and summer). Propagation may also be disturbed and enhanced during periods of intense solar activity. The upshot of this is that HF propagation has considerable vagaries and is far less predictable than propagation at VHF.

Frequencies chosen for a particular radio path are usually set roughly mid-way between the lowest usable frequency (LUF) and the maximum usable frequency (MUF). The daytime LUF is usually between 4 to 6 MHz during the day, falling rapidly after sunset to around 2 MHz. The MUF is dependent on the season and sunspot cycle but is often between 8 MHz and 20 MHz. Hence a typical daytime frequency for aircraft communication might be 8 MHz whilst this might be as low as 3 MHz during the night. Typical ranges are in the region of 500 km to 2500 km and this effectively fills in the gap in VHF coverage (see Figure 6.2-1).

As an example of the need to change frequencies during a 24-hour period, Figure 6.2-2 shows how the service provided by the Santa Maria HF oceanic service makes use of different parts of the HF spectrum at different times of the day and night. Note the correlation between the service availability chart shown in Figure 6.2-2(a) and the typical variation in MUF for the radio path between Madrid and New York.

The following HF bands are allocated to the aeronautical service:

- 2850 to 3155 kHz
- 3400 to 3500 kHz
- 4650 to 4750 kHz
- 5480 to 5730 kHz
- 6525 to 6765 kHz
- 8815 to 9040 kHz
- 10,005 to 10,100 kHz
- 11,175 to 11,400 kHz
- 13,200 to 13,360 kHz
- 15,010 to 15,100 kHz
- 17,900 to 18,030 kHz
- 21,870 to 22,000 kHz
- 23,200 to 23,350 kHz.

6.2.2 SSB Modulation

Unfortunately, the spectrum available for aircraft communications at HF is extremely limited. As a result, steps are taken to restrict the bandwidth of transmitted signals, for both voice and data. **Double sideband** (DSB) amplitude modulation requires a bandwidth of at least 7 kHz but this can be reduced by transmitting only one of the two sidebands. Note that either the **upper sideband** (USB) or the **lower sideband** (LSB) can be used because they both

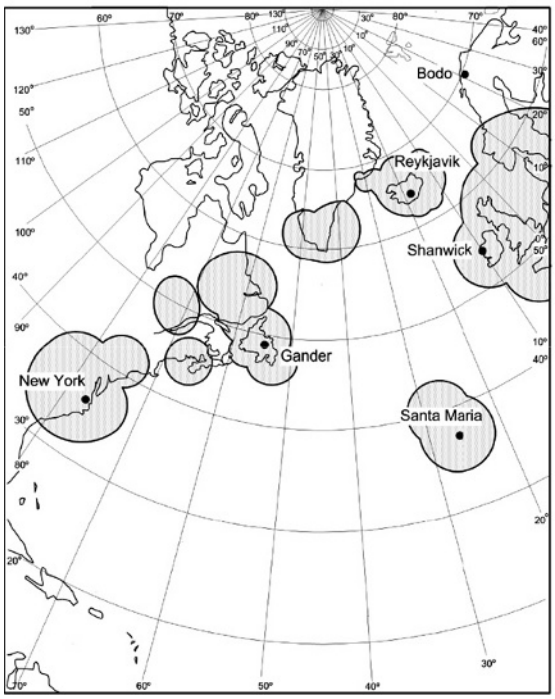


Figure 6.2-1 VHF aircraft coverage in the North Atlantic area.

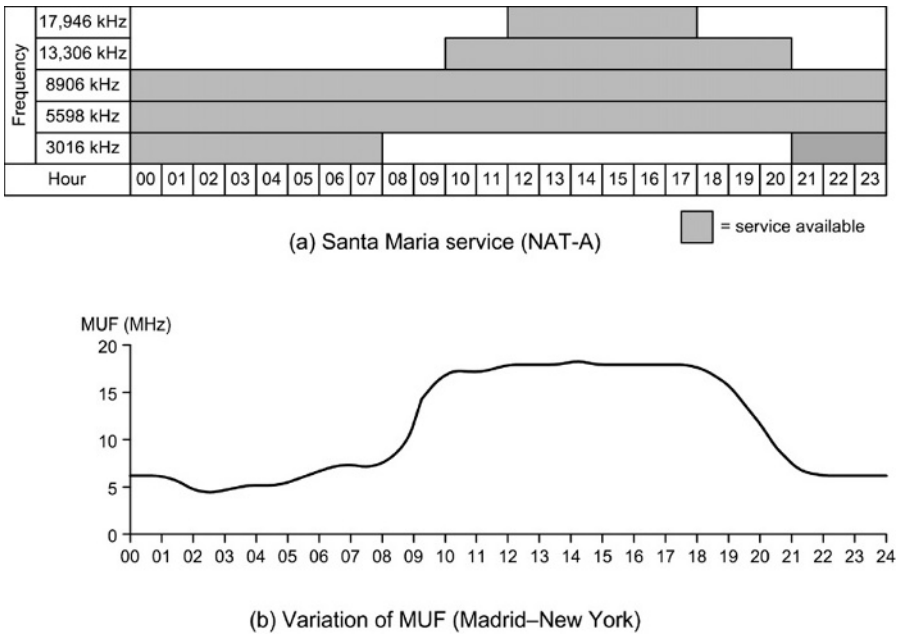


Figure 6.2-2 Santa Maria oceanic service (NAT-A) showing operational frequencies and times together with typical variation of MUF for a path from Madrid to New York.

contain the same modulating signal information. In addition, it is possible to reduce (or 'suppress') the carrier as this, in itself, does not convey any information.

In order to demodulate a signal transmitted without a carrier it is necessary to reinsert the carrier at the receiving end (this is done in the demodulator stage where a beat frequency oscillator or **carrier insertion oscillator** replaces the missing carrier signal at the final intermediate frequency—see Figure 6.2-9). The absence of the carrier means that less power is wasted in the transmitter which consequently operates at significantly higher efficiency.

Figure 6.2-3 shows the frequency spectrum of an RF signal using different types of amplitude modulation, with and without a carrier.

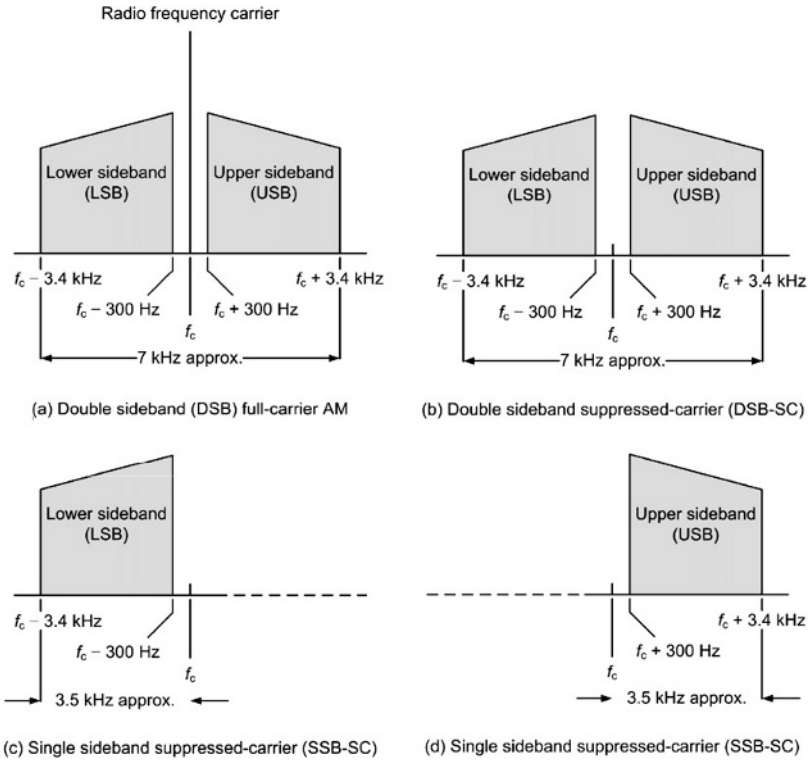


Figure 6.2-3 Frequency spectrum of an RF carrier using DSB and SSB modulation.

In Figure 6.2-3(a) the mode of transmission is conventional DSB amplitude modulation with full-carrier. This form of modulation is used for VHF aircraft communications and was described earlier in Chapter 6.1.

Figure 6.2-3(b) shows the effect of suppressing the carrier. This type of modulation is known as **double sideband suppressed-carrier** (DSB-SC). In practical DSB-SC systems the level of the carrier is typically reduced by 30 dB, or more. The DSB-SC signal has the same overall bandwidth as the DSB full-carrier signal but the reduction in carrier results in improved efficiency as well as reduced susceptibility to heterodyne interference.

Figure 6.2-3(c) shows the effect of removing both the carrier and the USB. The resulting signal is referred to as **single sideband** (SSB), in this case using only the LSB. Note how the overall bandwidth has been reduced to only around 3.5 kHz, i.e. half that of the comparable DSB AM signal shown in Figure 6.2-3(a).

Finally, Figure 6.2-3(d) shows the effect of removing the carrier and the lower sideband. Once again, the resulting signal is referred to as SSB, but in this case we are using only the USB. Here again, the overall bandwidth has been reduced to around 3.5 kHz. Note that aircraft HF communication requires the use of the USB. DSB AM may also be available but is now very rarely used due to the superior performance offered by SSB.

Test your understanding 6.2.1

1. Explain why HF radio is used on trans-oceanic routes.
2. Explain why different frequencies are used for HF aircraft communications during the day and at night.
3. State TWO advantages of using SSB modulation for aircraft HF communications.

6.2.3 SELCAL

Selective calling (SELCAL) reduces the burden on the flight crew by alerting them to the need to respond to incoming messages. SELCAL is available at HF and VHF but the system is more used on HF. This is partly due to the intermittent nature of voice communications on long oceanic routes and partly due to the fact that squelch systems are more difficult to operate when using SSB because there is no transmitted carrier to indicate that a signal is present on the channel.

The aircraft SELCAL system is defined in Annex 10 to the Convention on International Civil Aviation (ICAO), Volume 1, 4th edition of 1985 (amended 1987). The system involves the transmission of a short burst of audio tones.

Each transmitted code comprises two consecutive tone pulses, with each pulse containing two simultaneously transmitted tones. The pulses are of 1 second duration separated by an interval of about 0.2 seconds. To ensure

Table 6.2-1 SELCAL tone frequencies.

Character	Frequency
A	312.6 Hz
B	346.7 Hz
C	384.6 Hz
D	426.6 Hz
E	473.2 Hz
F	524.8 Hz
G	582.1 Hz
H	645.7 Hz
J	716.1 Hz
K	794.3 Hz
L	881.0 Hz
M	977.2 Hz
P	1083.9 Hz
Q	1202.3 Hz
R	1333.5 Hz
S	1479.1 Hz

proper operation of the SELCAL decoder, the frequency of the transmitted tones must be held to an accuracy of better than ±0.15%.

SELCAL codes are uniquely allocated to particular aircraft by Air Traffic Control (ATC). As an example, a typical transmitted SELCAL code might consist of a 1 second burst of 312.6 Hz and 977.2 Hz followed by a pause of about 0.2 seconds and a further 1 second burst of tone comprising 346.7 Hz and 977.2 Hz. Table 6.2-1 indicates that the corresponding transmitted SELCAL code is 'AM-BM' and only the aircraft with this code would then be alerted to the need to respond to an incoming message.

The RF signal transmitted by the ground radio station should contain (within 3 dB) equal amounts of the two modulating tones and the combination of tones should result in a modulation envelope having a nominal modulation percentage as high as possible (and in no case less than 60%).

The transmitted tones are made up from combinations of the tones listed in Table 6.2-1. Note that the tones have been chosen so that they are not harmonically related (thus avoiding possible confusion within the SELCAL decoder when harmonics of the original tone frequencies might be present in the demodulated waveform).

6.2.4 HF Data Link

ARINC's global HFDL coverage provides a highly cost-effective data link capability for carriers on remote oceanic routes, as well as the polar routes at high latitudes where SATCOM coverage is unavailable. HFDL is lower in cost than SATCOM and many carriers are using HFDL instead of satellite services, or as a backup system. HFDL is still the only data link technology that works over the North Pole, providing continuous, uninterrupted data link coverage on the popular polar routes between North America and eastern Europe and Asia.

The demand for HFDL has grown steadily since ARINC launched the service in 1998, and today HFDL avionics are offered as original equipment by all the major airframe manufacturers. HFDL offers a cost-effective solution for global data link service. The demand for HFDL service is currently growing by more than several hundred aircraft per year.

Advantages of HFDL can be summarised as:

- wide coverage due to the extremely long range of HF signals
- simultaneous coverage on several bands and frequencies (currently 60)
- multiple ground stations (currently 14) at strategic locations around the globe
- relatively simple avionics using well-tried technology
- rapid network acquisition
- exceptional network availability.

Disadvantages of HFDL are:

- very low data rates (making the system unsuitable for high-speed wideband communications).

As a result of the above, the vast majority of HFDL messages are related to **airline operational control** (AOC) (see Figure 6.2-4) but HFDL is also expected to play an important part in **future air navigation systems** (FANS) where it

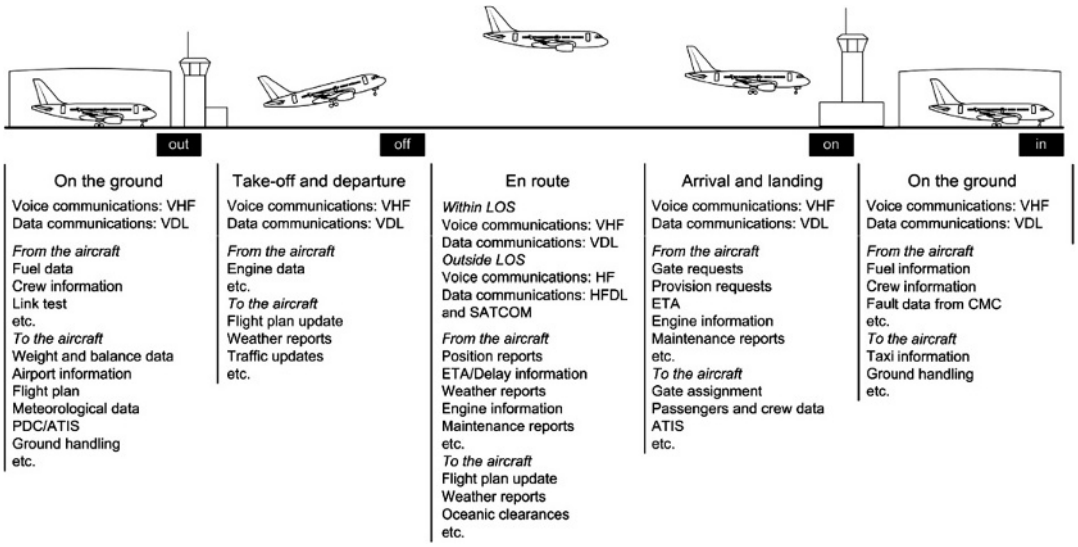


Figure 6.2-4 Aircraft operational control at various 'out-off-on-in' (OOOI) stages.

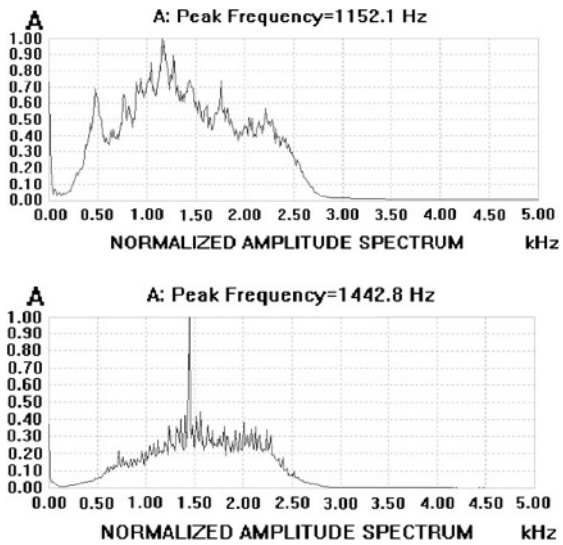


Figure 6.2-5 Frequency spectra of voice (upper trace) and HFDL signals (lower trace).

will provide a further means of data linking with an aircraft, supplementing VHF data link (VDL), GPS, and SATCOM systems. Note that SATCOM can support much faster data rates but it can also be susceptible to interruptions and may not be available at high latitudes.

HFDL uses **phase shift keying** (PSK) at data rates of 300, 600, 1200 and 1800 bps. The rate used is dependent on the prevailing propagation conditions. HFDL is based on **frequency division multiplexing** (FDM) for access to ground station frequencies and **time division multiplexing** (TDM) within individual communication channels. Figure 6.2-5 shows how the frequency spectrum of a typical HFDL signal at 300 bps compares with an HF voice signal.

Figure 6.2-6 shows typical HFDL messages sent from the four aircraft shown in Figure 6.2-7 to the Shannon HFDL ground station using the same communications channel. The radio path from one of the aircraft (LH8409) is illustrated in Figure 6.2-8. The first two of the messages shown in Figure 6.2-6 are **log-on requests** and the maximum bit rate is specified in the header. In each log-on request, the aircraft is identified by its unique 24-bit **ICAO address**. Once logged on, the aircraft is allocated an 8-bit **address code** (AF hex in the case of the third message and AD hex in the case of the fourth message). Each aircraft also transmits its current location data (longitude and latitude).

The system used for HFDL data exchange is specified in ARINC 635. Each ground station transmits a frame called a 'squitter' every 32 seconds. The **squitter frame** informs aircraft of the system status, provides a timing reference and provides protocol control. Each ground station has a time offset for its squitters. This allows aircraft to jump

```
Preamble 300  bps 1.8 sec Interleaver FREQ ERR 5.398116 Hz Errors 0
[MPDU AIR CRC PASS]
Nr LPDUs = 1 Ground station ID SHANNON - IRELAND SYNCHED
Aircraft ID LOG-ON
Slots Requested medium = 0 Low = 0
Max Bit rate 1800 bps U(R) = 0 UR(R)vect = 0
[LPDU LOG ON DLS REQUEST] ICAO AID 0A123C
[HFNPDU FREQUENCY DATA]
14:45:24  UTC  Flight ID = AB3784  LAT 39 37 10  N  LON 0 21 20  W
07 87 FF 00 04 00 14 85 92 BF 3C 12 0A FF D5     ..............
41 42 33 37 38 34 C8 C2 31 BF FF C2 67 88 8C     A B 3 7 8 4 ..1 ...g ..
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00     ..............
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00     ..............
00 00 00 00 00 00 00             .......

Preamble 300  bps 1.8 sec Interleaver FREQ ERR -18.868483 Hz Errors 19
[MPDU AIR CRC PASS]
Nr LPDUs = 1 Ground station ID SHANNON - IRELAND SYNCHED
Aircraft ID LOG-ON
Slots Requested medium = 0 Low = 0
Max Bit rate 1200 bps U(R) = 0 UR(R)vect = 0
[LPDU LOG ON DLS REQUEST] ICAO AID 4A8002
[HFNPDU FREQUENCY DATA]
14:45:30  UTC  Flight ID = SU0106  LAT 54 42 16  N  LON 25 50 42  E
07 87 FF 00 03 00 14 80 1E BF 02 80 4A FF D5     ............J ..
53 55 30 31 30 36 6A 6E F2 60 12 C5 67 33 FB     S U 0 1 0 6 j n ....g 3 .
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00     ..............
00 00 00 00 00 00 00 00 00 00 00 00 00 00 00     ..............
00 00 00 00 00 00 00             .......

Preamble 300  bps 1.8 sec Interleaver FREQ ERR 15.059247 Hz Errors 2
[MPDU AIR CRC PASS]
Nr LPDUs = 1 Ground station ID SHANNON - IRELAND SYNCHED
Aircraft ID AF
Slots Requested medium = 0 Low = 0
Max Bit rate 1200 bps U(R) = 0 UR(R)vect = 0
[LPDU UNNUMBERED DATA]
[HFNPDU PERFORMANCE]
14:45:30  UTC  Flight ID = LH8409  LAT 46 42 34  N  LON 21 22 55  E
07 87 AF 00 03 00 31 4D 1D 0D FF D1 4C 48 38     ......1 M ....L H 8
34 30 39 73 13 82 34 0F C5 67 01 36 03 02 02     4 0 9 s ..4 ..g .6 ...
00 B6 00 00 00 00 00 00 00 03 00 00 00 00 00     ..............
02 00 00 00 00 00 01 00 00 00 01 01 D3 EA 00     ..............
00 00 00 00 00 00 00             .......

Preamble 300  bps 1.8 sec Interleaver FREQ ERR 8.355845 Hz Errors 0
[MPDU AIR CRC PASS]
Nr LPDUs = 1 Ground station ID SHANNON - IRELAND SYNCHED
Aircraft ID AD
Slots Requested medium = 0 Low = 0
Max Bit rate 1200 bps U(R) = 0 UR(R)vect = 0
[LPDU UNNUMBERED DATA]
[HFNPDU PERFORMANCE]
14:43:30  UTC  Flight ID = LH8393  LAT 52 37 27  N  LON 16 46 41  E
07 87 AD 00 03 00 31 C5 0B 0D FF D1 4C 48 38     ......1 .....L H 8
33 39 33 BF 56 62 EE 0B 89 67 01 8A 07 01 B8     3 9 3 .V b ...g .....
00 7E 00 00 00 00 00 00 00 06 0F 00 00 00 00     ..............
2E 00 00 00 00 00 05 00 00 00 05 07 08 27 00     ..............
00 00 00 00 00 00 00             .......
```

Figure 6.2-6 Examples of aircraft communication using HFDL.

between ground stations finding the best one before logging on. When passing traffic, dedicated TDM time slots are used. This prevents two aircraft transmitting at the same time causing **data collisions**.

6.2.5 HF Radio Equipment

The block schematic of a simple HF transmitter/receiver is shown in Figure 6.2-9. Note that, whilst this equipment uses a single intermediate frequency (IF), in practice most modern aircraft HF radios are much more complex and use two or three intermediate frequencies.

Figure 6.2-7 Ground station and aircraft locations for the HFDL communications in Figure 6.2-6.

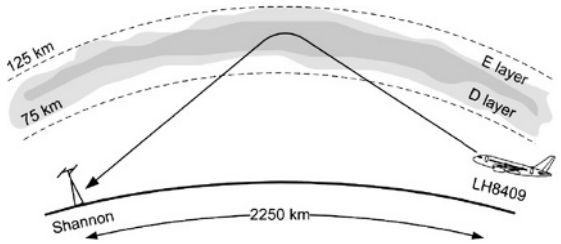


Figure 6.2-8 Radio path for LH8409.

On transmit mode, the DSB suppressed carrier (Figure 6.2-2b) is produced by means of a **balanced modulator** stage. The balanced modulator rejects the carrier and its output just comprises the upper and lower sidebands. The DSB signal is then passed through a multiple-stage crystal or mechanical filter. This filter has a very narrow pass-band (typically 3.4 kHz) at the intermediate frequency (IF) and this rejects the unwanted sideband. The resulting SSB signal is then mixed with a signal from the digital frequency synthesiser to produce a signal on the wanted channel. The output from the mixer is then further amplified before being passed to the output stage. Note that, to avoid distortion, all of the stages must operate in linear mode.

When used on receive mode, the incoming signal frequency is mixed with the output from the digital frequency synthesiser in order to produce the intermediate frequency signal. Unwanted adjacent channel signals are removed by means of another multiple-stage crystal or mechanical filter which has a pass-band similar to that used in the transmitter. The IF signal is then amplified before being passed to the demodulator.

The (missing) carrier is reinserted in the demodulator stage. The carrier signal is derived from an accurate crystal controlled carrier oscillator which operates at the IF frequency. The recovered audio signal from the demodulator is then passed to the audio amplifier where it is amplified to an appropriate level for passing to a loudspeaker.

The typical specification for an aircraft HF radio is shown in Table 6.2-2. One or two radios of this type are usually fitted to a large commercial aircraft (note that at least one HF radio is a requirement for *any* aircraft following a transoceanic route). Figure 6.2-10 shows the flight deck location of the **HF radio controller**.

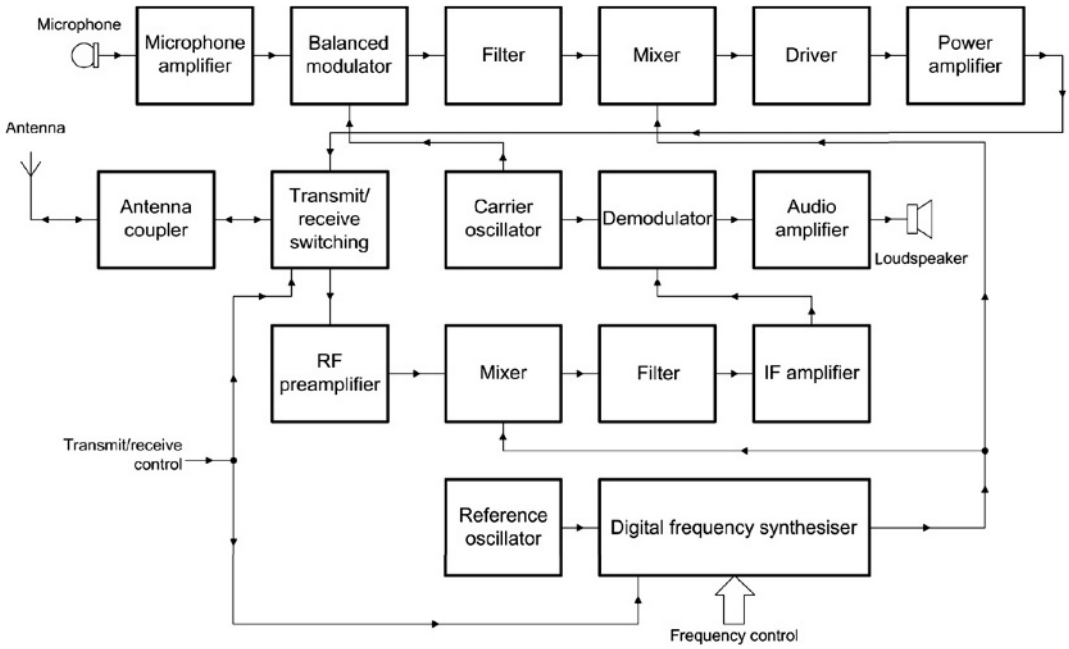


Figure 6.2-9 A simple SSB transmitter/receiver.

Table 6.2-2 Aircraft HF radio specifications.

Parameter	Specification
Frequency range	2.0000 MHz to 29.9999 MHz
Turning steps	100 Hz
Operating modes	SSB SC analogue voice (ARINC 719) and analogue data (ARINC 735 and ARINC 635) at up to 1800 bps; DSB AM (full carrier)
Sensitivity	1 μV for 10 dB (S + N)/N SSB; 4 μV for 10 dB (S + N)/N AM
Selectivity	6 dB max. attenuation at +2.5 kHz 60 dB min. attenuation at +3.4 kHz
Audio output	50 mW into 600 Ω
SELCAL output	50 mW into 600 Ω
RF output power	200 W pep min. SSB; 50 W min. DSB AM
Frequency stability	±20 Hz
Audio response	350 Hz to 2500 Hz at −6 dB
Mean time between failure	Greater than 50,000 hours

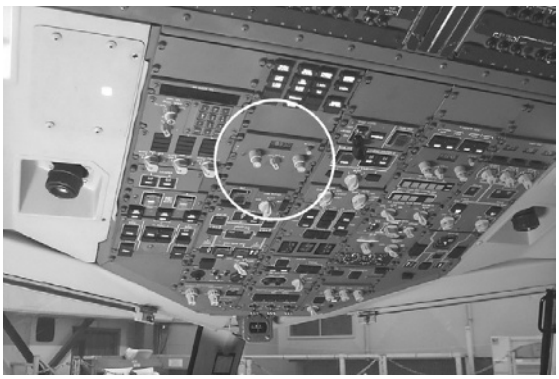


Figure 6.2-10 HF radio control unit.

Test your understanding 6.2.2

1. Explain how HFDL differs from VDL. Under what circumstances is HFDL used and what advantages does it offer?
2. Explain briefly how an aircraft logs on to the HFDL system. How are data collisions avoided?

6.2.6 HF Antennas and Coupling Units

External wire antennas were frequently used on early aircraft. Such antennas would usually run from the fuselage to the top of the vertical stabiliser and they were sufficiently long to permit resonant operation on one or more of the aeronautical HF bands. Unfortunately this type of antenna is unreliable and generally unsuitable for use with a modern high-speed passenger aircraft. The use of a large probe antenna is unattractive due to its susceptibility to static discharge and lightning strike. Hence an alternative solution in which the HF antenna is protected within the airframe is highly desirable. Early experiments (see Figure 6.2-13) showed that the vertical stabiliser (tail fin) would be a suitable location and is now invariably used to house the HF antenna and its associated coupling unit on most large transport aircraft—see Figures 6.2-11 and 6.2-12.

Due to the restriction in available space (which mitigates against the use of a resonant antenna such as a quarter-wave Marconi antenna) the HF antenna is based on a notch which uses part of the airframe in order to radiate effectively. The notch itself has a very high-Q factor and its resistance and reactance varies very widely over the operating frequency range (i.e. 3 MHz to 24 MHz). The typical variation of **standing wave ratio** (SWR) against frequency for an HF notch antenna is shown in Figure 6.2-14. For comparison, the variation of SWR with frequency for a typical quarter-wave VHF blade antenna is shown in Figure 6.2-15.

From Figures 6.2-14 and 6.2-15 it should be obvious that the HF antenna, whilst well matched at 21 MHz, would be severely mismatched to a conventional 50 Ω feeder/transmitter at most other HF frequencies. Because of this, and because the notch antenna is usually voltage fed, it is necessary to use an HF coupling/tuning unit between the HF

Figure 6.2-11 HF antenna location.

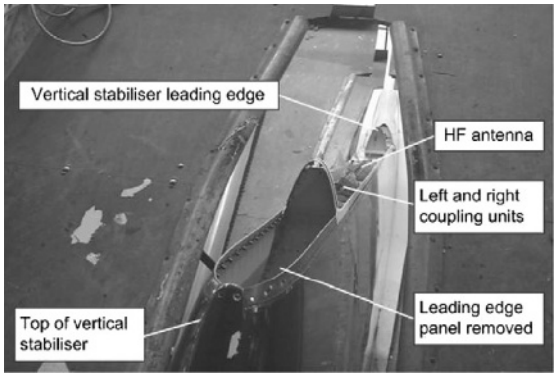


Figure 6.2-12 View from the top of the vertical stabiliser (leading edge panel removed).

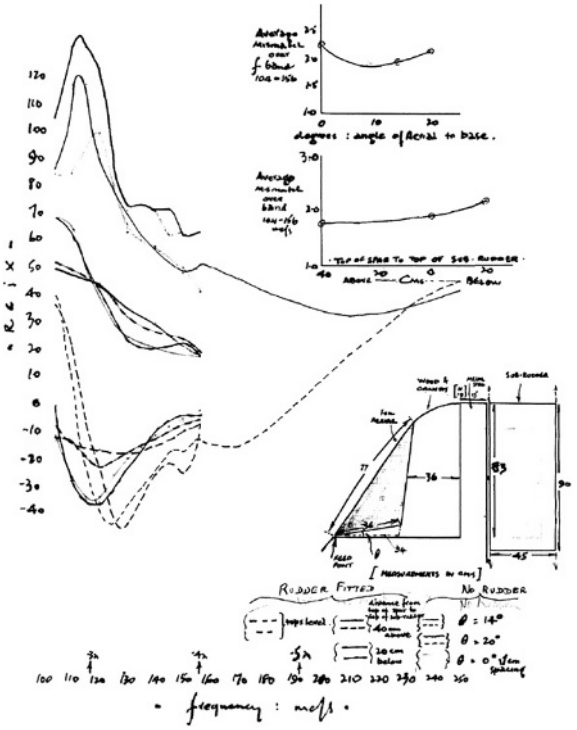


Figure 6.2-13 Original sketches for a tail-mounted antenna from work carried out by E. H. Tooley in 1944.

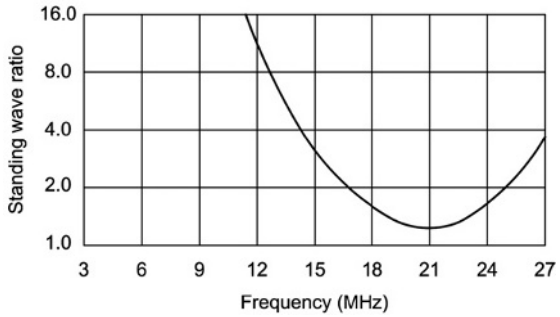


Figure 6.2-14 Variation of SWR with frequency for an HF notch antenna (note the logarithmic scale used for SWR).

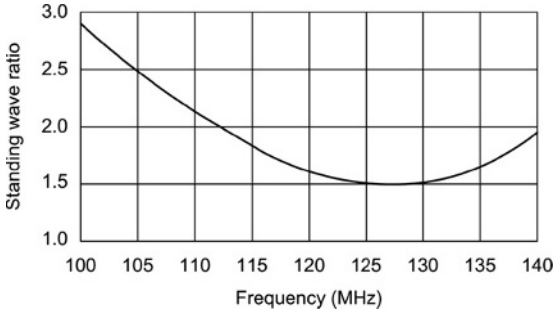


Figure 6.2-15 Variation of SWR with frequency for a VHF quarter-wave blade antenna (note the linear scale used for SWR).

radio feeder and the notch antenna. This unit is mounted in close proximity to the antenna, usually close to the top of the vertical stabiliser (see Figure 6.2-12). Figure 6.2-16 shows the effect of using a coupling/tuning unit on the SWR-frequency characteristic of the same notch antenna that was used in Figure 6.2-14. Note how the SWR has been reduced to less than 2:1 for most (if not all) of the HF range.

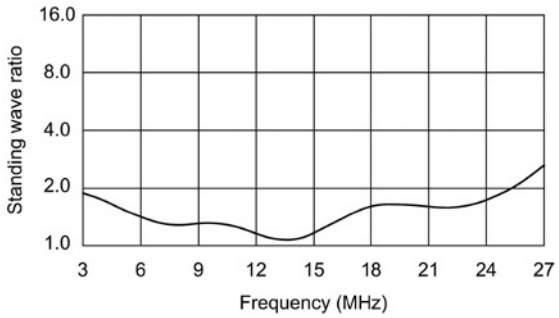


Figure 6.2-16 Variation of SWR with frequency for an HF notch antenna fitted with an antenna coupling/tuning unit.

The tuning adjustment of HF antenna coupler is entirely automatic and only requires a brief signal from the transmitter to retune to a new HF frequency. The HF antenna coupler unit incorporates an SWR bridge and a feedback control system (see Figure 6.2-17) to adjust a roller coater inductor (L1) and high-voltage vacuum variable capacitor (C1) together with a number of switched high-voltage capacitors (C1 to C4). The internal arrangement of a typical HF antenna coupler is shown in Figures 6.2-18 and 6.2-19. The connections required between the HF antenna coupler, HF radio and control unit are shown in Figure 6.2-20.

Voltages present in the vicinity of the HF antenna (as well as the field radiated by it) can be extremely dangerous. It is therefore **essential** to avoid contact with the antenna and to maintain a safe working distance from it (at least 5 metres) whenever the HF radio system is 'live'.

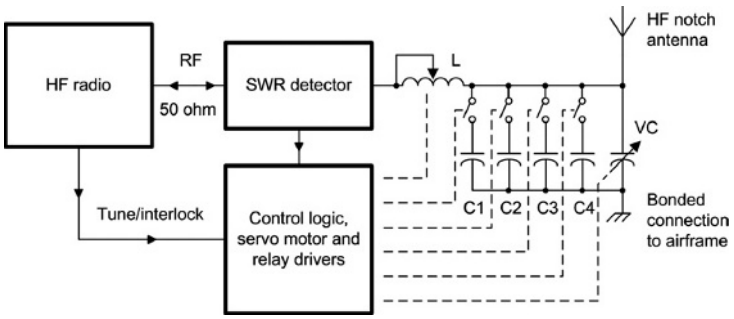


Figure 6.2-17 Typical feedback control system used in an HF antenna coupler.

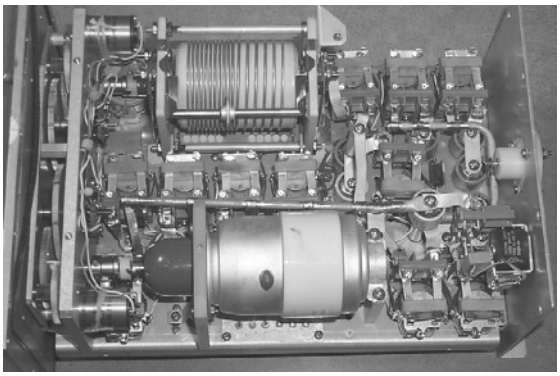


Figure 6.2-18 Interior view of an HF antenna coupler showing the roller coaster inductor (top) and vacuum variable capacitor (bottom). The high-voltage antenna connector is shown in the extreme right.

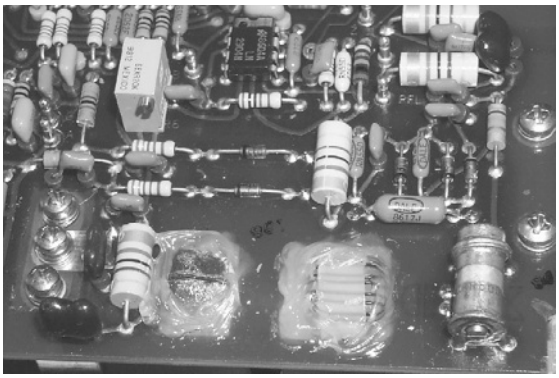


Figure 6.2-19 SWR bridge circuit incorporated in the HF antenna coupler. The output from the SWR bridge provides the error signal input to the automatic feedback control system.

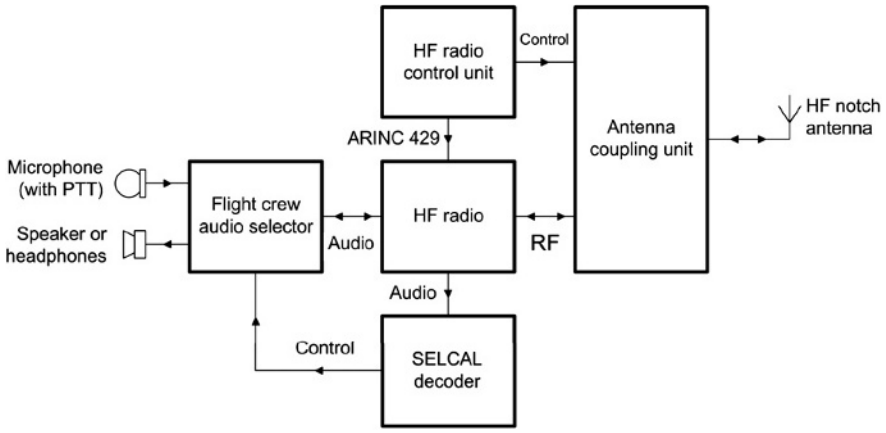


Figure 6.2-20 Connections to the HF radio, control unit and antenna coupling unit.

Test your understanding 6.2.3

Explain the function of an HF antenna coupler. What safety precautions need to be observed when accessing this unit?

6.2.7 Multiple Choice Questions

1. The typical bandwidth of an aircraft HF SSB signal is:
   (a) 3.4 kHz
   (b) 7 kHz
   (c) 25 kHz.

2. The principal advantage of SSB over DSB AM is:
   (a) reduced bandwidth
   (b) improved frequency response
   (c) faster data rates can be supported.

3. HF data link uses typical data rates of:
   (a) 300 bps and 600 bps
   (b) 2400 bps and 4800 bps
   (c) 2400 bps and 31,500 bps.

4. The standard for HF data link is defined in:
   (a) ARINC 429
   (b) ARINC 573
   (c) ARINC 635.

5. Which one of the following gives the approximate range of audio frequencies used for SELCAL tones?
   (a) 256 Hz to 2048 Hz
   (b) 312 Hz to 1479 Hz
   (c) 300 Hz to 3400 Hz.

6. How many alphanumeric characters are transmitted in a SELCAL code?
   (a) 4
   (b) 8
   (c) 16.

7. How many bits are used in an ICAO aircraft address?
   (a) 16
   (b) 24
   (c) 32.

8. The typical RF output power from an aircraft HF transmitter is:
   (a) 25 W pep
   (b) 50 W pep
   (c) 400 W pep.

9. An HF radio is required for use on oceanic routes because:
   (a) VHF coverage is inadequate
   (b) higher power levels can be produced
   (c) HF radio is more reliable.

10. The function of an HF antenna coupler is to:
    (a) reduce static noise and interference
    (b) increase the transmitter output power
    (c) match the antenna to the radio.

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.3 Aircraft Navigation

Mike Tooley and David Wyatt

Navigation is the science of conducting journeys over land and/or sea. Whether the journey is to be made across deserts or oceans, we need to know the ultimate destination and how the journey's progress will be checked along the way. Finding a position on the earth's surface and deciding on the direction of travel can be simply made by observations or by mathematical calculations. Aircraft navigation is no different, except that the speed of travel is much faster! Navigation systems for aircraft have evolved with the nature and role of the aircraft itself. Starting with visual references and the basic compass, leading onto radio ground aids and self-contained systems, many techniques and methods are employed.

Although the basic requirement of a navigation system is to guide the crew from point A to point B, increased traffic density and airline economics means that more than one aircraft is planning a specific route. Flight planning takes into account such things as favourable winds, popular destinations and schedules. Aircraft navigation is therefore also concerned with the management of traffic and safe separation of aircraft. This chapter reviews some basic features of the earth's geometry as it relates to navigation, and introduces some basic aircraft navigation terminology. The chapter concludes by reviewing a range of navigation systems used on modern transport and military aircraft.

6.3.1 The Earth and Navigation

Before looking at the technical aspects of navigation systems, we need to review some basic features of the earth and examine how these features are employed for aircraft navigation purposes. Although we might consider the earth to be a perfect sphere, this is not the case. There is a flattening at both the poles such that the earth is shaped more like an orange. For short distances, this is not significant; however, for long-range (i.e. global) navigation we need to know some accurate facts about the earth. The mathematical definition of a sphere is where the distance (radius) from the centre to the surface is equidistant. This is not the case for the earth, where the actual shape is referred to as an oblate spheroid.

6.3.1.1 Position

To define a unique two-dimensional position on the earth's surface, a coordinate system using imaginary lines of **latitude** and **longitude** is drawn over the globe, see Figure 6.3-1. Lines of longitude join the poles in **great circles** or **meridians**. A great circle is defined as the intersection of a sphere by a plane passing through the centre of the sphere; this has a radius measured from the centre to the surface of the earth. These north–south lines are spaced around the globe and measured in angular distance from the **zero (or prime) meridian**, located in Greenwich, London. Longitude referenced to the prime-meridian extends east or west up to 180 **degrees**. Note that the distance between lines of longitude converge at the poles. Latitude is the angular distance north or south of the equator; the poles are at latitude 90 degrees.

For accurate navigation, the degree (symbol ° after the value, e.g. 90° north) is divided by 60 giving the unit of **minutes** (using the symbol ′ after numbers), e.g. one half of a degree will be 30′. This can be further refined into smaller units by dividing again by 60 giving the unit of **seconds** (using the symbol " after numbers), e.g. one half of a minute will be 30". A second of latitude (or longitude at the equator) is approximately 31 metres, just over 100 feet. Defining a unique position on the earth's surface, e.g. Land's End in Cornwall, UK, using latitude and longitude is written as

Latitude N 50° 04′ 13″ Longitude W 5° 42′ 42″

6.3.1.2 Direction

Direction to an observed point (**bearing**) can be referenced to a known point on the earth's surface, e.g. **magnetic north**. Bearing is defined as the angle from the vertical plane of the reference point through to the vertical plane of the observed point. Basic navigational information is expressed in terms of **compass points** from zero referenced to

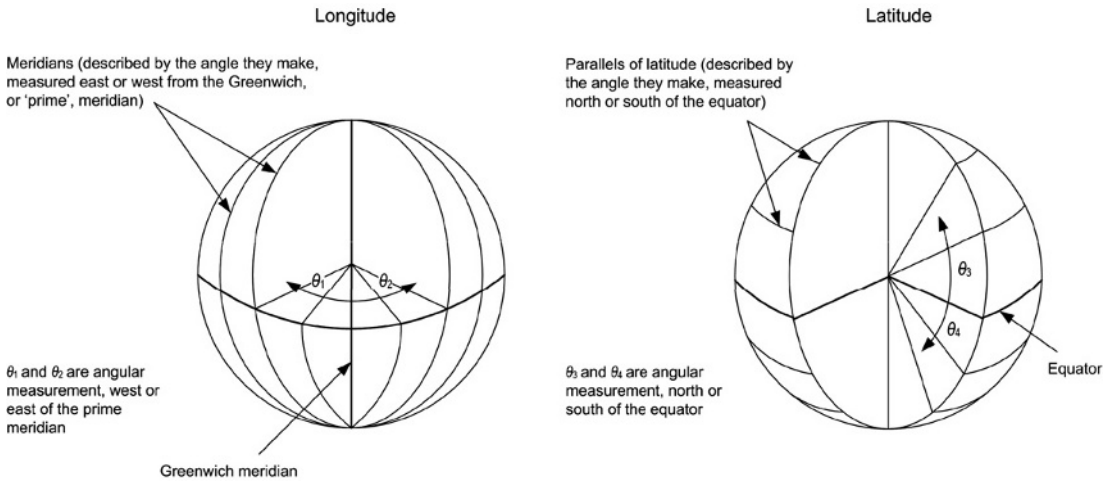


Figure 6.3-1 Longitude and latitude.

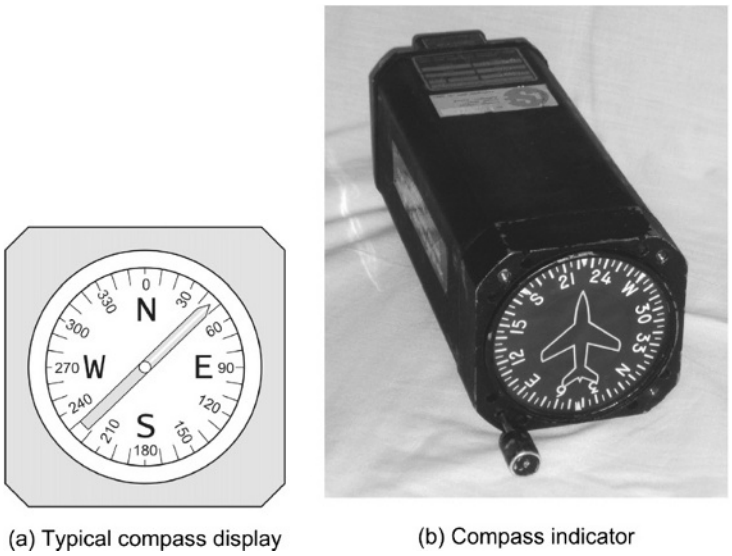


Figure 6.3-2 Compass indications.

north through 360° in a clockwise direction, see Figure 6.3-2. For practical navigation purposes, north has been taken from the natural feature of the earth's magnetic field; however, magnetic north is not at 90° latitude; the latter defines the position of **true north**. The location of magnetic north is in the Canadian Arctic, approximately 83° latitude and 115° longitude west of the prime meridian, see Figure 6.3-3. Magnetic north is a natural feature of the earth's geology; it is slowly drifting across the Canadian Arctic at approximately 40 km northwest per year. Over a long period of time, magnetic north describes an elliptical path. The Geological Survey of Canada keeps track of this motion by periodically carrying out magnetic surveys to redetermine the pole's location. In addition to this long-term change, the earth's magnetic field is also affected on a random basis by the weather, i.e. electrical storms.

Navigation charts based on magnetic north have to be periodically updated to consider this gradual drift. Compass-based systems are referenced to magnetic north; since this is not at 90° latitude there is an angular difference between magnetic and true north. This difference will be zero if the aircraft's position happens to be on the same longitude as magnetic north, and maximum at longitudes ±90° either side of this longitude. The angular difference between magnetic north and true north is called **magnetic variation**. It is vital that when bearings or headings are used, we are clear on what these are referenced to.

Figure 6.3-3 Location of magnetic north.

The imaginary lines of latitude and longitude described above are curved when superimposed on the earth's surface; they also appear as straight lines when viewed from above. The shortest distance between points A and B on a given route is a straight line. When this route is examined, the projection of the path (the **track**) flown by the aircraft over the earth's surface is described by a great circle.

Flying in a straight line implies that we are maintaining a constant heading, but this is not the case. Since the lines of longitude converge, travelling at a constant angle at each meridian yields a track that actually curves as illustrated in Figure 6.3-4. A track that intersects the lines of longitude at a constant angle is referred to as a **rhumb line**. Flying a rhumb line is readily achieved by reference to a fixed point, e.g. magnetic north. The great circle route, however, requires that the direction flown (with respect to the meridians) changes at any given time, a role more suited to a navigation computer.

6.3.1.3 Distance and Speed

The standard unit of measurement for distance used by most countries around the world (the exceptions being the UK and USA) is the kilometre (km). This quantity is linked directly to the earth's geometry; the distance between the poles and equator is 10,000 km. The equatorial radius of the earth is 6378 km; the polar radius is 6359 km.

For aircraft navigation purposes, the quantity of distance used is the **nautical mile (nm)**. This quantity is defined by distance represented by 1 minute of arc of a great circle (assuming the earth to be a perfect sphere). The nautical mile (unlike the statute mile) is therefore directly linked to the geometry of the earth. Aircraft speed, i.e. the rate of change of distance with respect to time, is given by the quantity '**knots**'; nautical miles per hour.

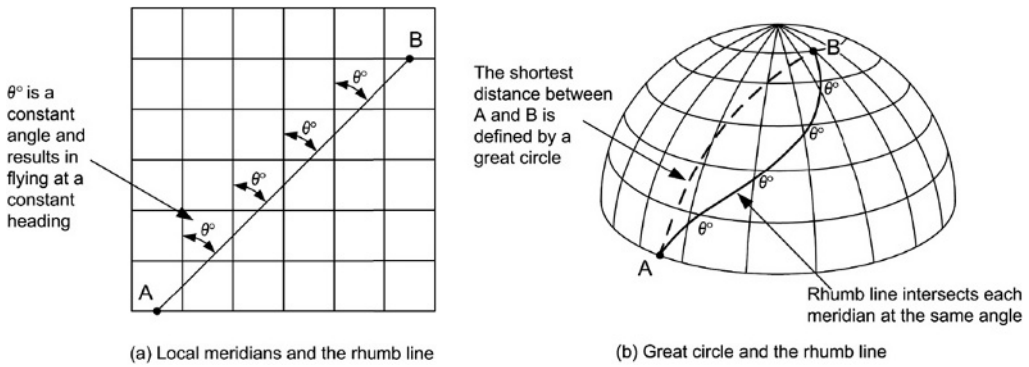


Figure 6.3-4 Flying a constant heading.

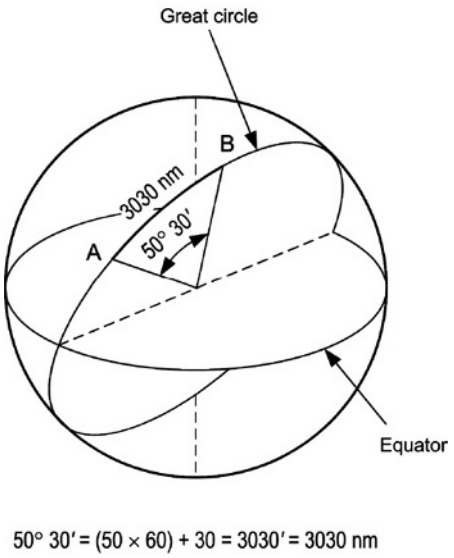


Figure 6.3-5 Calculation of great circle distances.

Calculating the great circle distance between two positions defined by an angle is illustrated in Figure 6.3-5. The distance between two positions defined by their respective latitudes and longitudes, ($lat1$, $lon1$) and ($lat2$, $lon2$), can be calculated from the formula

$$d = \cos^{-1}(\sin(lat1) \times \sin(lat2) + \cos(lat1) \times \cos(lat2) \times \cos(lon1 - lon2))$$

Test your understanding 6.3.1

Explain each of the following terms:

1. Latitude
2. Longitude
3. Great circle
4. Rhumb line.

Key Point

Although we might consider the earth to be a perfect sphere, this is not the case. The actual shape of the earth is referred to as an oblate spheroid.

Key Point

Longitude referenced to the prime-meridian extends east or west up to 180 degrees. Latitude is the angular distance north or south of the equator; the poles are at a latitude of 90 degrees.

Key Point

The nautical mile (unlike the statute mile) is directly linked to the geometry of the earth. This quantity is defined by distance represented by 1 minute of arc of a great circle (assuming the earth to be a perfect sphere).

Key Point

Both latitude and longitude are angular quantities measured in degrees. For accurate navigation, degrees can be divided by 60 giving the unit of ‘minutes’; this can be further divided by 60 giving the unit of ‘seconds’.

6.3.2 Dead Reckoning

Estimating a position by extrapolating from a known position and then keeping note of the direction, speed and elapsed time is known as **dead reckoning**. An aircraft passing over a given point on a heading of 90° at a speed of 300 knots will be 5 miles due east of the given point after 1 minute. If the aircraft is flying in zero wind conditions, this simple calculation holds true. In realistic terms, the aircraft will almost certainly be exposed to wind at some point during the flight and this will affect the navigation calculation. With our aircraft flying on a heading of 90° at a speed of 300 knots, let's assume that the wind is blowing from the south at 10 knots, see Figure 6.3-6. In a 1 hour time period, the air that the aircraft is flying in will have moved north by 10 nautical miles. This means that the aircraft's path (referred to as its **track**) over the earth's surface is not due east. In other words, the aircraft track is not the same as the direction in which the aircraft is heading. This leads to a horizontal displacement (**drift**) of the aircraft from the track it would have followed in zero wind conditions.

The angular difference between the heading and track is referred to as the **drift angle** (quoted as being to port/left or starboard/right of the heading). If the wind direction were in the same direction as the aircraft heading, i.e. a tail wind, the aircraft speed of 300 knots through the air would equate to a ground speed of 310 knots. Likewise, if the wind were from the east (a headwind) the ground speed would be 290 knots.

Knowledge of the wind direction and speed allows the crew to steer the aircraft into the wind such that the wind actually moves the aircraft onto the desired track. For dead reckoning purposes, we can resolve these figures in mathematical terms and determine a position by triangulation as illustrated in Figure 6.3-7. Although the calculation is straightforward, the accuracy of navigation by dead reckoning will depend on up to date knowledge of wind speed and direction. Furthermore, we need accurate measurements of speed and direction. Depending on the accuracy of measuring these parameters, positional error will build up over time when navigating by dead reckoning. We therefore need a means of checking our calculated position on a periodic basis; this is the process of **position fixing**.

Key Point

Dead reckoning is used to estimate a position by extrapolating from a known position and then keeping note of the direction and distance travelled.

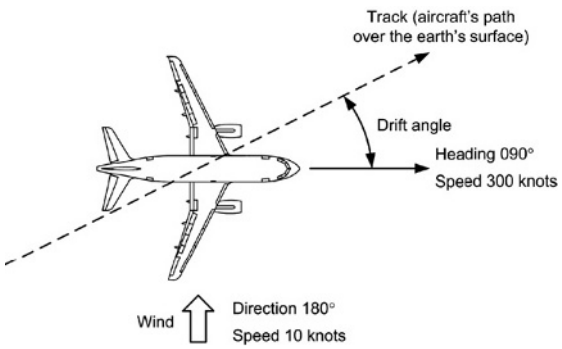


Figure 6.3-6 Effect of crosswind.

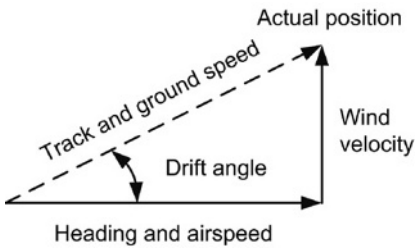


Figure 6.3-7 Resolving actual position.

Key Point

The angular difference between the heading and track is referred to as the drift angle.

6.3.3 Position Fixing

When travelling short distances over land, natural terrestrial features such as rivers, valleys, hills etc. can be used as direct observations to keep a check on (**pinpointing**) the journey's progress. If the journey is by sea, then we can use the coastline and specific features such as lighthouses to confirm our position. If the journey is now made at night or out of sight of the coast, we need other means of fixing our position.

The early navigators used the sun, stars and planets very effectively for navigation purposes; if the position of these celestial objects is known, then the navigator can confirm a position anywhere on the earth's surface. **Celestial navigation** (or **astronavigation**) was used very effectively in the early days of long distance aircraft navigation. Indeed, it has a number of distinct advantages when used by the military: the aircraft does not radiate any signals; navigation is independent of ground equipment; the references cannot be jammed; navigation references are available over the entire globe.

The disadvantage of celestial navigation for aircraft is that the skies are not always clear and it requires a great deal of skill and knowledge to be able to fix a position whilst travelling at high speed. Although automated celestial navigation systems were developed for use by the military, they are expensive; modern avionic equipment has now phased out the use of celestial navigation for commercial aircraft.

The earliest ground-based references (**navigation aids**) developed for aircraft navigation are based on radio beacons. These beacons can provide angular and/or distance information; when using this information to calculate a position fix, the terms are referred to mathematically as theta (θ) and rho (ρ). By utilising the directional properties of radio waves, the intersection of signals from two or more navigation aids can be used to fix a position (**theta–theta**), see Figure 6.3-8. Alternatively, if we know the distance and direction (bearing) to a navigation aid, the aircraft position can be confirmed (**rho–theta**). Finally, we can establish our position if we know the aircraft's distance (**rho–rho**) from any two navigation aids, i.e. without knowledge of the bearing.

6.3.4 Maps and Charts

Maps provide the navigator with a representative diagram of an area that contains a variety of physical features, e.g. cities, roads and topographical information. Charts contain lines of latitude and longitude, together with essential data such as the location of navigation aids. Creating charts and maps requires that we transfer distances and geographic features from the earth's spherical surface onto a flat piece of paper. This is not possible without some kind of compromise in geographical shape, surface area, distance or direction. Many methods of producing charts have been developed over the centuries; the choice of projection depends on the intended purpose.

In the sixteenth century Gerhardus Mercator, the Flemish mathematician, geographer and cartographer, developed what was to become the standard chart format for nautical navigation: the **Mercator projection**. This is a cylindrical map projection where the lines of latitude and longitude are projected from the earth's centre, see Figure 6.3-9. Imagine a cylinder of paper wrapped around the globe and a light inside the globe; this projects the lines of latitude and longitude onto the paper. When the cylinder is unwrapped, the lines of latitude appear incorrectly as

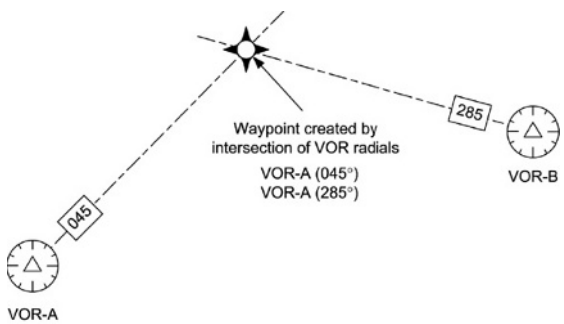


Figure 6.3-8 Position fixing.

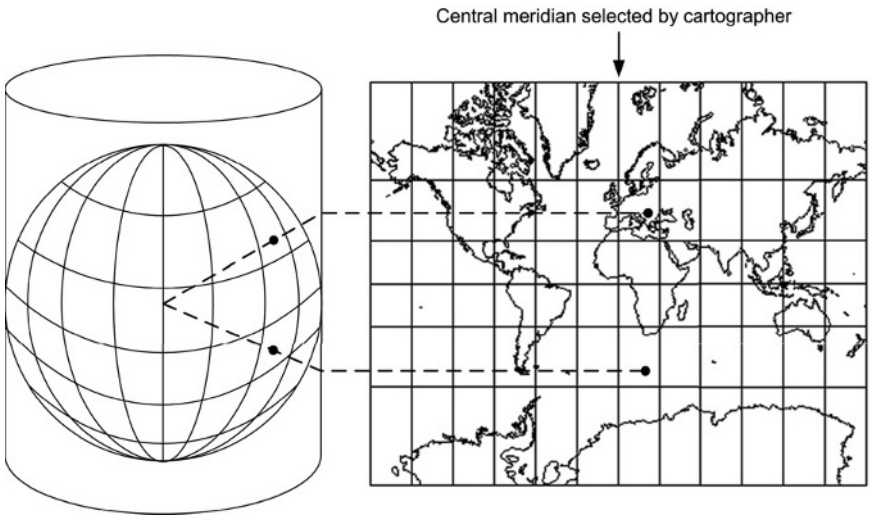


Figure 6.3-9 Mercator projection.

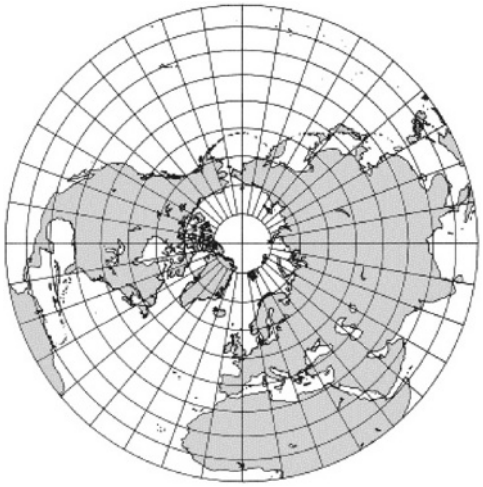


Figure 6.3-10 Lambert projection (viewed from true north).

having equal length. Directions and the shape of geographic features remain true; however, distances and sizes become distorted. The advantage of using this type of chart is that the navigator sets a constant heading to reach the destination. The meridians of the Mercator projected chart are crossed at the same angle; the track followed is referred to as a **rhumb line** (see Figure 6.3-4).

For aircraft navigation the Mercator projection might be satisfactory; however, if we want to navigate by great circle routes then we need true directions. An alternative projection used for aircraft navigation, and most popular maps and charts, is the **Lambert** azimuthul equal-area projection. This projection was developed by Johann Heinrich Lambert (1728–77) and is particularly useful in high latitudes. The projection is developed from the centre point of the geographic feature to be surveyed and charted. Using true north as an example, Figure 6.3-10 illustrates the Lambert projection.

6.3.5 Navigation Terminology

The terms shown in Table 6.3-1 are used with numerous navigation systems including INS and RNAV; computed values are displayed on a control display unit (CDU) and/or primary flight instruments. These terms are illustrated in Figure 6.3-11, all terms are referenced to true north.

Table 6.3-1 Navigation terminology.

Term	Abbreviation	Description
Cross track distance	XTK	Shortest distance between the present position and desired track
Desired track angle	DSRTK	Angle between north and the intended flight path of the aircraft
Distance	DIS	Great circle distance to the next waypoint or destination
Drift angle	DA	Angle between the aircraft's heading and ground track
Ground track angle	TK	Angle between north and the flight path of the aircraft
Heading	HDG	Horizontal angle measured clockwise between the aircraft's centreline (longitudinal axis) and a specified reference
Present position	POS	Latitude and longitude of the aircraft's position
Track angle error	TKE	Angle between the actual track and desired track (equates to the desired track angle minus the ground track angle)
Wind direction	WD	Angle between north and the wind vector
True airspeed	TAS	True airspeed measured in knots
Wind speed	WS	Measured in knots
Ground speed	GS	Measured in knots

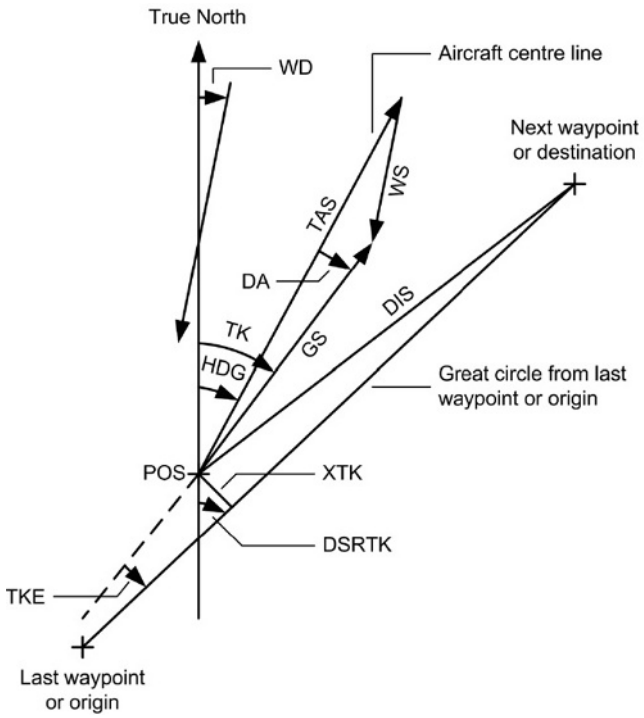


Figure 6.3-11 Navigation terminology.

6.3.6 Navigation Systems Development

This section provides a brief overview of the development of increasingly sophisticated navigation systems used on aircraft.

6.3.6.1 Gyro-Magnetic Compass

The early aviators used visual aids to guide them along their route; these visual aids would have included rivers, roads, rail tracks, coastlines, etc. This type of navigation is not possible at high altitudes or in low visibility and so the earth's magnetic field was used as a reference leading to the use of simple **magnetic compasses** in aircraft. We have seen that magnetic variation has to be taken into account for navigation; there are additional considerations to be addressed for compasses in aircraft. The earth's magnetic field around the aircraft will be affected by:

- the aircraft's own 'local' magnetic fields, e.g. those caused by electrical equipment
- sections of the aircraft with high permeability causing the field to be distorted.

Magnetic compasses are also unreliable in the short-term, i.e. during turning manoeuvres. **Directional gyroscopes** are reliable for azimuth guidance in the short term, but drift over longer time periods. A combined magnetic compass stabilised by a directional gyroscope (referred to as a **gyro-magnetic compass**) can overcome these deficiencies. The gyro-magnetic compass (see Figure 6.3-12), together with an airspeed indicator, allowed the crew to navigate by dead reckoning, i.e. estimating their position by extrapolating from a known position and then keeping note of the direction and distance travelled.

In addition to directional references, aircraft also need an attitude reference for navigation, typically from a vertical gyroscope. Advances in sensor technology and digital electronics have led to combined attitude and heading reference systems (AHRS) based on laser gyros and micro-electromechanical sensors.

Instrumentation errors inevitably lead to deviations between the aircraft's actual and calculated positions; these deviations accumulate over time. Crews therefore need to be able to confirm and update their position by means of a fixed ground-based reference, e.g. a radio navigation aid.

6.3.6.2 Radio Navigation

Early airborne navigation systems using ground-based navigation aids consisted of a loop antenna in the aircraft tuned to amplitude modulated (AM) commercial radio broadcast stations transmitting in the low-/medium-frequency (LF/MF) bands. Referring to Figure 6.3-13, pilots would know the location of the radio station (indeed, it would invariably have been located close to or even in the town/city that the crew wanted to fly to) and this provided a means of fixing a position. Although technology has moved on, these **automatic direction finder** (ADF) systems are still in use today.

Operational problems are encountered using low-frequency (LF) and medium-frequency (MF) transmissions. During the mid to late 1940s, it was evident to the aviation world that an accurate and reliable short-range navigation system was needed. Since radio communication systems based on very high frequency (VHF) were being successfully deployed, a decision was made to develop a radio navigation system based on VHF. This system became the

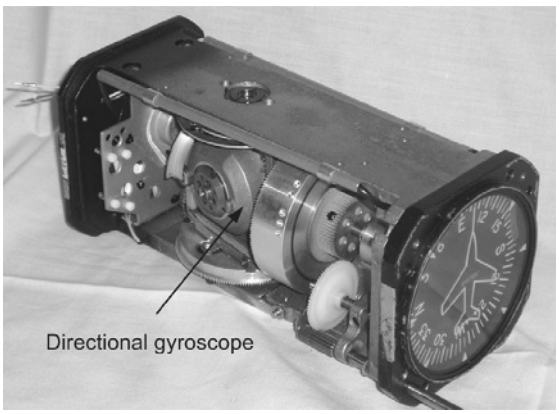


Figure 6.3-12 Gyro-magnetic compass.

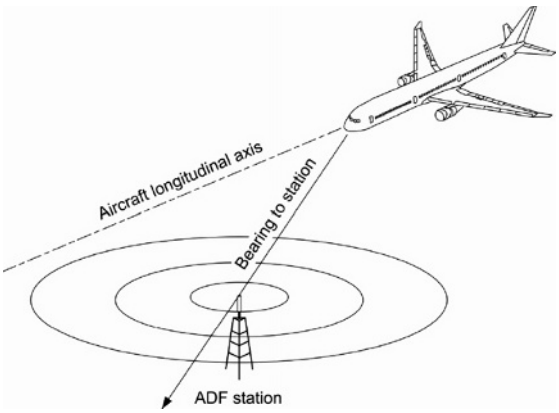


Figure 6.3-13 ADF radio navigation.

VHF omnidirectional range (VOR) system, see Figure 6.3-14; a system that is in widespread use throughout the world today. VOR is the basis of the current network of 'airways' that are used in navigation charts.

The advent of radar in the 1940s led to the development of a number of navigation aids including **distance measuring equipment** (DME). This is a short-/medium-range navigation system, often used in conjunction with the VOR system to provide accurate navigation fixes. The system is based on secondary radar principles, see Figure 6.3-15.

Navigation aids such as ADF, VOR and DME are used to define airways for en route navigation, see Figure 6.3-16. They are also installed at airfields to assist with approaches to those airfields. These navigation aids cannot, however, be used for precision approaches and landings. The standard approach and landing system installed at airfields around the world is the **instrument landing system** (ILS), see Figure 6.3-17. The ILS uses a combination of VHF and UHF radio waves and has been in operation since 1946. There are a number of shortcomings with ILS; in 1978 the **microwave landing system** (MLS) was adopted as the long-term replacement. The system is based on the principle of time referenced scanning beams and provides precision navigation guidance for approach and landing. MLS provides three-dimensional approach guidance, i.e. azimuth, elevation and range. The system provides multiple approach angles for both azimuth and elevation guidance. Despite the advantages of MLS, it has not yet been introduced on a worldwide basis for commercial aircraft. Military operators of MLS often use mobile equipment that can be deployed within hours.

The aforementioned radio navigation aids have one disadvantage in that they are land based and only extend out beyond coastal regions. Long-range radio navigation systems based on **hyperbolic navigation** were introduced in the

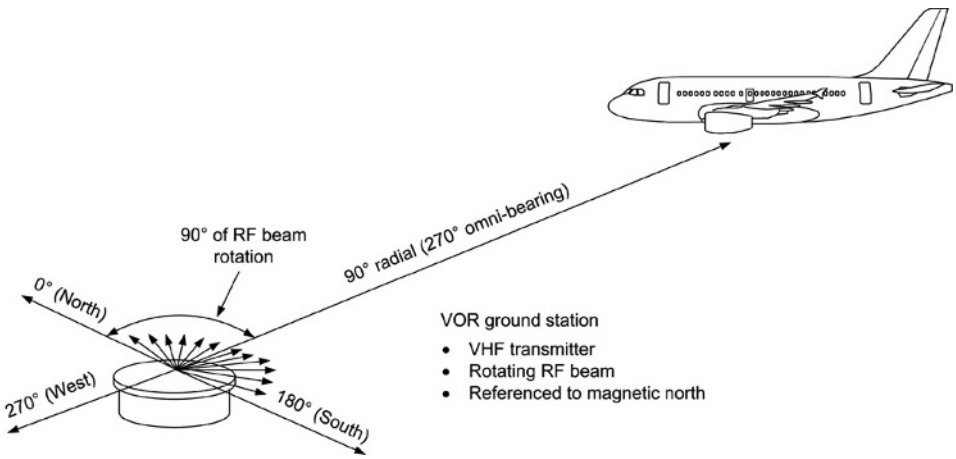


Figure 6.3-14 VOR radio navigation.

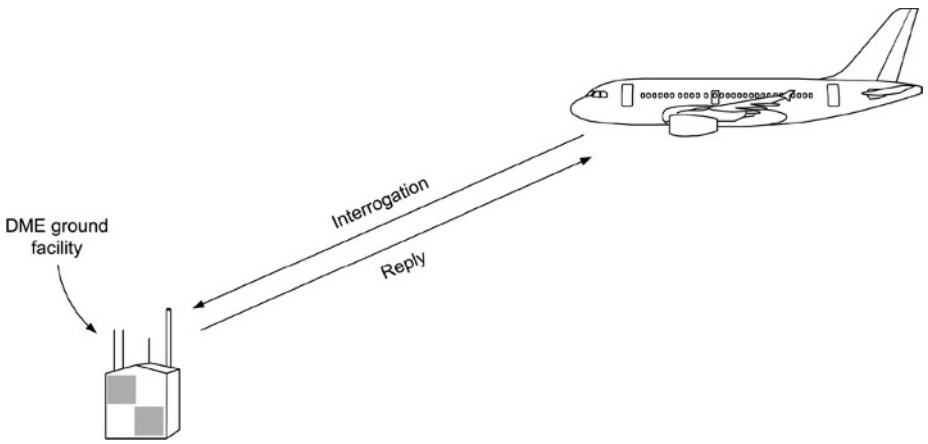


Figure 6.3-15 Distance measuring equipment (DME).

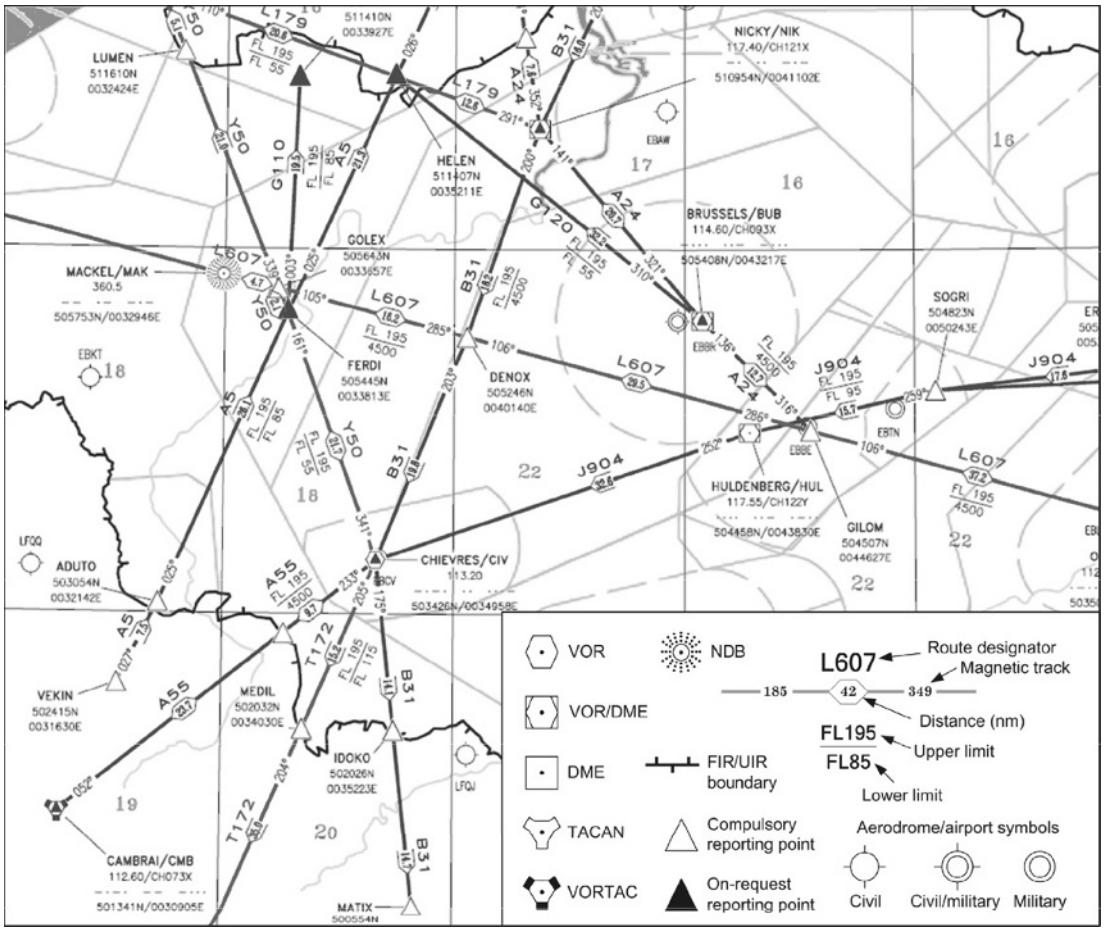


Figure 6.3-16 Airways defined by navigation aids.

1940s to provide for en route operations over oceans and unpopulated areas. Several hyperbolic systems have been developed since, including Decca, Omega and Loran. The operational use of Omega and Decca navigation systems ceased in 1997 and 2000 respectively. Loran systems are still very much available today as stand-alone systems; they are also being proposed as a complementary navigation aid for global navigation satellite systems. The **Loran-C** system is based on a master station and a number of secondary stations; the use of VLF radio provides an increased area of coverage, see Figure 6.3-18.

The advent of computers, in particular the increasing capabilities of integrated circuits using digital techniques, has led to a number of advances in aircraft navigation. One example of this is the **area navigation system** (RNAV); this is a means of combining, or filtering, inputs from one or more navigation sensors and defining positions that are not necessarily co-located with ground-based navigation aids. Typical navigation sensor inputs to an RNAV system can be from external ground-based navigation aids such as VOR and DME, see Figure 6.3-19.

6.3.6.3 Dead Reckoning Systems

Dead reckoning systems require no external inputs or references from ground stations. Doppler navigation systems were developed in the mid-1940s and introduced in the mid-1950s as a primary navigation system. Ground speed and drift can be determined using a fundamental scientific principle called **Doppler shift**. Being self-contained, the system can be used for long distance navigation over oceans and undeveloped areas of the globe.

A major advance in aircraft navigation came with the introduction of the **inertial navigation system** (INS). This is an autonomous dead reckoning system, i.e. it requires no external inputs or references from ground stations. The system was developed in the 1950s for use by the US military and subsequently the space programmes. INSs were introduced into commercial aircraft service during the early 1970s. The system is able to compute navigation data such as present position, distance to waypoint, heading, ground speed, wind speed, wind direction etc. It does not

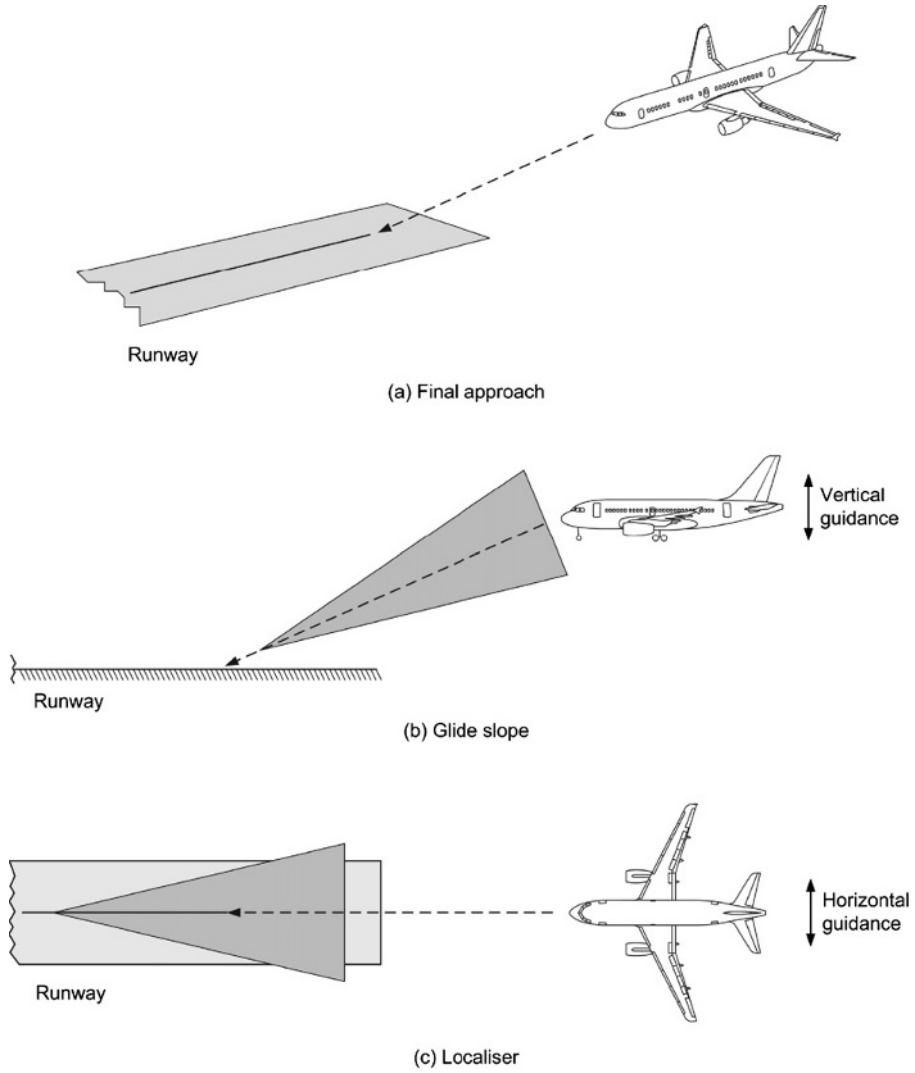


Figure 6.3-17 Instrument landing system.

need radio navigation inputs and it does not transmit radio frequencies. Being self-contained, the system can be used for long distance navigation over oceans and undeveloped areas of the globe.

6.3.6.4 Satellite Navigation

Navigation by reference to the stars and planets has been employed since ancient times; commercial aircraft used to have periscopes to take celestial fixes for long distance navigation. An artificial constellation of navigation aids was initiated in 1973 and referred to as **Navstar** (navigation system with timing and ranging). The global positioning system (GPS) was developed for use by the US military; the first satellite was launched in 1978 and the full constellation was in place and operating by 1994. GPS is now widely available for use by many applications including aircraft navigation; the system calculates the aircraft position by triangulating the distances from a number of satellites, see Figure 6.3-20.

6.3.6.5 Radar Navigation

The planned journey from A to B could be affected by adverse weather conditions. Radar was introduced onto passenger aircraft during the 1950s to allow pilots to identify weather conditions (see Figure 6.3-21) and subsequently reroute around these conditions for the safety and comfort of passengers. A secondary use of weather radar is a terrain-mapping mode that allows the pilot to identify features on the ground, e.g. rivers, coastlines and mountains.

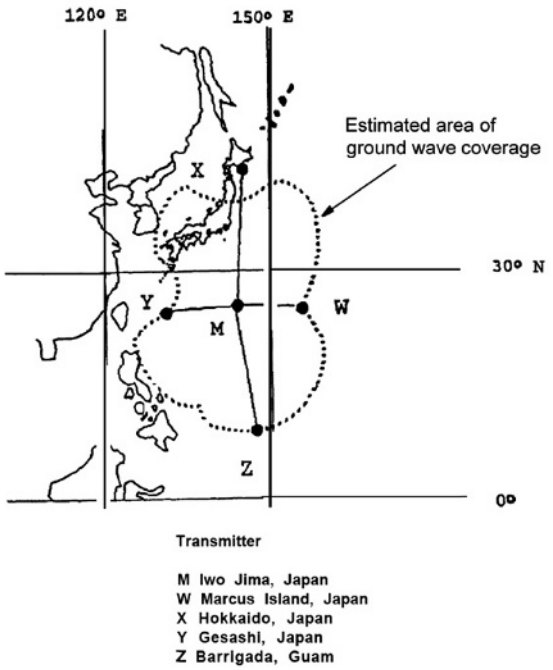


Figure 6.3-18 Loran-C oceanic coverage using VLF transmissions.

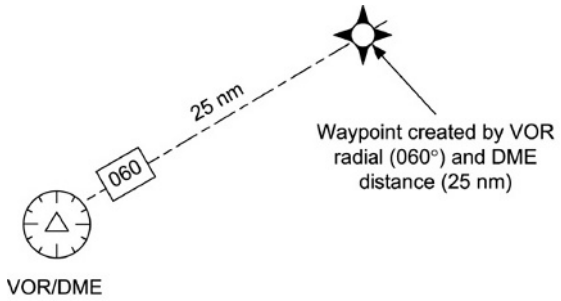


Figure 6.3-19 Area navigation.

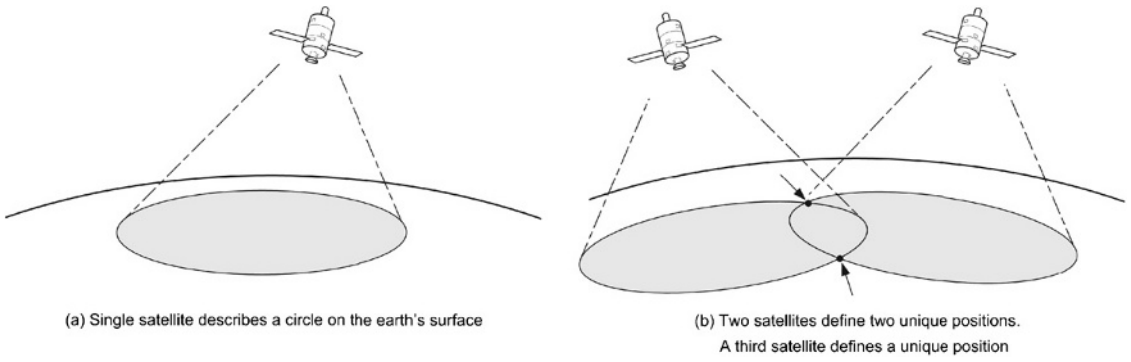


Figure 6.3-20 Satellite navigation.

6.3.6.6 Air Traffic Control

Increasing traffic density, in particular around airports, means that we need a method of **air traffic control** (ATC) to manage the flow of traffic and maintain safe separation of aircraft. The ATC system is based on secondary surveillance radar (SSR) facilities located at strategic sites, at or near airfields. Ground controllers use the SSR system to identify individual aircraft on their screens, see Figure 6.3-22.

With ever increasing air traffic congestion, and the subsequent demands on ATC resources, the risk of a mid-air collision increases. The need for improved traffic flow led to the introduction of the **traffic alert and collision avoidance system** (TCAS). This is an automatic surveillance system that helps aircrews and ATC to maintain safe

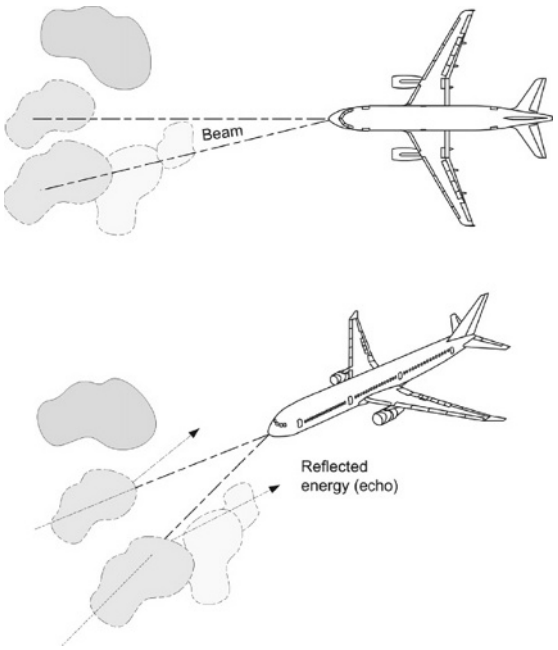


Figure 6.3-21 Weather radar.

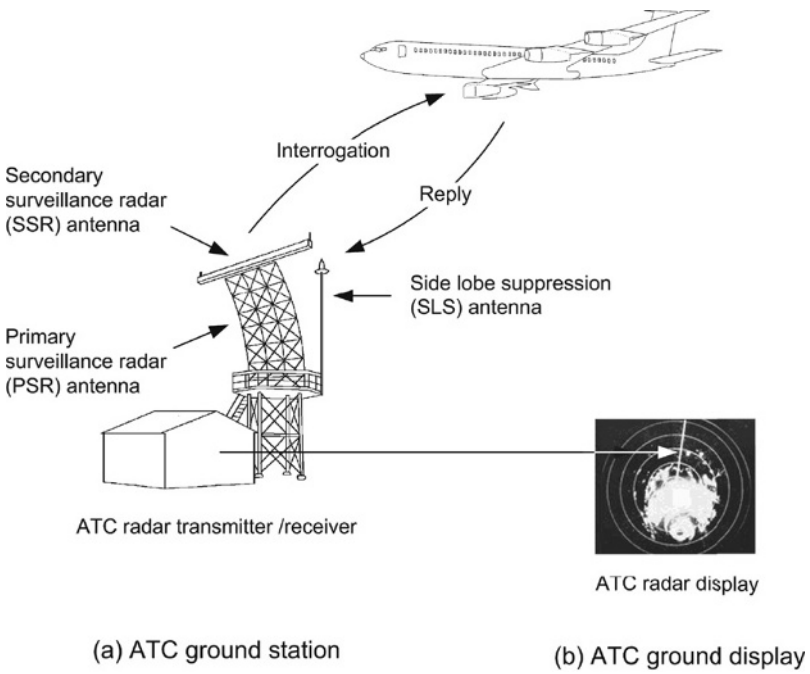


Figure 6.3-22 Secondary surveillance radar.

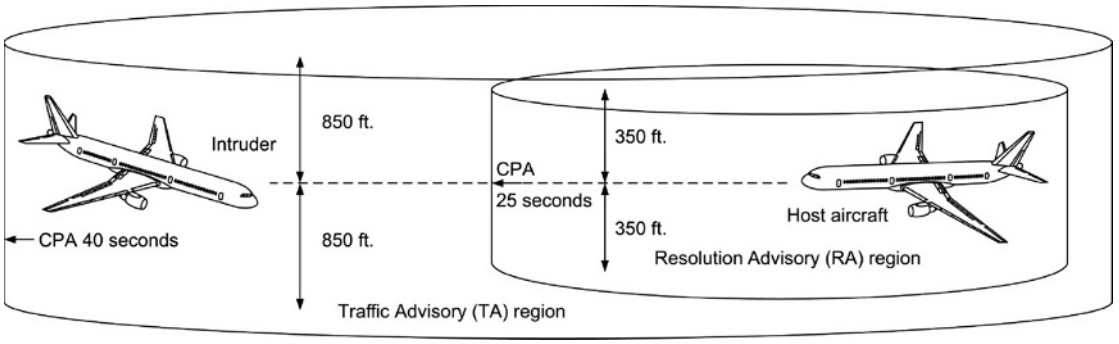


Figure 6.3-23 Traffic alert and collision avoidance system.

separation of aircraft. It is an airborne system (see Figure 6.3-23) based on secondary radar that interrogates and replies directly with aircraft via a high-integrity data link. The system is functionally independent of ground stations and alerts the crew if another aircraft comes within a predetermined time to a potential collision. It is important to note that TCAS is a back-up system, i.e. it provides warnings when other navigation systems (including ATC) have failed to maintain safe separation of aircraft that could lead to a collision.

6.3.7 Navigation Systems Summary

Navigation systems for aircraft have evolved with the nature and role of the aircraft itself. Some of these individual systems are described in detail in the following chapters. Each system has been developed to meet specific requirements within the available technology and cost boundaries. Whatever the requirement, all navigation systems are concerned with several key factors:

- **Accuracy**: conformance between calculated and actual position of the aircraft
- **Integrity**: ability of a system to provide timely warnings of system degradation
- **Availability**: ability of a system to provide the required function and performance
- **Continuity**: probability that the system will be available to the user
- **Coverage**: geographic area where each of the above are satisfied at the same time.

Test your understanding 6.3.2

The nautical mile is directly linked to the geometry of the earth; how is a nautical mile defined?

Test your understanding 6.3.3

Explain the difference between dead reckoning and position fixing.

Test your understanding 6.3.4

For a given airspeed, explain how tailwinds and headwinds affect groundspeed.

Test your understanding 6.3.5

Explain the following terms: accuracy, integrity, availability.

Test your understanding 6.3.6

Describe three ways that bearings and ranges can be used for position fixing.

Test your understanding 6.3.7

1. Explain the difference between Mercator and Lambert projections.
2. Where on the earth's surface is the difference between a rhumb line and great circle route the greatest?

6.3.8 Multiple Choice Questions

1. Longitude referenced to the prime meridian extends:
   (a) north or south up to 180°
   (b) east or west up to 180°
   (c) east or west up to 90°.

2. Latitude is the angular distance:
   (a) north or south of the equator
   (b) east or west of the prime meridian
   (c) north or south of the prime meridian.

3. The distance between lines of longitude converge at the:
   (a) poles
   (b) equator
   (c) great circle.

4. Lines of latitude are always:
   (a) converging
   (b) parallel
   (c) the same length

5. Degrees of latitude can be divided by 60 giving the unit of:
   (a) longitude
   (b) minutes
   (c) seconds.

6. The location of magnetic north is approximately:
   (a) 80° latitude and 110° longitude, east of the prime meridian
   (b) 80° longitude and 110° latitude, west of the prime meridian
   (c) 80° latitude and 110° longitude, west of the prime meridian.

7. One minute of arc of a great circle defines a:
   (a) nautical mile
   (b) kilometre
   (c) knot.

8. The angular difference between magnetic north and true north is called the:
   (a) magnetic variation
   (b) great circle
   (c) prime meridian.

9. Mercator projections produce parallel lines of:
   (a) the earth's magnetic field
   (b) longitude
   (c) great circle routes.

10. With respect to the polar radius, the equatorial radius of the earth is:
    (a) equal
    (b) larger
    (c) smaller.

11. Dead reckoning is the process of:
    (a) fixing the aircraft's position
    (b) correcting the aircraft's position
    (c) estimating the aircraft's position.

12. The angle between the aircraft's heading and ground track is known as the:
    (a) drift angle
    (b) cross track distance
    (c) wind vector.

13. Magnetic compasses are unreliable in the:
    (a) long-term, flying a constant heading
    (b) short-term, during turning manoeuvres
    (c) equatorial regions.

14. The angle between north and the flight path of the aircraft is the:
    (a) ground track angle
    (b) drift angle
    (c) heading.

15. When turning into a 25 knot head wind at constant indicated airspeed, the ground speed will:
    (a) increase by 25 knots
    (b) remain the same
    (c) decrease by 25 knots.

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.4 Automatic Direction Finder

Mike Tooley and David Wyatt

Radio waves have directional characteristics. This is the basis of the automatic direction finder (ADF), one of earliest forms of radio navigation that is still in use today. ADF is a short-/medium-range (200 nm) navigation system providing directional information; it operates within the frequency range 190–1750 kHz, i.e. low and medium frequency bands. The term 'automatic' is somewhat misleading in today's terms; this refers to the introduction of electromechanical equipment to replace manually operated devices. In this chapter we will look at the historical background to radio navigation, review some typical ADF hardware that is fitted to modern commercial transport aircraft, and conclude with some practical aspects associated with the operational use of ADF.

6.4.1 Introducing ADF

The early aviators used visual aids to guide them along their route; these visual aids would have included rivers, roads, rail tracks, coastlines, etc. This type of navigation is not possible in low visibility and so magnetic compasses were introduced. Magnetic compasses are somewhat unreliable in the short term, i.e. during turning manoeuvres. Directional gyroscopes are reliable in the short term, but drift over longer time periods. A combined magnetic compass stabilised by a directional gyroscope (referred to as a **gyromagnetic compass**) can overcome these deficiencies. The gyro-magnetic compass, together with an airspeed indicator, allowed the crew to navigate by dead reckoning, i.e. estimating their position by extrapolating from a known position and then keeping note of the direction and distance travelled. Instrumentation errors inevitably lead to deviations between the aircraft's actual and calculated positions; these deviations accumulate over time. Crew therefore need to be able to confirm and update their position by means of a fixed ground-based reference.

The early airborne navigation systems using ground-based navigation aids consisted of a fixed-loop antenna in the aircraft tuned to an amplitude modulated (AM) commercial radio broadcast station. Pilots would know the location of the radio station (indeed, it would invariably have been located close to or even in the town/city that the crew wanted to fly to). The fixed-**loop antenna** was aligned with the longitudinal axis of the aircraft, with the pilot turning the aircraft until he received the minimum signal strength (null reading). By maintaining a null reading, the pilot could be sure that he was flying towards the station. This constant turning was inefficient in terms of fuel consumption and caused inherent navigation problems in keeping note of the aircraft's position during these manoeuvres! The effects of crosswind complicated this process since the aircraft's heading is not aligned with its track.

6.4.2 ADF Principles

The introduction of an ADF system addresses this problem. A loop antenna that the pilot could rotate by hand solves some of these problems; however, this still requires close attention from the crew. Later developments of the equipment used an electrical motor to rotate the loop antenna. The received signal strength is a function of the angular position of the loop with respect to the aircraft heading and bearing to the station, see Figure 6.4-1(a) and (b). If a plot is made of loop angle and signal strength, the result is a sine wave as shown in Figure 6.4-1(c). The **null point** is easier to determine than the maximum signal strength since the rate of change is highest. Rotating the antenna (rather than turning the aircraft) to determine the null reading from the radio station was a major advantage of the system. The pilot read the angular difference between the aircraft's heading and the direction of the radio station, see Figure 6.4-2(a), from a graduated scale and a bearing to the station could then be determined. The industry drive towards solid-state components, i.e. with no moving parts, has led to the equipment described in Section 6.4.3.

Navigation based on ADF (using AM commercial radio stations broadcasting in the frequency range 540–1620 kHz) became an established method of travelling across country. With the growth of air travel, dedicated radio navigation aids were installed along popular air transport routes. These radio stations, known as **non-directional beacons** (NDBs), gradually supplemented the commercial radio stations and a network of NDBs sprang up in the nations developing their aeronautical infrastructure. These NDBs broadcast in the low-frequency (LF) range 190–415 kHz and medium-frequency (MF) range 510–535 kHz. As the quantity of NDBs increased, air navigation charts

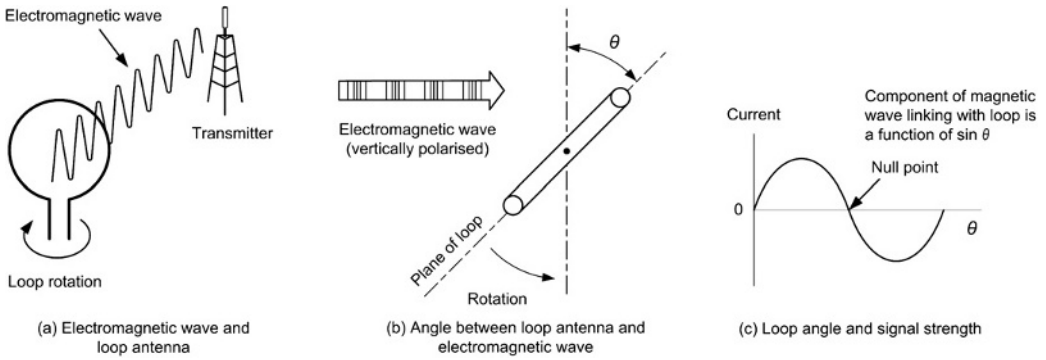


Figure 6.4-1 Loop antenna output.

were produced and the NDBs were identified by a two- or three-letter alpha code linked to the location and frequency. In Figure 6.4-2(c), the NDB located at Mackel in Belgium transmits on 360.5 kHz and is identified as MAK; note the Morse code, latitude and longitude details on the chart. Beacons are deployed with varying power outputs classified as high (2 kW), medium (50 W to 2 kW) and low (less than 50 W).

Table 6.4-1 provides a list of typical NDBs associated with airports and cities in a typical European country (note that these are provided for illustration purposes only). Beacons marked with an asterisk in this table are referred to as locator beacons; they are part of the final approach procedures for an airfield.

6.4.3 ADF Equipment

6.4.3.1 Antenna

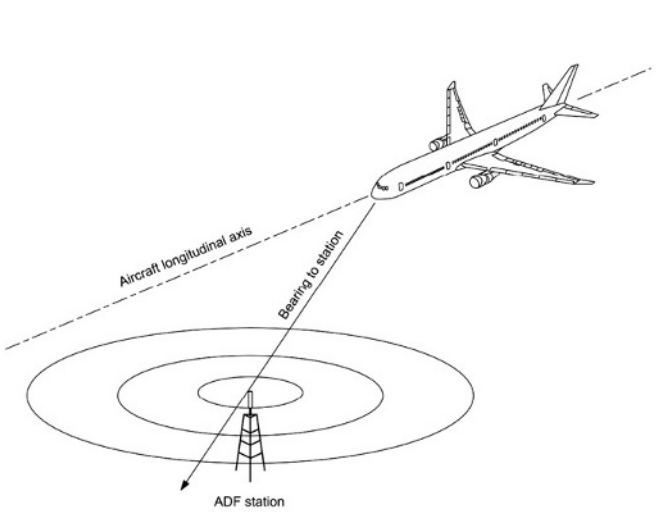
The rotating loop antenna was eventually replaced with a fixed antenna consisting of two loops combined into a single item: one aligned with the centreline of the fuselage the other at right angles as shown in Figure 6.4-3. This orthogonal antenna is still referred to as the 'loop' antenna. Measuring the signal strength from each of the loops, and deriving an angular position in a dedicated ADF receiver, determines the direction to the selected beacon (or commercial radio station). The loop antenna resolves the directional signal; however, this can have two possible solutions 180 degrees apart. A second **'sense' antenna** is therefore required to detect nondirectional radio waves from the beacon; this signal is combined with the directional signals from the loop antenna to produce a single directional solution. The polar diagram for a loop and sense antenna is shown in Figure 6.4-4; when the two patterns are combined, it forms a **cardioid**. Most commercial transport aircraft are fitted with two independent ADF systems typically identified as left and right systems; the antenna locations for a typical transport aircraft are shown in Figure 6.4-5.

6.4.3.2 Receiver

ADF receivers are located in the avionic equipment bay. The signal received at the antenna is coupled to the receiver in three ways:

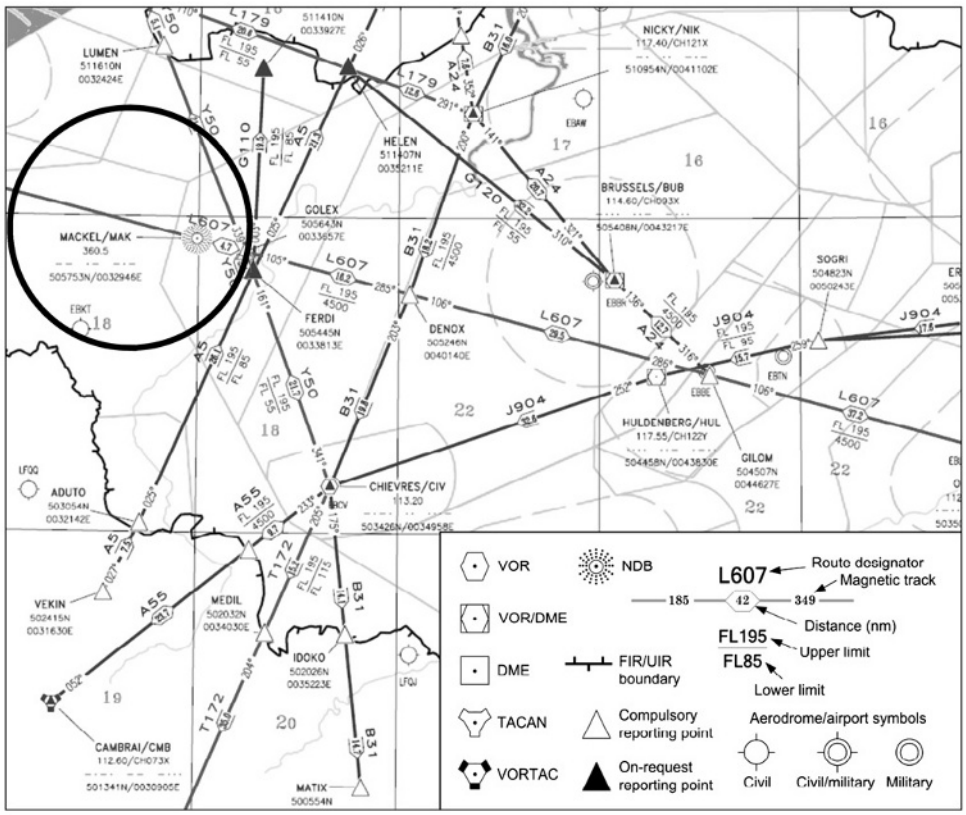
- The sense signal
- A loop signal proportional in amplitude to the cosine of the relative angle of the aircraft centreline and received signal
- A loop signal proportional in amplitude to the sine of the relative angle of the aircraft centreline and received signal.

The sense antenna signal is processed in the receiver via a superhet receiver which allows weak signals to be identified, together with discrimination of adjacent frequencies. The output of the superhet receiver is then integrated into the aircraft's audio system. Loop antenna signals are summed with the sense antenna signal; this forms a phase-modulated (PM) carrier signal. The superhet intermediate frequency (IF) is coupled with the PM signal into a coherent demodulator stage that senses the presence of a sense antenna signal from the IF stage. The PM component of the signal is recovered from the voltage controlled oscillator (VCO) phase lock circuit. This recovered signal contains the bearing information received by the antenna and is compared to a reference modulation control signal.



(a) Using an ADF system for navigation

(b) ADF non-directional beacon (NDB) (photo courtesy of T. Diamond)



(c) MACKEL NDB shown on a navigation chart

Figure 6.4-2 Navigation by non-directional beacons (NDBs).

Receivers based on analog technology send bearing data to the flight deck displays using synchro systems. Digital receivers transmit bearing data to the displays using a data bus system, typically ARINC 429. The ADF receiver is often incorporated into a multi-mode receiver along with other radio navigation systems.

Table 6.4-1 Examples of NDB codes and frequencies.

Name	Identification code	Frequency (kHz)
Eelde	SO	330.00
Eindhoven	EHN	397.00
Eindhoven	PH	316.00
Gull	GUL	383.50
Maastricht	NW	373.00
Maastricht*	ZL	339.00
Rotterdam	ROT	350.50
Rotterdam*	PS	369.00
Rotterdam*	RR	404.50
Schiphol	CH	388.50
Schiphol*	NV	332.00
Stad	STD	386.00
Stadskanaal	STK	315.00
Thorn	THN	434.00
Twenthe	TWN	335.50

*Locator beacons

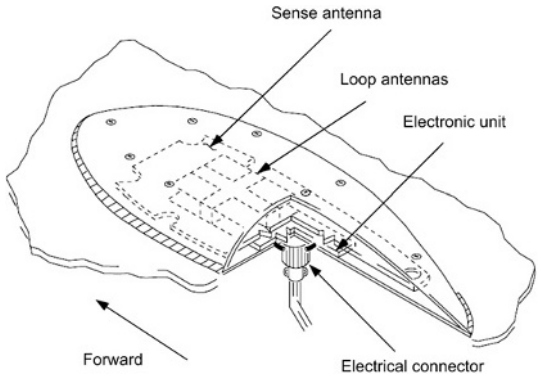


Figure 6.4-3 ADF antenna.

6.4.3.3 Control Panel

Aircraft with analog (electromechanical) avionics have a dedicated ADF control panel, located at the centre pedestal, see Figure 6.4-6(a). An alternative panel shown in Figure 6.4-6(b) enables the crew to select a range of functions including: frequency selectors/displays and the beat frequency oscillator (BFO). This function is used when they want to create an audio frequency for carrier wave transmissions through their audio panel.

NDB carrier waves that are not modulated with an audio component use the BFO circuit in the ADF receiver. To produce an audio output, the receiver heterodynes (beats) the carrier wave signal with a separate signal derived from an oscillator in the receiver.

Some ADF panels have an ADF/ANT switch where 'ADF' selects normal operation, i.e. combined sense and loop antennas; and 'ANT' selects the sense antenna by itself so that the crew can confirm that a station is broadcasting, i.e. without seeking a null. General aviation products combine the control panel and receiver into a single item, see Figure 6.4-6(c). A changeover switch is used to select the active and standby frequencies.

6.4.3.4 ADF Bearing Display

The output from the ADF receiver is transmitted to a display that provides the pilot with both magnetic heading and direction to the tuned NDB, this can either be a dedicated ADF instrument as shown in Figure 6.4-7(a), or be a **radio magnetic indicator** (RMI), see Figure 6.4-7(b).

In the RMI, two bearing pointers (coloured red and green) are associated with the two ADF systems and allow the crew to tune into two different NDBs at the same time. RMIs can have a dual purpose; pilots use a switch on the RMI

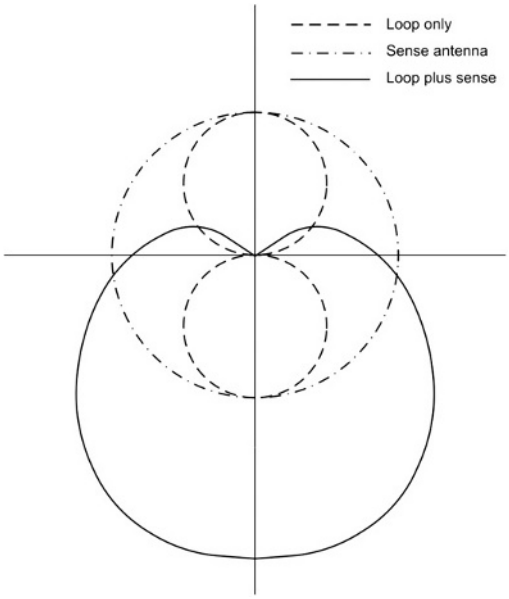


Figure 6.4-4 Polar diagram for the ADF loop and sense antennas.

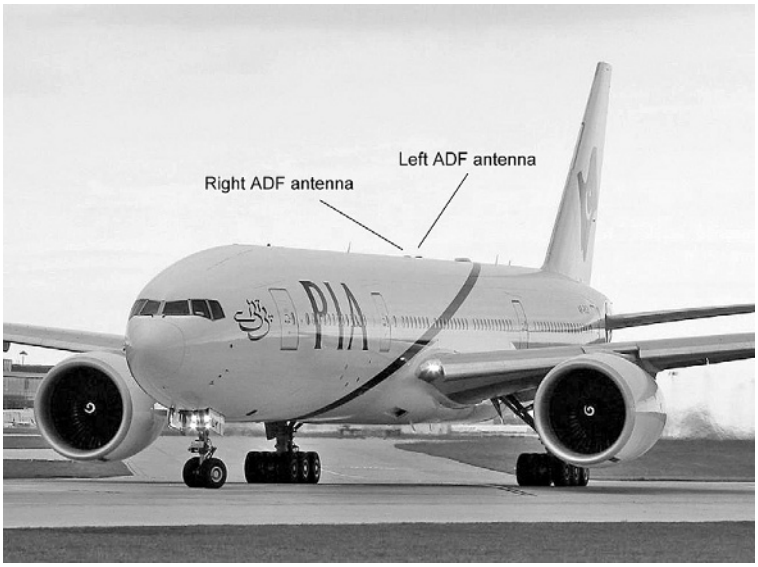


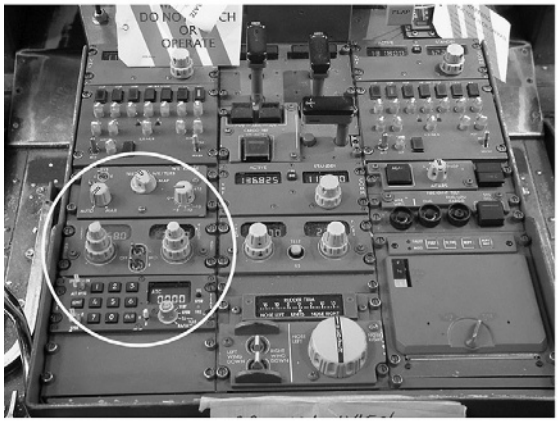
Figure 6.4-5 Location of left and right ADF antennas on a typical transport aircraft.

to select either ADF and/or VHF omnidirectional range (VOR) bearings (see Chapter 6.5 for the latter). Referring to Figure 6.4-7(c), some aircraft have a bearing source indicator (located adjacent to the RMI) that confirms ADF or VOR selection.

The evolution of digital electronics together with integration of other systems has led to the introduction of the flight management system (FMS: see Chapter 6.7) control display unit (CDU) which is used to manage the ADF system. Aircraft fitted with electronic flight instrument systems (EFISs) have green NDB icons displayed on the electronic horizontal situation indicator (EHSI) as shown in Figure 6.4-7(c).

Key Point

ADF is a short-medium-range (200 nm) navigation system operating within the frequency range 190 to 1750 kHz, i.e. low and medium frequency bands. The ADF system uses an orthogonal antenna consisting of two loops: one aligned with the centreline of the fuselage, the other at right angles.



(a) Location of ADF control panel

(b) Typical ADF control panel

(c) ADF panel/receiver for general aviation

Figure 6.4-6 ADF control panels.

Test your understanding 6.4.1

Why does the ADF system seek a null rather than the maximum signal strength from a transmitting station?

Test your understanding 6.4.2

Explain the function of the ADF/ANT switch that is present on some ADF panels.

Test your understanding 6.4.3

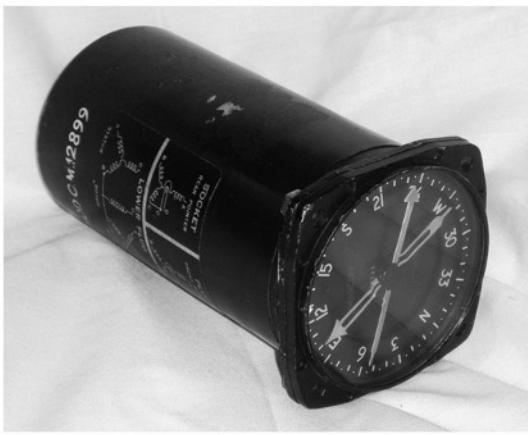
Explain the purpose of a BFO and why it is needed in an ADF receiver.

6.4.4 Operational Aspects of ADF

ADF radio waves are propagated as ground waves and/or sky waves. Problems associated with ADF are inherent in the frequency range that the system uses. ADF transmissions are susceptible to errors from:

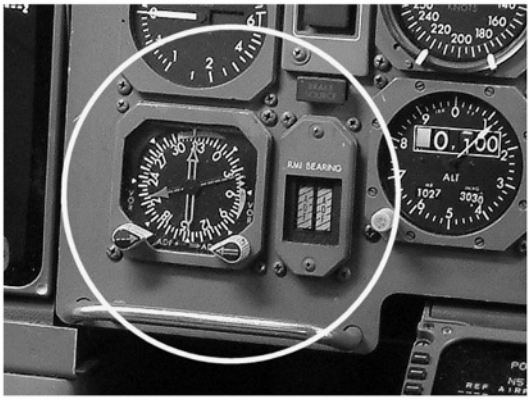
- **Atmospheric conditions:** the height and depth of the ionosphere will vary depending on solar activity. The sky waves (see Figure 6.4-8) will be affected accordingly since their associated skip distances will vary due to refraction in the ionosphere. This is particularly noticeable at sunrise and sunset.

(a) ADF bearing indicator

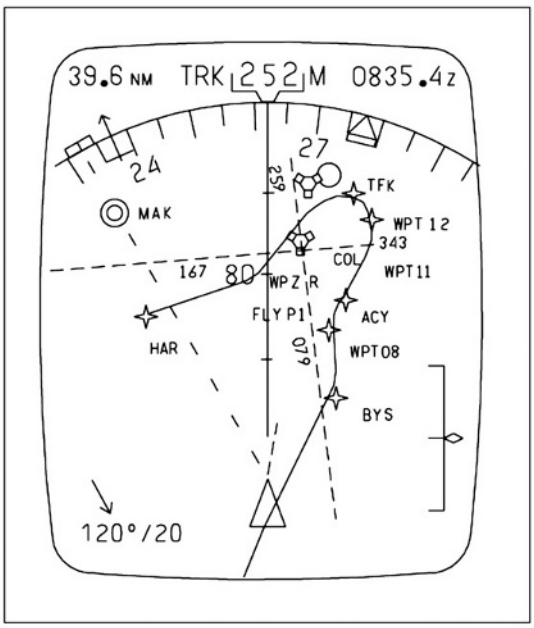


(b) RMI with two bearing indications

(c) Location of RMI and source indicator



(d) RMI and source indicator



(e) EHSI with an NDB icon (shown as MAK on the upper left of the display)

Figure 6.4-7 ADF displays.

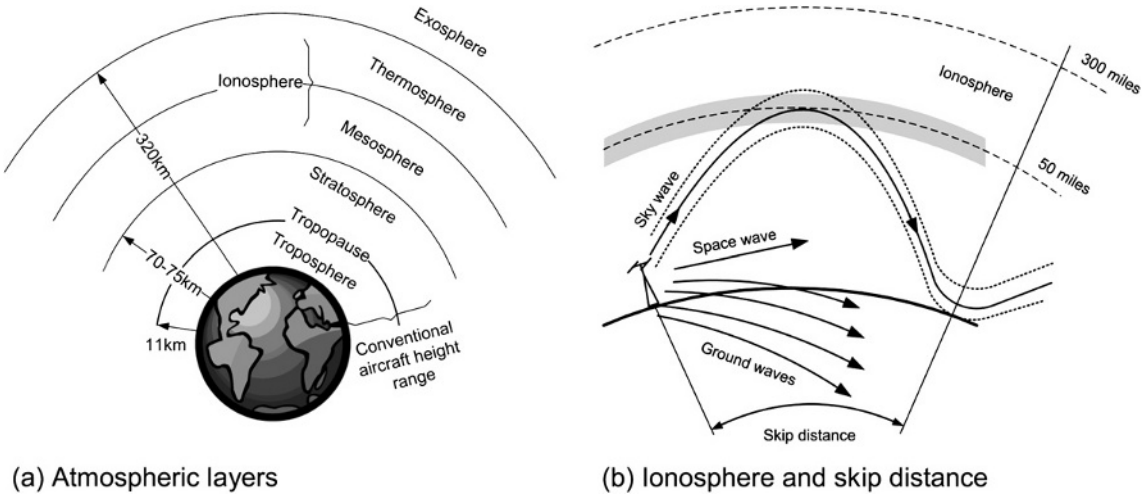


Figure 6.4-8 Sky waves and the ionosphere.

- **Physical aspects of terrain:** mountains and valleys will reflect the radio waves causing multipath reception.
- **Coastal refraction:** LF waves that are propagated across the surface of the earth as ground waves will exhibit different characteristics when travelling over land versus water. This is due to the attenuation of the ground wave being different over land and water. The direction of a radio wave across land will change (see Figure 6.4-9) when it reaches the coast and then travels over water. This effect depends on the angle between the radio wave and the coast.
- **Quadrantal error (QE):** many parts of the aircraft structure, e.g. the fuselage and wings, are closely matched in physical size to the wavelength of the ADF radio transmissions. Radiated energy is absorbed in the airframe and re-radiated causing interference; this depends on the relative angle between the direction of travel, the physical aspects of the aircraft and location of the ADF transmitter. Corrections can be made for QE in the receiver.
- **Interference:** this can arise from electrical storms, other radio transmissions, static buildup/discharges and other electrical equipment on the aircraft.

The accuracy of an ADF navigation system is in the order of ±5 degrees for locator beacons and ±10 degrees for en route beacons. Any of the above conditions will lead to errors in the bearing information displayed on the RMI. If these conditions occur in combination then the navigation errors will be significant. Pilots cannot use ADF for precision navigation due to these limitations.

The increased need for more accuracy and reliability of navigation systems led to a new generation of en route radio navigation aids; this is covered in the next chapter. In the meantime, ADF transmitters remain installed throughout the world and the system is used as a secondary radio navigation aid. The equipment remains installed on modern aircraft, albeit integrated with other radio navigation systems.

Key Point

ADF radio waves are propagated as ground waves and/or sky waves. Problems associated with ADF are inherent in the frequency range that the system uses.

Key Point

ADF cannot be used for precision navigation due to inherent performance limitations; it remains as a backup to other navigation systems.

Test your understanding 6.4.4

Why do ADF antennas need a sense loop?

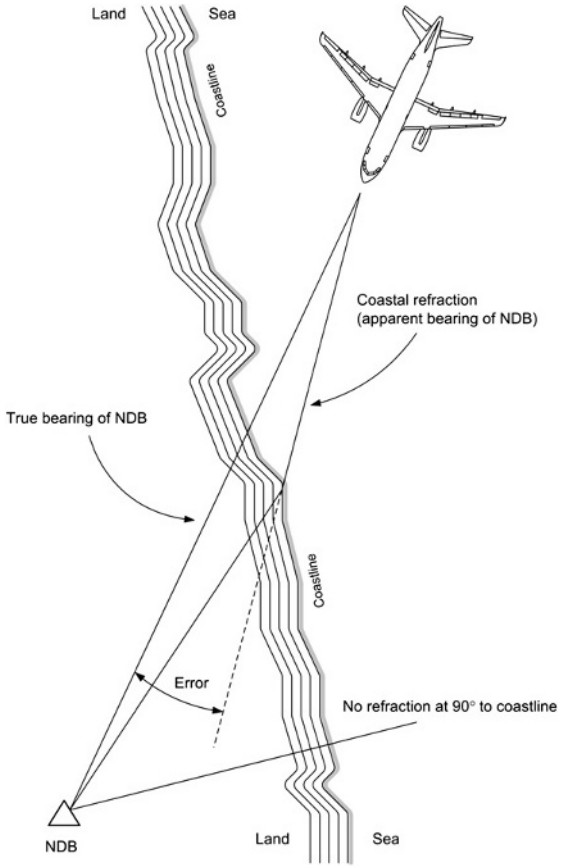


Figure 6.4-9 Effect of coastal refraction.

Test your understanding 6.4.5

How are NDBs identified on navigation charts?

Test your understanding 6.4.6

Where would locator beacons be found?

Test your understanding 6.4.7

Why are there two pointers on the RMI?

Test your understanding 6.4.8

Describe how ground and sky waves are affected by:

(a) local terrain
(b) the ionosphere
(c) attenuation over land and water
(d) electrical storms.

Test your understanding 6.4.9

Explain, in relation to an ADF system, what is meant by quadrantal error. What steps can be taken to reduce this error?

6.4.5 Multiple Choice Questions

1. ADF antennas are used to determine what aspect of the transmitted signal?
   (a) wavelength
   (b) null signal strength
   (c) maximum signal strength.

2. The ADF antennas include
   (a) one sense loop and two directional loops
   (b) two sense loops and two directional loops
   (c) two sense loops and one directional loop.

3. ADF operates in the following frequency range:
   (a) MF to VHF
   (b) LF to MF
   (c) VLF.

4. Bearing to the tuned ADF station is displayed on the
   (a) RMI
   (b) NDB
   (c) HSI.

5. The purpose of an ADF sense antenna is to
   (a) provide directional information to the receiver
   (b) discriminate between NDBs and commercial broadcast stations
   (c) combine with the loop antenna to determine a station bearing.

6. The RMI has two pointers coloured red and green; these are used to indicate
   (a) two separately tuned ADF stations
   (b) AM broadcast stations (red) and NDBs (green)
   (c) heading (red) and ADF bearing (green).

7. The bearing source indicator adjacent to the RMI confirms
   (a) ADF or VOR selection
   (b) the NDB frequency
   (c) the NDB bearing.

8. NDBs on navigation charts can be identified by
   (a) five-letter codes
   (b) two-three-letter codes
   (c) triangles.

9. Morse code is used to confirm the NDB's
   (a) frequency
   (b) name
   (c) bearing.

10. During sunrise and sunset, ADF transmissions are affected by
    (a) coastal refraction
    (b) static build-up in the airframe
    (c) variations in the ionosphere.

11. NDBs associated with the final approach to an airfield are called:
    (a) locator beacons
    (b) reporting points
    (c) en route navigation aids.

12. Quadrantal error (QE) is associated with the:
    (a) ionosphere
    (b) physical aspects of terrain
    (c) physical aspects of the aircraft structure.

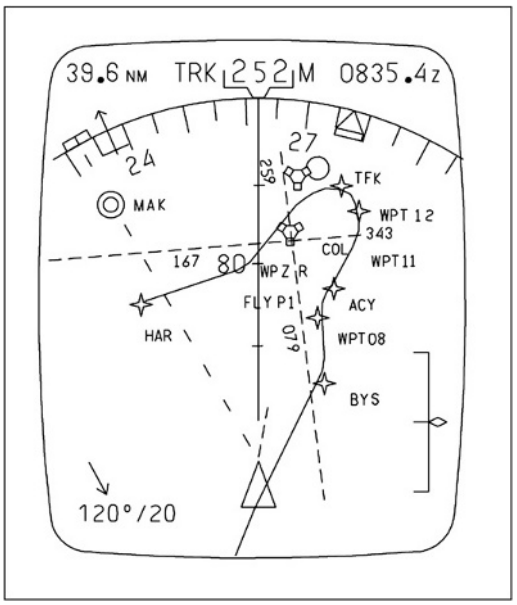


Figure 6.4-10 See Question 16.

13. ADF ground waves are affected by
    (a) the ionosphere
    (b) coastal refraction and terrain
    (c) solar activity.

14. ADF sky waves are affected by
    (a) the ionosphere
    (b) coastal refraction
    (c) the local terrain.

15. A BFO can be used to establish
    (a) the non-directional output of an NDB
    (b) which loop antenna is receiving a null
    (c) an audio tone for an NDB.

16. Referring to Figure 6.4-10, which icon is for the NDB?
    (a) HAR
    (b) MAK
    (c) WPT08.

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.5 VHF Omnidirectional Range

Mike Tooley and David Wyatt

It is well known that radio waves have directional characteristics. In Chapter 6.4, we looked at the early use of radio navigation, and some of the operational problems in using low-frequency (LF) and medium-frequency (MF) transmissions. During the mid- to late 1940s, it was evident to the aviation world that an accurate and reliable short-range navigation system was needed. Since radio communication systems based on very high frequency (VHF) were being successfully deployed, a decision was made to develop a radio navigation system based on VHF. This system became the VHF omnidirectional range (VOR) system, a system that is in widespread use throughout the world today and is the basis for the current network of 'airways' that are used for navigation.

6.5.1 VOR Principles

6.5.1.1 Overview

VOR is a short/medium-range navigation system operating in the 108–117.95 MHz range of frequencies. This means that the radio waves are now propagated as space waves. The problems that were encountered with ground and sky waves in the LF and MF ranges are no longer present with a VHF system. VOR navigation aids are identified by unique three-letter codes (derived from their name, e.g. London VOR is called LON, Dover VOR is called DVR, etc.). The code is modulated onto the carrier wave as a 1020 Hz tone that the crew can listen to as a Morse code signal.

Some VOR navigation aids have an automatic voice identification announcement that provides the name of the station; this alternates with the Morse code signal. The location of the VOR navigation aids (specified by latitude and longitude) together with their carrier wave frequencies is provided on navigation charts as with ADF.

VOR operates in the same frequency range as the instrument landing system (ILS). Although the two systems are completely independent and work on totally different principles, they often share the same receiver. The two systems are differentiated by their frequency allocations within this range. ILS frequencies are allocated to the odd tenths of each 0.5 MHz increment, e.g. 109.10 MHz, 109.15 MHz, 109.30 MHz, etc. VOR frequencies are allocated even tenths of each 0.5 MHz increment, e.g. 109.20 MHz, 109.40 MHz, 109.60 MHz, etc. Table 6.5-1 provides an illustration of how these frequencies are allocated within the 109 MHz range. This pattern is applied from 108 to 111.95 MHz.

In addition to the inherently improved system performance and navigation reliability, VOR has another feature that makes it extremely useful for air navigation. The VOR system has the ability to transmit specific bearing information, referred to as a **'radial'**, see Figure 6.5-1(a). The pilot can select any radial from a given VOR navigation aid and fly to or from that aid. Since the system is 'line of sight', i.e. receiving signals as space waves, the altitude of the aircraft will have a direct relationship with the range within which the system can be used, see Figure 6.5-1(b).

Using VHF navigation aids imposes a limit on the theoretical working range that can be obtained. The maximum theoretical line-of-sight (LOS) distance between an aircraft and the ground station is given by the relationship

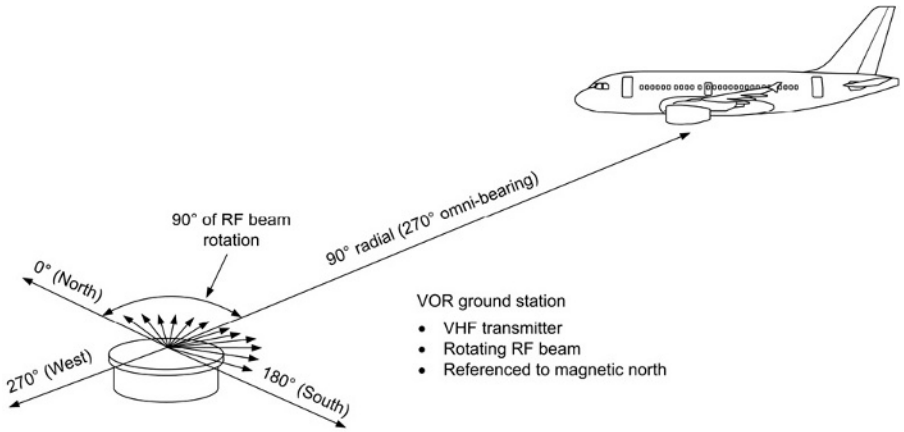
$$d = 1.1\sqrt{h}$$

where d is the distance in nautical miles, and h is the altitude in feet above ground level (assumed to be flat terrain). The theoretical LOS range for altitudes up to 20,000 feet is given in Table 6.5-2.

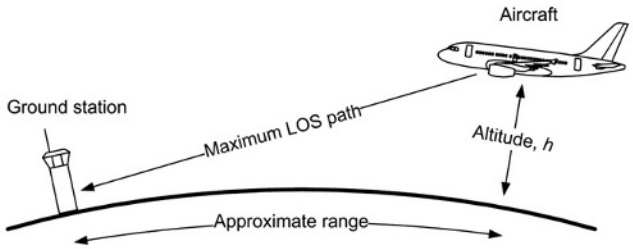
At higher altitudes, it is possible to receive VOR signals at greater distances but with reduced signal integrity. Although the actual range also depends on transmitter power and receiver sensitivity, the above relationship provides a good approximation. In practice, navigation aids have a designated standard service volume (SSV); this defines the reception limits within an altitude envelope as shown in Table 6.5-3.

Table 6.5-1 Allocation of ILS and VOR frequencies.

ILS frequency (MHz)	VOR frequency (MHz)
	109.00
109.10	
109.15	
	109.20
109.30	
109.35	
	109.40
109.50	
109.55	
	109.60
109.70	
109.75	
	109.80
109.90	
109.95	



(a) VHF omni-range (VOR) overview



(b) VHF omni-range—line of sight

Figure 6.5-1 VOR overview.

Table 6.5-2 Theoretical LOS range.

Altitude (feet)	Range (nm)
100	10
1,000	32
5,000	70
10,000	100
20,000	141

Table 6.5-3 Navigation aid classifications.

Classification	Altitude (feet)	Range (nm)
Terminal	1,000–12,000	25
Low altitude	1,000–18,000	40
High altitude	18,000–45,000	130

Key Point

VOR radials are referenced to magnetic north; they are the basis of airways for en route navigation.

Key Point

VOR transmissions are 'line of sight'; therefore, range increases with increased altitude.

6.5.1.2 Conventional VOR (CVOR)

There are two types of VOR ground navigation transmitter: conventional and Doppler. The conventional VOR (CVOR) station radiates two signals: omnidirectional and directional on a continuous basis. The omnidirectional (or reference) signal is the carrier wave frequency of the station, amplitude modulated to 30 Hz. The directional signal is radiated as a cardioid pattern rotating at 30 revolutions per second. The subcarrier frequency is 9960 Hz, frequency modulated in the range 9960 ±480 Hz at 30 Hz.

The directional signal is arranged to be in phase with the reference signal when the aircraft is due north (magnetic) of the VOR station. As the cardioid pattern rotates around the station, the two signals become out of phase on a progressive basis, see Figure 6.5-2. The bearing between any given angle and magnetic north is determined by the receiver as the phase angle difference between the reference and directional signals. This difference in phase angle is resolved in the aircraft receiver and displayed to the crew as a radial from the VOR station, see Figure 6.5-3.

Locations of CVOR ground stations have to be carefully planned to take into account local terrain and obstacles. Mountains and trees can cause multipath reflections resulting in distortion (known as siting errors) of the radiated signal. These errors can be overcome with an enhanced second-generation system known as Doppler VOR (DVOR).

6.5.1.3 Doppler VOR (DVOR)

Doppler is usually associated with self-contained navigation systems. The Doppler effect is also applied to the second-generation version of VOR ground transmitters. The Doppler effect can be summarised here as: '…the frequency of a wave apparently changes as its source moves closer to, or farther away from an observer'.

The DVOR ground station has an omnidirectional transmitter in the centre, amplitude modulated at 30 Hz; this is the reference phase. The directional signal is derived from a 44 feet diameter circular array comprising up to 52 individual antennas, see Figure 6.5-4(a) and (b). Each antenna transmits in turn to simulate a rotating antenna.

Consider two aircraft using the DVOR station as illustrated in Figure 6.5-4(c). The effect of the rotating 9960 Hz signal is to produce a Doppler shift; aircraft A will detect a decreased frequency, aircraft B will detect an increased frequency. Doppler shift creates a frequency modulated (FM) signal in the aircraft receiver over the range 9960 Hz ±480 Hz varying at 30 Hz in a sine wave. Note that the perceived frequency will be 9960 Hz when the aircraft are in the positions shown in Figure 6.5-4(c). The phase difference measured in the airborne equipment depends on the bearing of the aircraft relative to the station. Since the FM variable signal is less prone to interference, DVOR performance is superior to CVOR.

DVOR actually uses two rotating patterns as shown in Figure 6.5-4(d). These patterns (diagonally opposite each other) rotate at 30 revolutions per second; one pattern is 9960 Hz above the reference, the other is 9960 Hz below the reference frequency. The diameter of the array, together with the speed of pattern rotation, creates a Doppler shift of 480 Hz (at VOR frequencies).

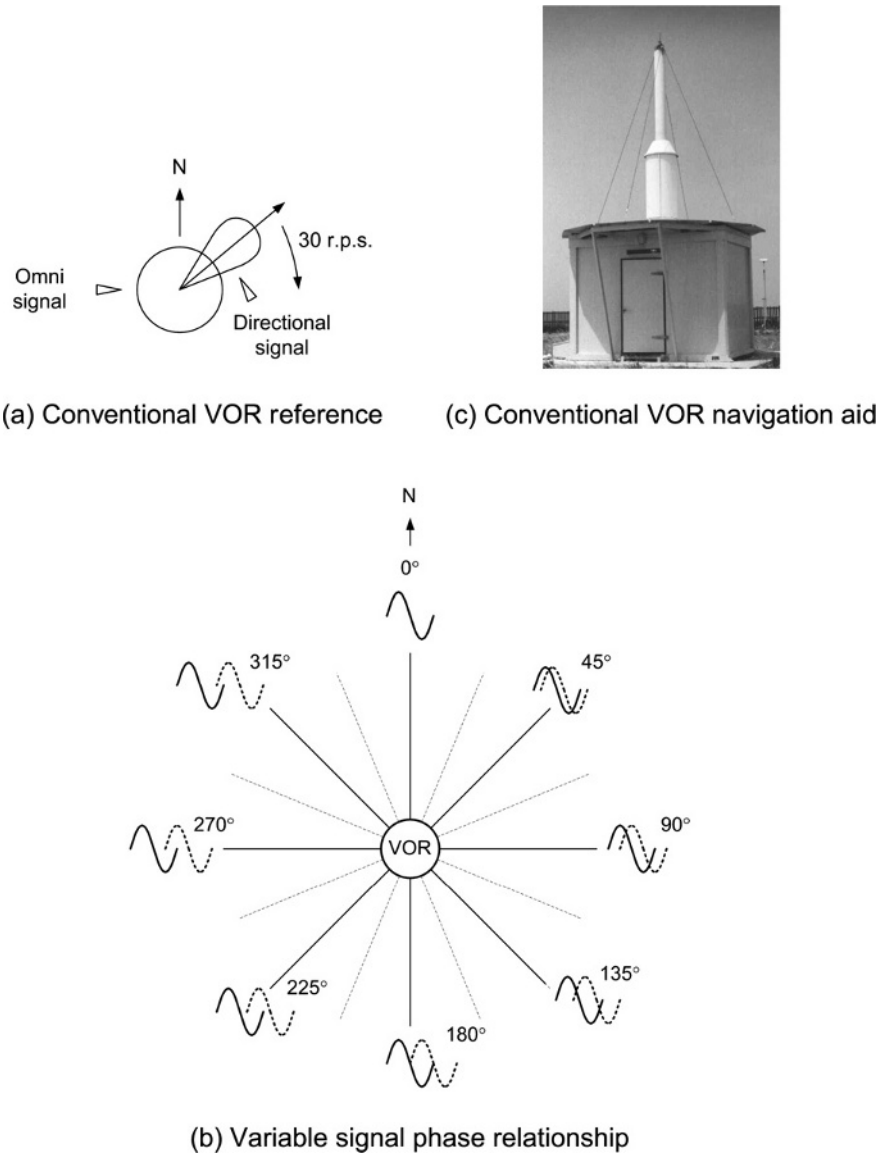


Figure 6.5-2 Conventional VOR (CVOR).

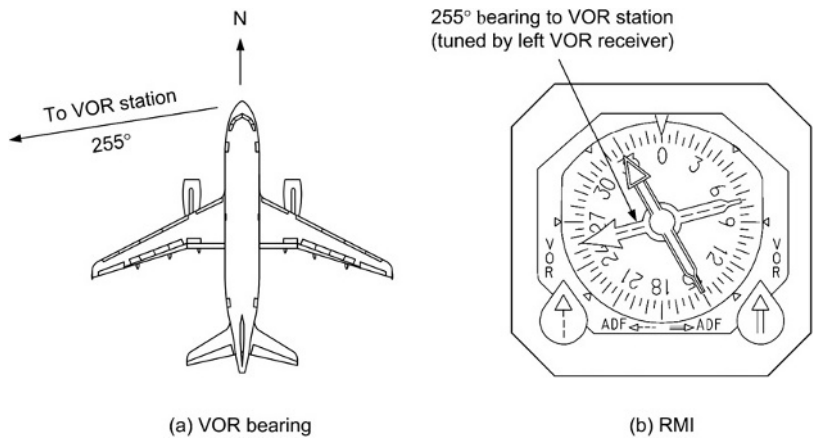
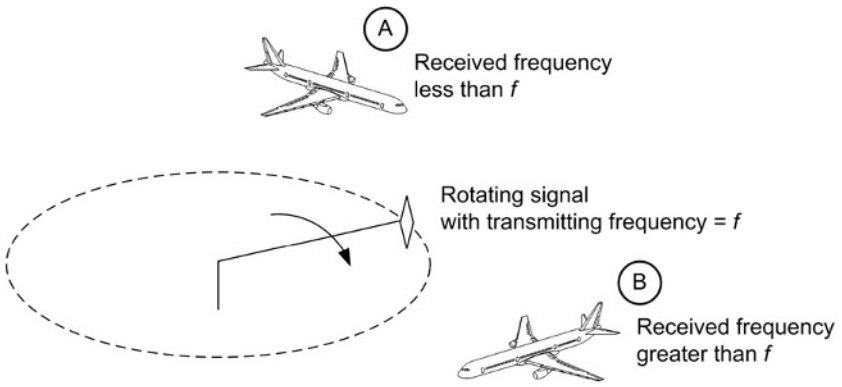


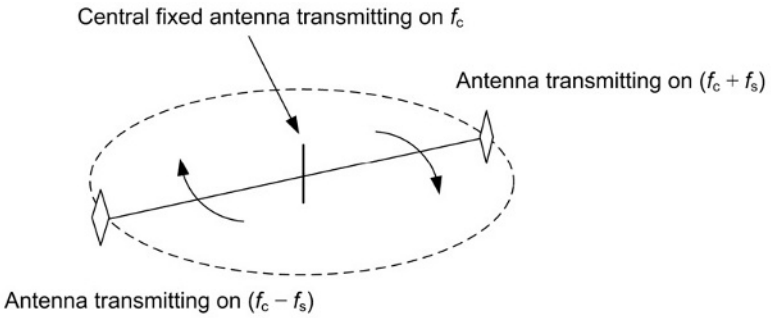
Figure 6.5-3 VOR bearings and displays.

(a) Doppler VOR (DVOR) navigation aid
(photo courtesy of T. Diamond)

(b) Doppler VOR (DVOR) navigation aid
(photo courtesy of T. Diamond)



(c) Rotating transmitting signal



(d) Double sideband Doppler VOR

Figure 6.5-4 Doppler VOR (DVOR).

6.5.2 Airborne VOR Equipment

Modern transport aircraft have two VOR systems often designated left and right; note that airborne equipment is the same for conventional and DVOR. Radio frequency (RF) signals from the antenna are processed in the receiver as determined by frequency and course selections from the control panel; outputs are sent to various displays.

6.5.2.1 Antennas

The VOR antenna is a horizontally polarised, omnidirectional half-wave dipole, i.e. a single conductor with a physical length equal to half the wavelength of the VOR signals being received. Two such antennas are formed into a single package and usually located in the aircraft fin as indicated in Figure 6.5-5. They are packaged within a composite fairing for aerodynamic streamline purposes. The antennas are connected to the receivers via coaxial cables.

6.5.2.2 Receivers

VOR receivers are often combined with other radio navigation functions, e.g. the ILS receivers are located in the avionic equipment bay, see Figure 6.5-6.

VOR receivers are based on the super-heterodyne principle with tuning from the control panel. The received radio frequency signal is passed through an amplitude modulation filter to separate out the:

- 30 Hz tone from the rotating pattern
- voice identification (if provided from the navigation aid)
- Morse code tone; 9960 Hz signal FM by ±480 Hz at 30 Hz reference tone.

Voice and Morse code tones are integrated with the audio system. A comparison of the phase angles of the variable and reference 30 Hz signals produces the VOR radial signal. Receivers based on analog technology interface with the flight deck displays using synchro systems. Digital receivers interface with other systems using a data bus system, typically ARINC 429. Receivers usually combine both VOR and instrument landing system functions.

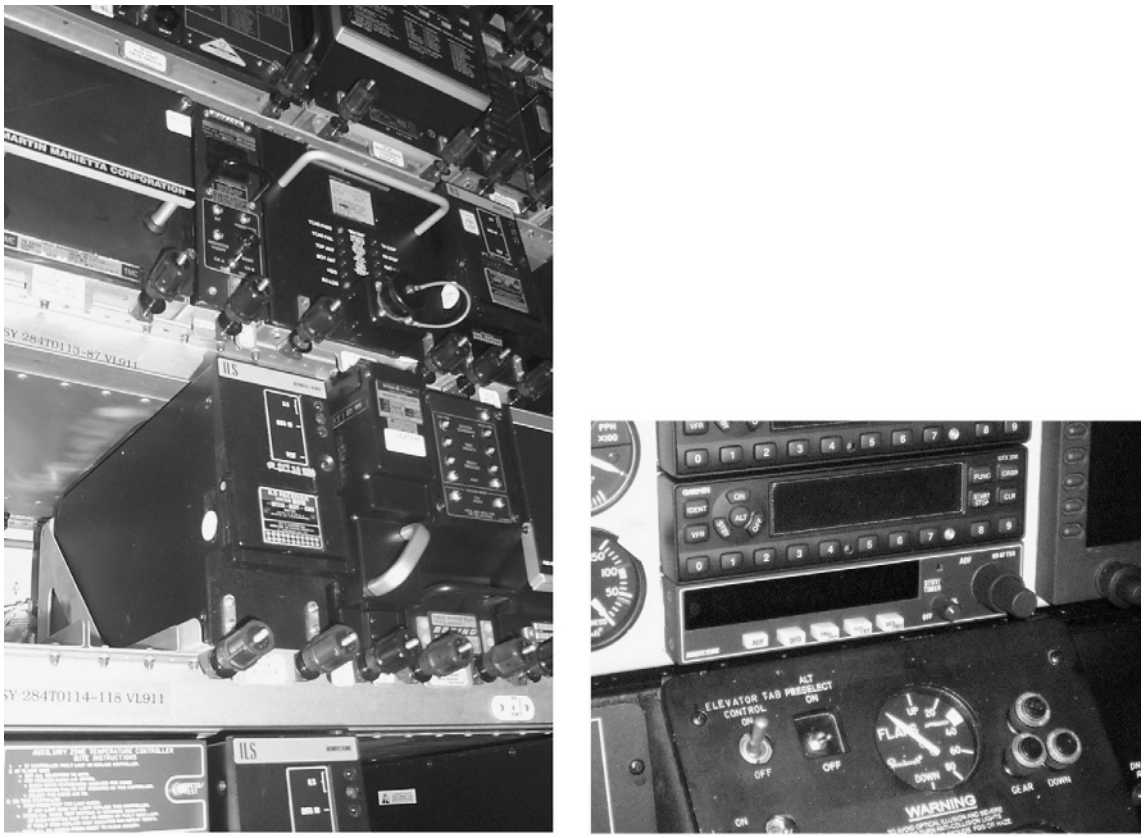
6.5.2.3 Control Panels

Control panels identified as 'VHF NAV' can be located on the glare shield (as shown in Figure 6.5-7) or centre pedestal. This panel is used to select the desired course and VOR frequencies. General aviation products have a combined VHF navigation and communications panel—see Figure 6.5-6(b)—this can be integrated with the GPS navigation panel.

6.5.2.4 VOR Displays

The bearing to a VOR navigation aid (an output from the receiver) can be displayed on the radio magnetic indicator (RMI); this is often shared with the ADF system as discussed in the previous chapter. The RMI, Figure 6.5-8(a), provides the pilot with both magnetic heading and direction to the tuned VOR navigation aid. The two bearing pointers (coloured red and green) are associated with the two VOR systems and allow the crew to tune into two different

Figure 6.5-5 Location of VOR antennas.



(a) VOR receiver (remotely located in the aircraft's avionic equipment bay)

(b) Navigation and communications panel/ receiver used in general aviation

Figure 6.5-6 VOR receivers.

VOR navigation aids at the same time. On some instruments, a switch on the RMI is used to select either ADF or VOR bearing information, see Figure 6.5-8(b). RMIs therefore have a dual purpose; pilots use a switch on the RMI to select either ADF and/or VOR bearings (see Chapter 6.4 for ADF). Some aircraft have a bearing source indicator adjacent to the RMI to confirm ADF or VOR selection, see Figure 6.5-8(c).

In order to fly along an airway, first it has to be intercepted. This is achieved by flying towards the desired radial on a specified heading. The method of displaying the VOR radial depends on the type of avionic fit. Electromechanical instruments include the omni-bearing selector (OBS) and course deviation indicator (CDI), see Figure 6.5-9. The OBS indicator has a number of features; the selector is used to manually rotate the course card. This card is calibrated from 0 to 360° and indicates the VOR bearing selected to fly TO or FROM.

In Figure 6.5-9(a), a VOR radial of 345° has been selected. The deviation pointer moves to the left or right to guide the pilot in the required direction to maintain the selected course. Each dot on the scale represents a 2° deviation from the selected course. The back-course (BC) indicator indicates when flying FROM the VOR navigation aid. On some instruments, the BC indicator is replaced by a TO/FROM display in the form of an arrow. A red OFF flag indicates when the:

- VOR navigation aid is beyond reception range
- pilot has not selected a VOR frequency
- VOR system is turned off, or is inoperative.

An updated version of this instrument is the CDI. This has a compass display and course selector as shown in Figure 6.5-9(b). The course selector (lower right-hand side of instrument) is set to the desired VOR radial; a deviation pointer moves left or right of the aircraft symbol to indicate if the aircraft is to the right or left of the selected radial.

For aircraft with electronic flight instruments, the desired radial is displayed on the electronic horizontal situation indicator (EHSI). This EHSI display can either be selected to show a conventional compass card (Figure 6.5-10(a))

Figure 6.5-7 Typical VOR control panel.

or expanded display (Figure 6.5-10(b)). As the radial is approached the deviation bar gradually aligns with the selected course.

If flying manually, the crew turn the aircraft onto the selected course whilst monitoring the deviation bar; when it is centred, the radial has been intercepted and the EHSI will display 'TO' confirming that the inbound radial is being followed. The flight continues until the VOR navigation is reached, and the radial to the next navigation aid is selected. If the EHSI were still selected to the original inbound radial, the EHSI would display 'FROM'. The lateral deviation bar therefore shows if the aircraft is flying on the selected radial, or if it is to the left or right of the radial.

Key Point

Two intersecting VOR radials can be used to define unique locations known as reporting points; these are used for air traffic control purposes.

Key Point

DVOR was introduced to overcome siting problems found with CVOR. The two systems operate on different principles; however, the airborne equipment is the same.

6.5.3 Operational Aspects of VOR

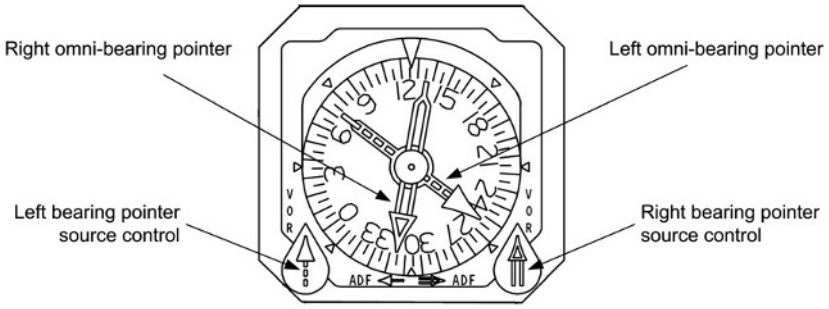
Radials from any given VOR navigation aid are the basis of **airways**; system accuracy is typically within one degree. These are the standard routes flown by aircraft when flying on instruments. When two VOR radials intersect, they provide a unique navigation position fix based (theta–theta). The accuracy of this fix is greatest when the radials intersect at right angles. Figure 6.5-11 illustrates how navigation charts are built up on a network of VOR radials; the accuracy of VOR radials is generally very good (±1 degree). In this illustration, the navigation aid located near Brussels (abbreviated BUB) transmits on 114.6 MHz. Three radials can be seen projected from this navigation aid on 136°, 310° and 321°. These radials are used to define airways A24 and G120. Note the Morse code output and latitude/longitude for the navigation aid.

The intersecting radials from navigation aids are used to define reporting points for en route navigation. These reporting points are given five-letter identification codes associated with their geographic location. For example, the reporting point HELEN is defined by airways G5 and A120. The flexibility of VOR is greatly increased when co-located distance measuring equipment (DME) is used, thereby providing rho–theta fixes from a single navigation aid. There are examples of VOR-only navigation aids, e.g. Perth in Scotland (identification code PTH, frequency 110.40 MHz). The majority of VOR navigation aids are paired with DME; this system is described in the next chapter.

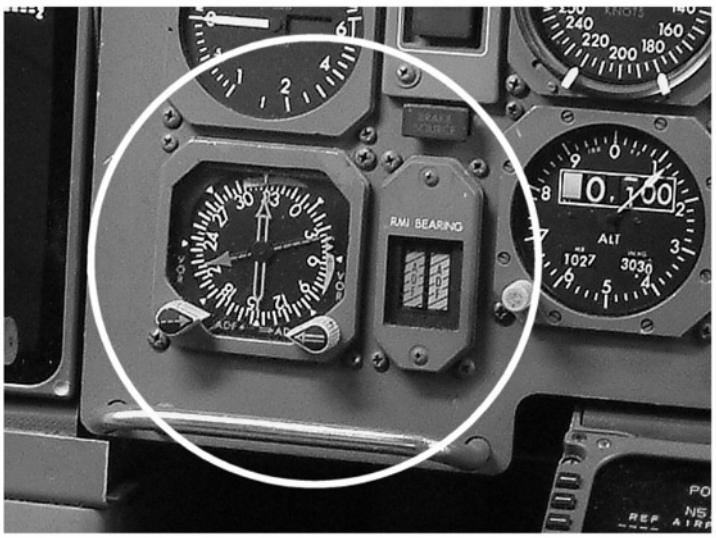
Test your understanding 6.5.1

Why are VOR transmissions 'line of sight' only?

(a) RMI with two bearing pointers



(b) RMI source control (VOR/ADF)



(c) RMI and bearing source annunciator

Figure 6.5-8 VOR displays and indicators.

Test your understanding 6.5.2

Calculate (a) the LOS range for an aircraft at an altitude at 7,500 feet and (b) the altitude of an aircraft that would yield a LOS range of 200 nautical miles.

Figure 6.5-9 Omni-bearing selector and course deviation indicator.

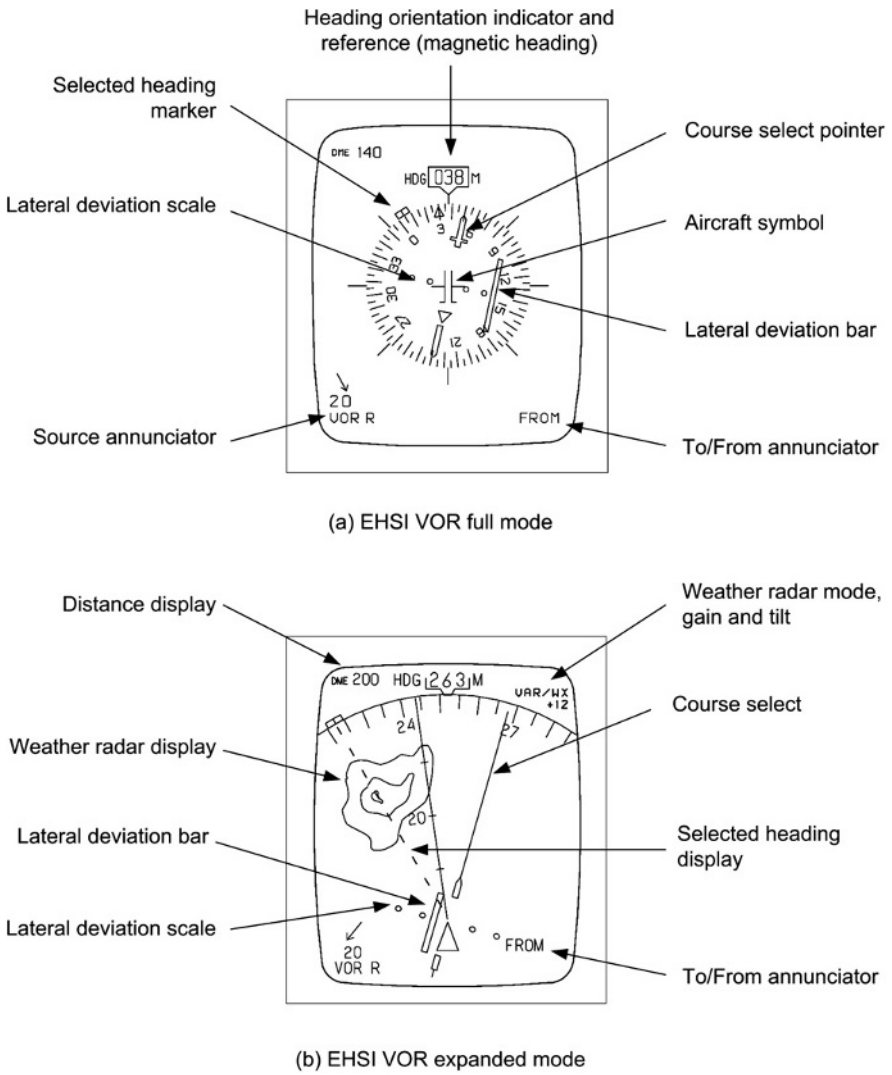


Figure 6.5-10 VOR electronic displays.

Test your understanding 6.5.3

How can the crew identify a specific VOR navigation aid?

Test your understanding 6.5.4

Where can a VOR radial be displayed?

Test your understanding 6.5.5

Explain how a VOR radial is captured.

Test your understanding 6.5.6

Why does an RMI have two VOR pointers?

Test your understanding 6.5.7

What is the difference in aircraft equipment beween conventional and DVOR?

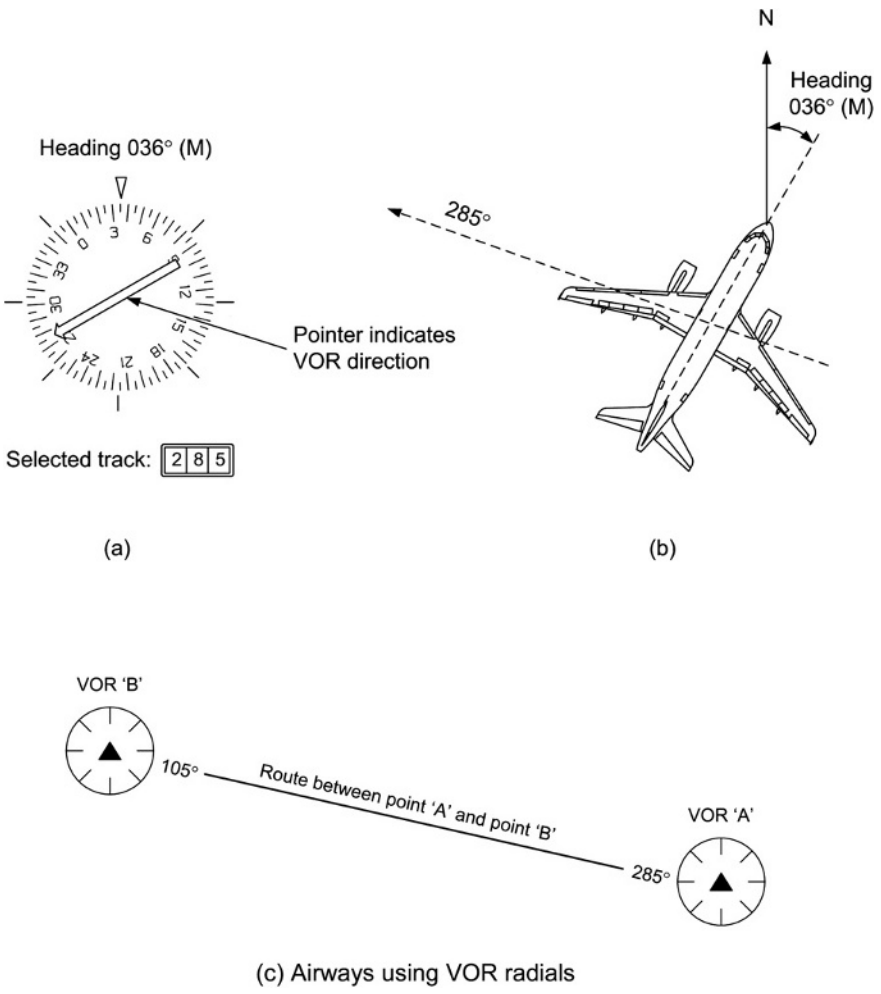
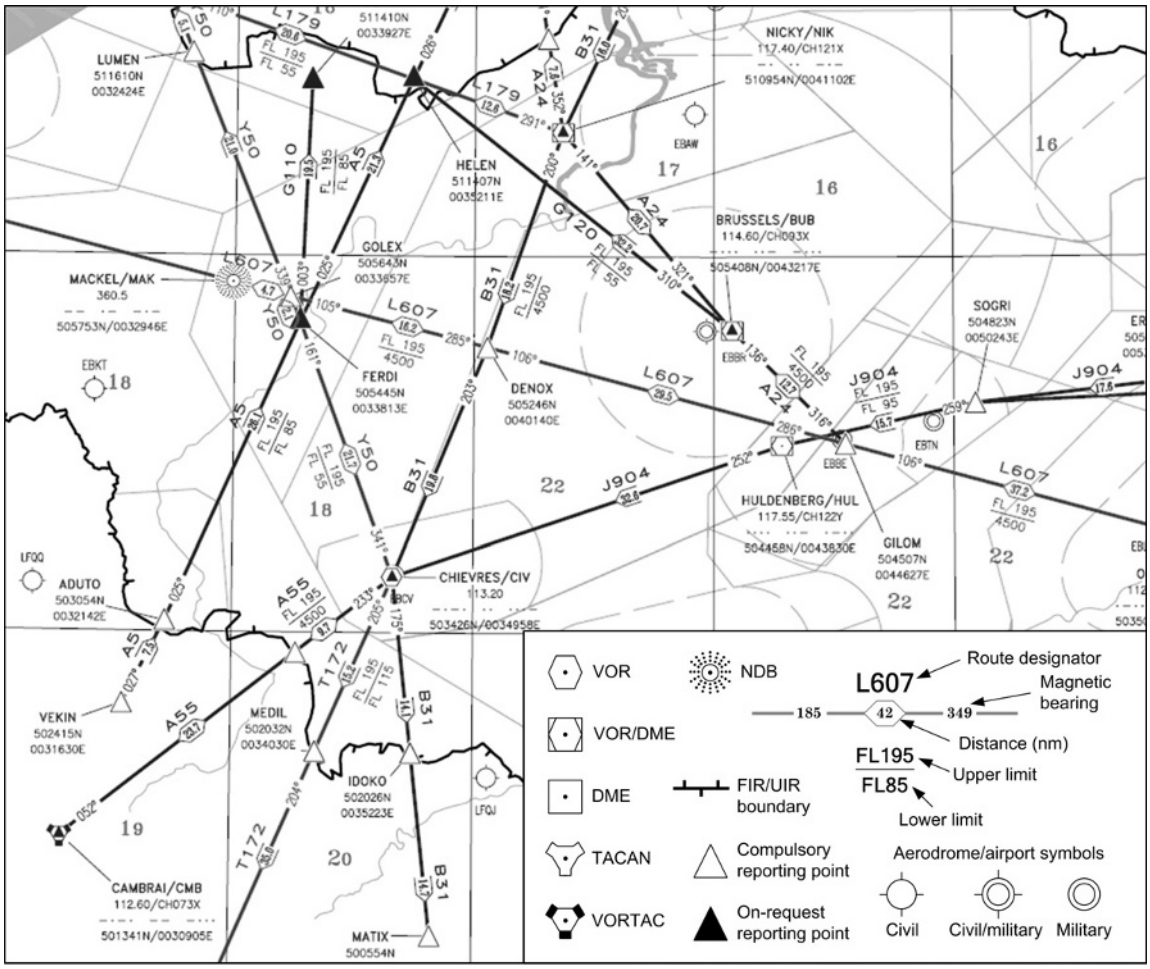


Figure 6.5-11 VOR radials and airways.



(c) Airway network over Belgium

Figure 6.5-11 (continued)

Key Point

Navigation charts are built up on a network of VOR radials.

Test your understanding 6.5.8

What is the Morse code output used for in a VOR transmission?

Key Point

VOR operates in the frequency band extending from 108 MHz to 117.95 MHz. Three alpha characters are used to identify specific VOR navigation aids.

6.5.4 Multiple Choice Questions

1. VOR operates in which frequency range?
   (a) LF
   (b) MF
   (c) VHF.

2. VOR signals are transmitted as what type of wave?
   (a) Sky wave
   (b) Ground wave
   (c) Space wave.
3. Where is the deviation from a selected VOR radial displayed?
   (a) RMI
   (b) HSI
   (c) NDB.
4. At which radial will the directional wave be out of phase by 90 degrees with the non-directional wave?
   (a) 090 degrees
   (b) 000 degrees
   (c) 180 degrees.
5. At which radial will the directional signal be in phase with the non-directional signal?
   (a) 090 degrees
   (b) 000 degrees
   (c) 180 degrees.
6. VOR navigation aids are identified by how many alpha characters?
   (a) Two
   (b) Three
   (c) Five.
7. VOR radials are referenced to:
   (a) non-directional signals from the navigation aid
   (b) magnetic north
   (c) true north.
8. The RMI has two pointers coloured red and green; these are used to indicate:
   (a) the bearing of two separately tuned VOR stations
   (b) directional (red) and non-directional transmissions (green)
   (c) the radials of two separately tuned VOR stations.
9. Morse code tones are used to specify the VOR:
   (a) identification
   (b) frequency
   (c) radial.
10. The intersection of two VOR radials provides what type of position fix?
    (a) Rho–rho
    (b) Theta–theta
    (c) Rho–theta.
11. An aircraft is flying on a heading of 090 degrees to intercept the selected VOR radial of 180 degrees; the HSI will display that the aircraft is:
    (a) right of the selected course
    (b) left of the selected course
    (c) on the selected course.
12. The DVOR navigation aid has an omnidirectional transmitter located in the:
    (a) centre
    (b) outer antenna array
    (c) direction of magnetic north.
13. When flying overhead a VOR navigation aid, the reliability of directional signals:
    (a) decreases
    (b) increases
    (c) stays the same.

Figure 6.5-12 See Question 16.

14. Reporting points using VOR navigation aids are defined by the:
    (a) identification codes
    (b) intersection of two radials
    (c) navigation aid frequencies.

15. With increasing altitude, the range of a VOR transmission will be:
    (a) increased
    (b) decreased
    (c) the same.

16. Referring to Figure 6.5-12, the instrument shown is called the:
    (a) omni-bearing selector (OBS)
    (b) radio magnetic indicator (RMI)
    (c) course deviation indicator (CDI).

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.6 Distance Measuring Equipment

Mike Tooley and David Wyatt

The previous two chapters have been concerned with obtaining directional information for the purposes of airborne navigation. In this chapter we will look at a system that provides the crew with the distance to a navigation aid. Distance measuring equipment (DME) is a short/medium range navigation system, often used in conjunction with the VOR system to provide accurate navigation fixes. The system is based on secondary radar principles, and operates in the L-band of radar. Before looking at what the system does and how it operates in detail, we need to take at look at some basic radar principles.

6.6.1 Radar Principles

The word **radar** is derived from radio detection and ranging; the initial use of radar was to locate aircraft and display their range and bearing on a monitor (either ground based or in another aircraft). This type of radar is termed **primary radar:** Energy is radiated via a rotating radar antenna to illuminate a 'target'; this target could be an aircraft, the ground or a cloud. Some of this energy is reflected back from the target and is collected in the same antenna, see Figure 6.6-1. The strength of the returned energy is measured and used to determine the range of the target. A rotating antenna provides the directional information such that the target can be displayed on a screen.

Primary radar has its disadvantages, one of which is that the amount of energy being transmitted is very large compared with the amount of energy reflected from the target. An alternative method is **secondary radar** that transmits a specific low energy signal (the interrogation) to a known target. This signal is analysed and a new (or secondary) reply signal, i.e. not a reflected signal, is sent back to the origin, see Figure 6.6-2(a). Secondary radar was developed during the Second World War to differentiate between friendly aircraft and ships: Identification Friend or Foe (IFF). The principles of secondary radar now have a number of applications including DME.

6.6.2 DME Overview

The DME navigation aid contains a transponder (receiver and transmitter) contained within a single navigation aid, Figure 6.6-2(b). The aircraft equipment radiates energy pulses to the DME navigation aid; secondary signals are then transmitted back to the aircraft. An on-board **interrogator** measures the time taken for the signals to be transmitted and received at the aircraft. Since we know the speed of radio wave propagation, the interrogator can calculate the distance to the DME navigation aid. DME navigation aids can either be self-contained ground stations, or co-located with a VOR navigation aid, Figure 6.6-2(c).

Since the system is 'line of sight', the altitude of the aircraft will have a direct relationship with the range that the system can be used, see Figure 6.6-3(a). Using DME navigation aids imposes a limit on the working range that can be obtained. The maximum line-of-sight (LOS) distance between an aircraft and the ground station is given by the relationship

$$d = 1.1\sqrt{h}$$

where d is the distance in nautical miles, and h is the altitude in feet above ground level (assumed to be flat terrain). The theoretical LOS range for altitudes up to 20,000 feet is given in Table 6.6-1.

Referring to Figure 6.6-3(b) it can be seen that the actual distance being measured by the interrogator is the '**slant**' **range**, i.e. not the true distance (horizontal range) over the ground. The effects of slant range in relation to the horizontal range are greatest at high altitudes and/or when the aircraft is close to the navigation aid. Taking this to the limit, when the aircraft is flying over a DME navigation aid, it would actually be measuring the aircraft's altitude!

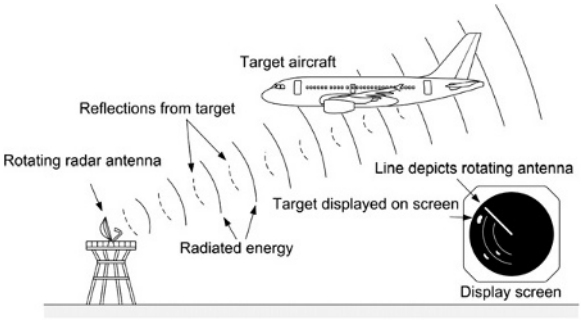
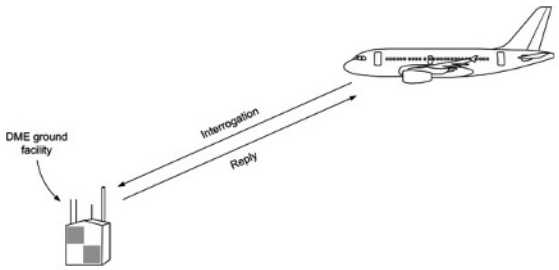


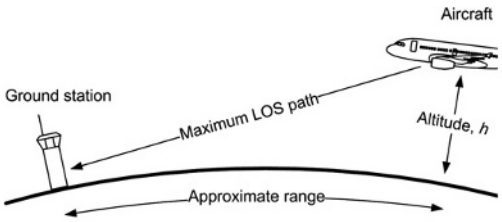
Figure 6.6-1 Primary radar.



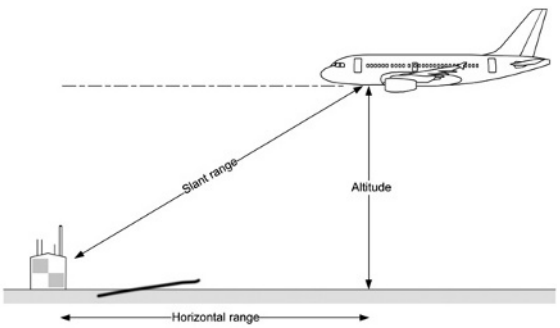
(a) Secondary radar used for DME

(b) DME transponder (right of photo)
(photo courtesy of T. Diamond)

Figure 6.6-2 DME overview.



(a) Line of sight versus altitude



(b) DME slant range

Figure 6.6-3 DME range terminology.

Test your understanding 6.6.1

What is the difference between primary and secondary radar?

Test your understanding 6.6.2

Distinguish between slant range and horizontal range.

Table 6.6-1 Theoretical LOS range.

Altitude (feet)	Range (nm)
100	10
1,000	32
5,000	70
10,000	100
20,000	141

6.6.3 DME Operation

The signals transmitted by the interrogator are a pair of pulses, each of 3.5 ms duration and 12 ms apart modulated on the DME navigation aid frequency. The interrogator generates a pulse-pair repetition rate between 5 and 150 pulse-pairs per second. At the DME navigation aid, the transponder receives these pulses and, after a 50 ms time delay, transmits a new pair of pulses at a frequency 63 MHz above or below the interrogator's frequency. The aircraft's interrogator receives the pulses and matches the time interval between the transmitted pair of pulses. This ensures that other aircraft interrogating the same DME navigation aid at the same time only process their own pulses.

By measuring the elapsed time between transmitting and receiving (and taking into account the 50 ms time delay) the interrogator calculates the distance to the navigation aid. DME is a LOS system with a maximum range of approximately 200 nm; this equates to approximately 2400 ms elapsed time taken for a pair of pulses to be transmitted and received, taking into account the 50 ms time delay in the ground station. System accuracy is typically ±0.5 nm, or 3% of the calculated distance, whichever is the greater.

Test your understanding 6.6.3

What is the typical accuracy and maximum range of a DME system?

Key Point

The varying interval between pulse-pairs ensures that the DME interrogator recognises its own signals and rejects other signals.

Key Point

DME is based on secondary radar; it operates in the L-band between 962 MHz and 1215 MHz (ultrahigh frequency – UHF) with channel spacing at 1 MHz.

6.6.4 Equipment Overview

Commercial transport aircraft are usually fitted with two independent DME systems, comprising antennas and interrogators.

The DME antennas are L-band blades, located on the underside of the aircraft fuselage, see Figure 6.6-4(a); note that the antenna is dual purpose in that it is used for both transmitting and receiving.

The interrogators are located in the equipment bays (Figure 6.6-4(b)) and provide three main functions: transmitting, receiving and calculation of distance to the selected navigation aid. Transmission is in the range 1025 to 1150 MHz; receiving is in the range 962 to 1215 MHz; channel spacing is 1 MHz. The interrogator operates in several modes:

- Standby
- Search
- Track
- Scan
- Memory
- Fault
- Self-test.

(a) Location of DME antennas

(b) Location of DME transceiver

Figure 6.6-4 DME equipment.

When the system is first powered up, it enters the standby mode; transmissions are inhibited, the receiver and audio are operative; the DME display is four dashes to indicate **no computed data** (NCD). The receiver monitors pulse-pairs received from any local ground stations. If sufficient pulse-pairs are counted, the interrogator enters the search mode. The transmitter now transmits pulse-pairs and monitors any returns; synchronous pulse-pairs are converted from time into distance and the system enters the track mode. Distance to the navigation aid will now be displayed on the DME indicator (see Figure 6.6-5). The scan mode has two submodes: directed scanning for multiple navigation aid tuning; up to five stations can be scanned in accordance with a predetermined area navigation auto-tuning programme. Alternatively, free scanning occurs for any DME navigation aids within range. If pulse-pairs from any navigation aids are not received after a short period of time (two seconds typical), the interrogator goes into memory mode whereby distance is calculated from the most recently received pulse-pairs. Memory mode expires after a short period of time, typically 10 seconds, or until pulse-pairs are received again. If the system detects any fault conditions, the distance display is blanked out. Self-test causes the system to run through a predetermined sequence causing the indicators to read: blank, dashes (NCD) and 0.0 nm.

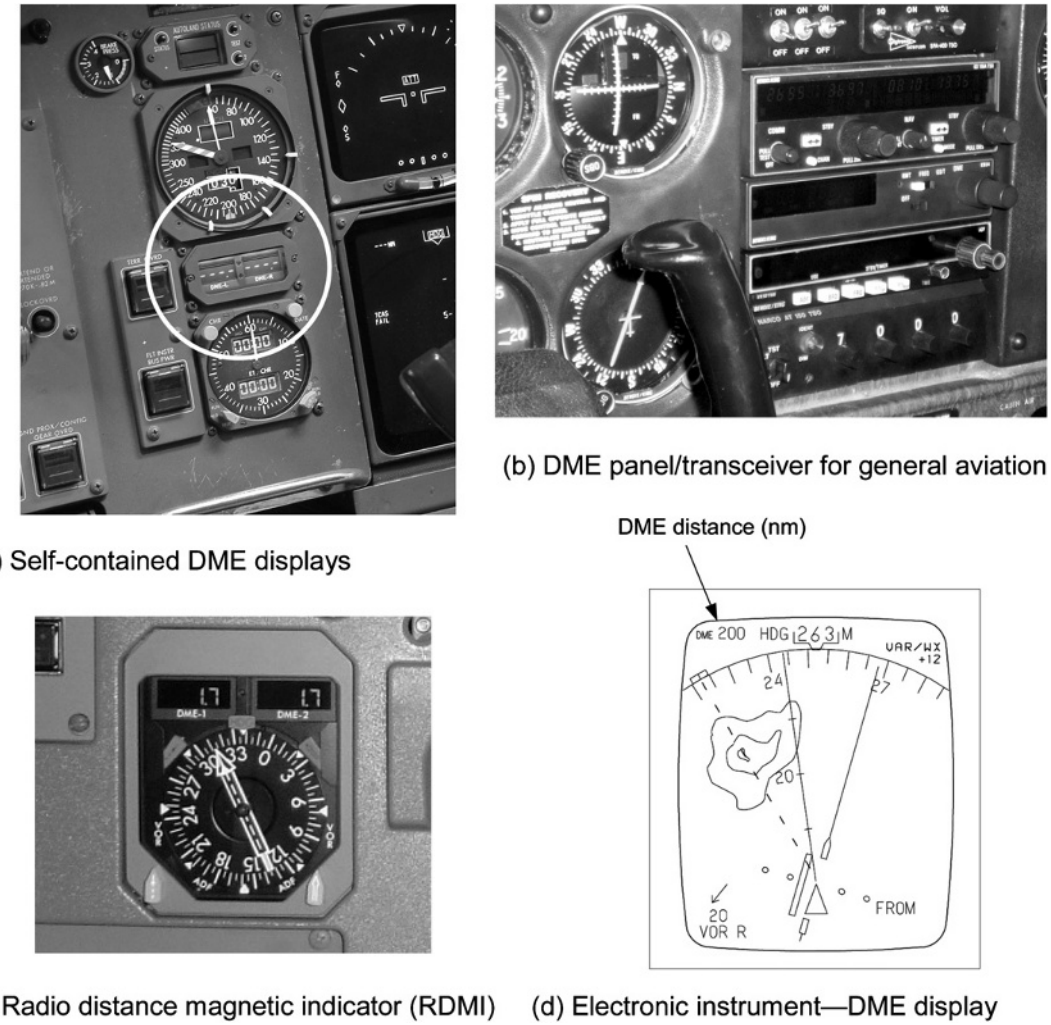


Figure 6.6-5 Various types of DME display.

DME outputs can be displayed in a variety of ways, see Figure 6.6-5. These displays include dedicated readouts, electronic flight instrument systems (EFISS), combined panels/transceivers (for general aviation) and radio distance magnetic indicators (RDMIS). When selecting a co-located VOR-DME navigation aid, the crew only needs to tune into the VOR frequency; the DME frequency is automatically selected.

Key Point

VOR and DME systems operate on different frequencies. When they are co-located, the DME frequency is automatically selected when the pilot tunes into the VOR frequency.

Key Point

When NCD is available this condition is displayed as four dashes.

Test your understanding 6.6.4

List and describe four modes in which a DME interrogator can operate.

6.6.5 En Route Navigation Using Radio Navigation Aids

Basic en route navigation guidance for commercial aircraft can be readily accomplished using co-located VOR and DME systems, thereby providing rho–theta fixes from a single navigation aid. The DME frequency is paired with the

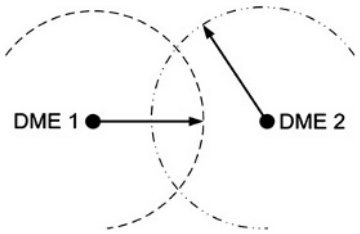


Figure 6.6-6 Ambiguous DME–DME position.

Table 6.6-2 Locations of VOR and DME navigation aids in Switzerland.

Name	Identification code	Type
Corvatsch	CVA	DME
Fribourg	FRI	VOR-DME
Geneva Cointrin	GVA	VOR-DME
Grenchen	GRE	VOR-DME
Hochwald	HOC	VOR-DME
Kloten	KLO	VOR-DME
La Praz	LAP	DME
Montana	MOT	VOR-DME
Passeiry	PAS	VOR-DME
Sion	SIO	VOR-DME
St. Prex	SPR	VOR-DME
Trasadingen	TRA	VOR-DME
Willisau	WIL	VOR-DME
Zurich East	ZUE	VOR-DME

VOR frequency; this means that only the VOR frequency needs to be tuned, the DME frequency is automatically tuned as a result. Alternatively, rho–rho fixes can be established from a pair of DME navigation aids. Note that this produces an ambiguous fix unless another DME is used, see Figure 6.6-6. An example of DME transponder locations and co-located VOR-DME navigation aids in Switzerland is provided in Table 6.6-2.

In the US, a combined rho–theta system was introduced for military aircraft known as **TACAN** (tactical air navigation). This system is a short-range bearing and distance navigation aid operating in the 962–1215 MHz band. TACAN navigation aids (see Figure 6.6-7) are often co-located with VOR navigation aids; these are identified on navigation charts as **'VORTAC'**.

The TACAN navigation aid is essentially a DME transponder (using the same pulse pair and frequency principles as the standard DME) to which directional information has been added; both operate in the same UHF band. An important feature of TACAN is that both distance and bearing are transmitted on the same frequency; this offers the potential for equipment economies. Furthermore, because the system operates at a higher frequency than VOR, the antennas and associated hardware can be made smaller. This has the advantage for military use since the TACAN equipment can be readily transported and operated from ships or other mobile platforms.

When co-located with a VOR navigation aid, military and commercial aircraft can share the VORTAC facility. Referring to Figure 6.6-8, military aircraft obtain their distance and bearing information from the TACAN part of the VORTAC; commercial aircraft obtain their distance information from the TACAN, and bearing information from the VOR part of the TACAN. Reporting points (shown as triangles) based on DME navigation aids, e.g. the VORTAC navigation aid located at Cambrai (CMB), northern France, are illustrated in Figure 6.6-9. The intersecting radials from navigation aids are used to define reporting points for en route navigation. These **reporting points** are given five-letter identification codes associated with their geographic location. For example, the reporting point 'HELEN' (at the top of the chart) is defined by a distance and bearing from the Brussels VOR/DME navigation aid.

TACAN frequencies are specified as channels that are allocated to specific frequencies, e.g. Raleigh–Durham VORTAC in North Carolina, USA, operates on channel 119X. This corresponds to a:

- VOR frequency of 117.2 MHz
- DME interrogation frequency of 1143 MHz
- DME reply frequency of 1206 MHz
- Pulse code of 12 ms.

Figure 6.6-7 TACAN navigation aid.

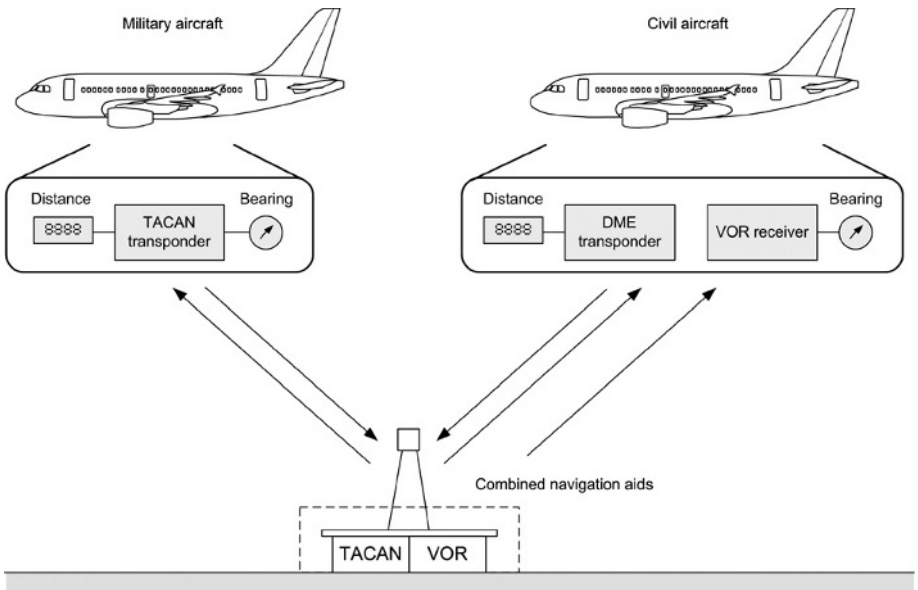


Figure 6.6-8 VORTAC navigation aid and associated aircraft functions.

Note that since DME, VOR and VORTAC navigation aids have to be located on land, the airways' network does not provide a great deal of coverage beyond coastal regions.

Referring to Figure 6.6-10, a combination of VOR, DME and VORTAC stations (see Figure 6.6-11) located in a number of European countries provides a certain amount of navigation guidance in the North Atlantic, Norwegian Sea and North Sea. This diagram assumes a LOS range of approximately 200 nm. The gaps in this radio navigation network can be overcome by the use of alternative navigation systems including: inertial navigation system (INS), Doppler, satellite navigation and Loran-C.

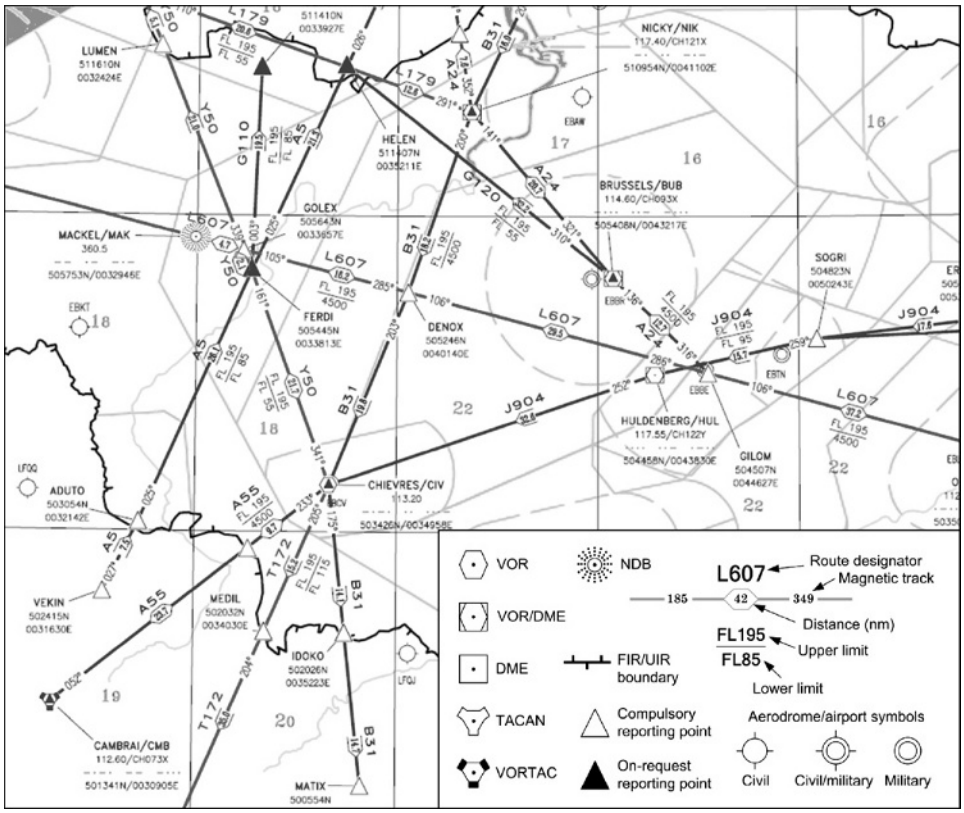


Figure 6.6-9 Reporting points defined by VOR–DME.

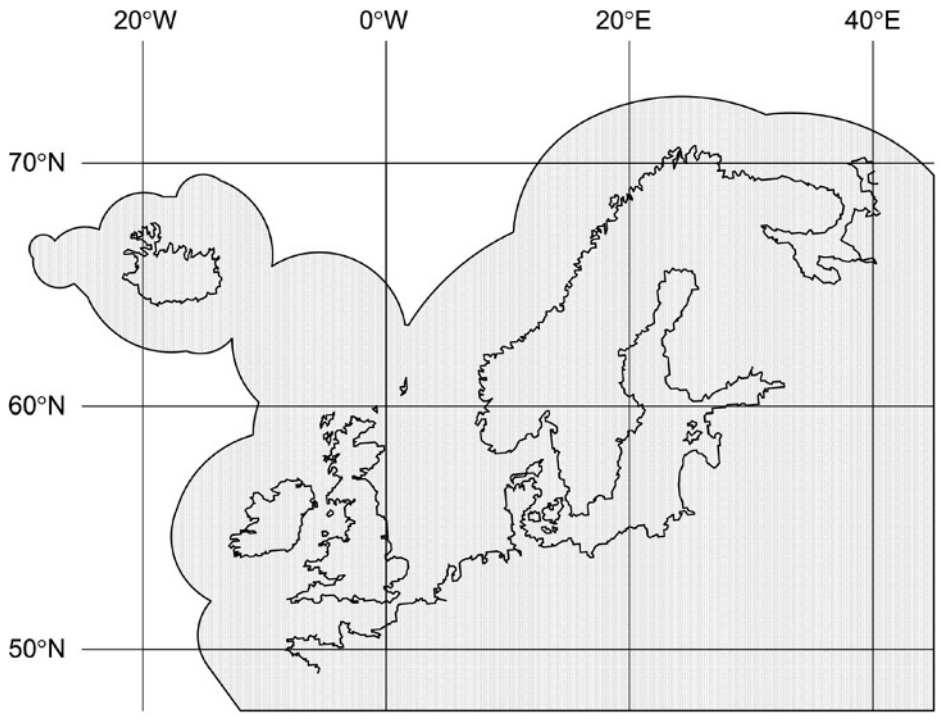


Figure 6.6-10 Approximate max. LOS navigation coverage for northern Europe.

Figure 6.6-11 Typical VOR–DME navigation aid (photo courtesy T. Diamond).

Test your understanding 6.6.5

Explain what is meant by frequency pairing.

Test your understanding 6.6.6

Describe two ways in which DME distance information is displayed.

Test your understanding 6.6.7

DME ground stations could be responding to numerous aircraft; how does the airborne DME system recognise its own signals and reject signals intended for other aircraft?

Test your understanding 6.6.8

What information does an RDMI provide to the crew?

Test your understanding 6.6.9

What type of information does a VORTAC provide?

6.6.6 Multiple Choice Questions

1. DME is based on what type of radar?
   (a) primary
   (b) secondary
   (c) VHF.
2. DME provides the following information to the crew:
   (a) bearing to a navigation aid
   (b) deviation from a selected course
   (c) distance to a navigation aid.
3. When tuned into a VORTAC, commercial aircraft obtain their distance and bearing information from the:
   (a) TACAN and VOR
   (b) DME and VOR
   (c) DME and TACAN.
4. DME signals are transmitted:
   (a) by line of sight
   (b) as ground waves
   (c) as sky waves.

5. An RDMI provides the following information:
   (a) distance and bearing to a navigation aid
   (b) deviation from a selected course
   (c) the frequency of the selected navigation aid.

6. Slant range errors are greatest when the aircraft is flying at:
   (a) high altitudes and close to the navigation aid
   (b) high altitudes and far from the navigation aid
   (c) low altitudes and far from the navigation aid.

7. To select a co-located VOR-DME navigation aid, the crew tunes into the:
   (a) DME frequency
   (b) VOR frequency
   (c) NDB frequency.

8. The DME interrogator is part of the:
   (a) airborne equipment
   (b) DME navigation aid
   (c) VORTAC.

9. The varying interval between pulse-pairs ensures that the interrogator:
   (a) recognises its own pulse-pairs and rejects other signals
   (b) recognises other pulse-pairs and rejects its own signal
   (c) tunes into a VOR station and DME navigation aid.

10. When a DME indicator is receiving NCD, it will display:
    (a) dashes
    (b) zeros
    (c) eights.

11. Using a collocated VOR-DME navigation aid produces what type of position fix?
    (a) rho–rho
    (b) rho–theta
    (c) theta–theta.

12. Distance and bearing signals from a TACAN navigation aid are transmitted on:
    (a) HF
    (b) UHF
    (c) VHF.

13. Using two DME navigation aids provides how many calculated positions?
    (a) two
    (b) one
    (c) three.

Figure 6.6-12 See Question 15.

Figure 6.6-13 See Question 16.

Figure 6.6-14 See Question 17.

14. DME operates in which frequency band?
    (a) UHF
    (b) VHF
    (c) LF/MF.

15. The instrument shown in Figure 6.6-12 is called the:
    (a) RMI
    (b) RDMI
    (c) CDI.

16. Referring to Figure 6.6-13, the display is providing:
    (a) maximum distance
    (b) minimum distance
    (c) no computed data.

17. Referring to Figure 6.6-14, the installation on the right is a DME:
    (a) transponder
    (b) transmitter
    (c) receiver.

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.7 Flight Management Systems

Mike Tooley and David Wyatt

The term 'navigation' can be applied in both the lateral and vertical senses for aircraft applications. Lateral navigation (LNAV) is effectively the area navigation function. Vertical navigation (VNAV) is concerned with optimising the performance of the aircraft to reduce operating costs. This has been traditionally achieved by the flight crew (particularly the flight engineer) making reference to data contained within charts, tables and performance manuals.

Aircraft performance data is based on a number of factors including aircraft weight, altitude and outside air temperature. Since these factors are constantly changing, the task of calculating optimum engine thrust limits, aircraft speed and altitude has gradually been automated with the advent of performance management systems. During the 1980s, lateral navigation and performance management functions were combined into a single system known as the flight management system (FMS). Various tasks previously performed by the crew can now be automated with the intention of reducing crew workload. In this chapter we will review the principles of flight management systems and explore some the key features and benefits.

6.7.1 FMS Overview

The FMS combines area navigation and performance management into a single system. The two primary components of the system are the flight management computer (FMC) and control display unit (CDU). Primary aircraft interfaces with the FMC are the inertial reference system and automatic flight control system, including the autothrottle. Flight management systems were introduced at a time of rising operating costs; the contributing factors to these costs include fuel and time. The cost of fuel is self-evident; the cost of time includes aircraft utilisation, e.g. if the aircraft is being leased on a cost per flying hour basis. Reducing aircraft speed will decrease fuel burn, but this leads to a longer flight time and increased 'cost of time'. Flying faster will reduce the cost of time but increase fuel burn.

Four-dimensional navigation is possible with flight management systems. The aircraft's latitude, longitude, altitude and arrival time requirements can be planned, calculated and subsequently predicted on an ongoing basis. Each airline will have its own financial model in terms of fuel and time costs; the FMS can be customised accordingly and expressed as a **cost-index**; this is entered into system within the range 0–100 to represent the extremes of minimum fuel through to minimum time. In order to perform the key functions of area navigation and performance management, the system interfaces with many other systems on the aircraft.

FMSs were the first examples of integrated multi-mode avionics. On transport category aircraft, the FMS integrates many systems including radio navigation systems, inertial navigation systems, global positioning systems, and centralised maintenance monitoring.

6.7.2 Flight Management Computer System (FMCS)

The two primary components of the system are the FMC and CDU; these are a subset of the FMS referred to as the flight management computer system (FMCS).

6.7.2.1 Flight Management Computer

The FMC contains an operational program, navigation database and performance database. The FMC's **navigation database** (see Table 6.7-1) is a comprehensive version of what has already been discussed in area navigation systems. The **performance database** (PDB) contains a detailed model of the aircraft's aerodynamic characteristics. This includes the aircraft's speed and altitude capabilities together with operating limits for both normal operation and abnormal conditions, e.g. engine failure. Engine parameters are also stored in the PDB; these include fuel flow and thrust models for the type of engine installed on the aircraft. Note that aircraft can be certified to fly with more than one engine type; these are all stored in the PDB.

An important feature of the FMC are the **program pins**. Rather than producing many different FMC software configurations for each aircraft type and each engine combination, one FMC part number can be installed with

Table 6.7-1 Navigation database.

Content	Details
Radio navigation aids	VOR, DME, VORTAC, ADF identification codes, frequencies, locations, elevations
Waypoints	Names and locations, pre-planned within company routes
Airports and runways	Locations, ILS frequencies, runway identifiers, lengths
Standard instrument departures (SIDs)	Published departure procedures including altitude restrictions
Standard terminal arrival routes (STARs)	Published arrival procedures including altitude restrictions
En route airways	Navigation aid references, bearings, distance between navigation aids
Holding patterns	Fix point, inbound course, turn direction
Company routes	A combination of all the above, as specified by the airline

software covering a number of aircraft and engine types in the PDB. The FMC (like most avionic computers) is installed in the equipment rack and connects to the wiring looms via pins/sockets at the rear of the computer.

Program pins are used to select various software options within the computer; these are connections that are made to the connector either to ground, 28 V DC power supply or not connected. Logic circuits inside the computer are thereby set into predetermined configurations depending on how the program pins are configured. For example, a program pin could be connected to ground for one engine type, and set to 28 V DC for another engine type. When the FMC is installed, it effectively recognises which engine type is installed and the relevant engine software is used. The same FMC installed on another aircraft with different engine type will recognise this via the program pin(s) and utilise the relevant engine software.

Certain functions are fixed and cannot be changed, e.g. the aircraft type/model. Other program pins are airline options; examples of these options are the use of metric or imperial units, e.g. Centigrade or Fahrenheit, pounds (×1000) or kilograms (×1000).

6.7.2.2 Control Display Unit

The CDU is the primary interface between the crew and FMC. It is designed such that data entry and displays are in the language used by ATC. The location of a CDU on a typical transport aircraft is shown in Figure 6.7-1. The CDU comprises a variety of features, referring to Figure 6.7-2; these include the

- data display area (typically a cathode ray tube—CRT)
- line-select keys (LSK)
- function and mode keys
- alpha-numeric key pad
- warning annunciators.

The display area is arranged in the form of chapters and pages of a book. When the system is first powered up, the CDU displays the IDENT page, see Figure 6.7-3.

The 'IDENT' page contains basic information as stored in the FMC including aircraft model, engine types, etc. Other pages are accessed from this page on a menu basis using the line-select keys, or directly from one of the function or mode select keys.

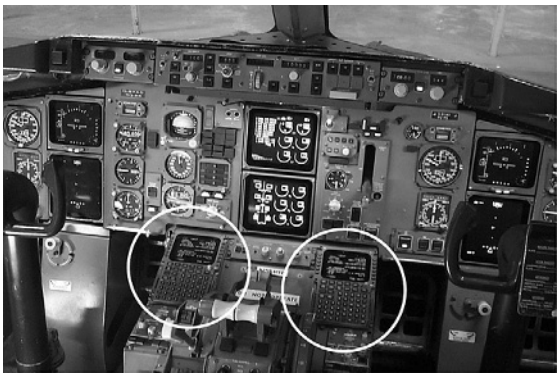


Figure 6.7-1 Location of FMCS control and display unit.

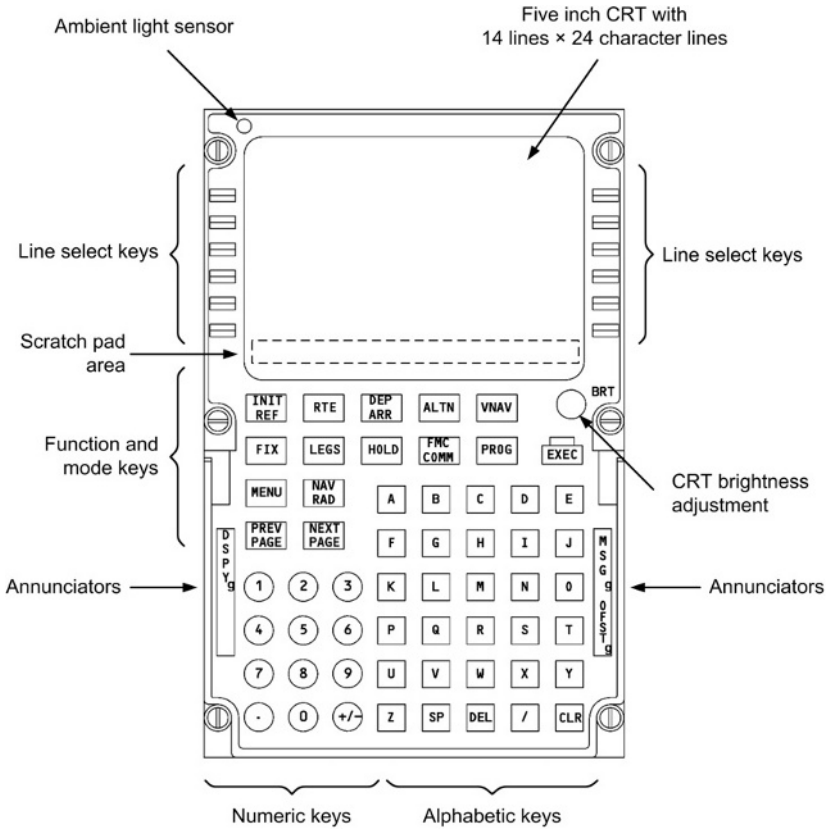


Figure 6.7-2 Location of FMCS control and display unit.

Figure 6.7-3 'IDENT' page displayed on system power-up.

6.7.3 System Initialisation

Before the system can be used for lateral and vertical navigation, the FMC needs some basic initialisation data. Certain information required by the system has to be entered by the crew; other information is stored as a default and can be overwritten by the crew. To simplify the process, information to be entered by the crew is displayed in **box prompts**. Information displayed as default information is displayed as **dash prompts**. There are a number of ways that

individual pages can be accessed, and there is a variety of information on each page. The description below illustrates an initialisation procedure, starting with position initialisation through to performance initialisation, following a logical process. During this initialisation description, we will be making reference to **fields**; these are specific areas on the CDU screen where data is either displayed and/or entered. In the following text, each **line-select key** (LSK) will be referred to by its location: left/right of the display and 1–6 from top to bottom.

6.7.3.1 Position Initialisation

When the system is powered up, the IDENT page is displayed, see Figure 6.7-4 (more details about this page are provided after this section). Pressing LSK-6R, identified POS INIT for **position initialisation**, displays the POS INIT page, see Figure 6.7-5. The information needed at this point (indicated by box prompts) is present position for inertial reference system (IRS) alignment. Position can be entered in a number of ways, but let's assume at this point that we want to load present position by manually keying in latitude and longitude. Using the alphanumeric keys, latitude and longitude are entered via the key pad, entries appear in the bottom of the display (referred to as the **scratch pad**). When this data is confirmed in the scratch pad, LSK-4R (adjacent to the position boxes) is pressed and the scratch pad data replaces the 'Set IRS position' boxes. Present position is automatically transferred to the IRS and the next stage of initialisation is prompted by LSK-6R; this leads to the next section of initialisation for the desired ROUTE.

6.7.3.2 ROUTE Selection

The ROUTE page requires that an origin and destination be entered; these are entered (via the scratch pad) to replace the box prompts adjacent to LSK-1L (origin) and LSK-1R (destination), see Figure 6.7-6(a). Origins and destinations are identified using the International Civil Aviation Organisation (ICAO) four-letter codes, e.g. London

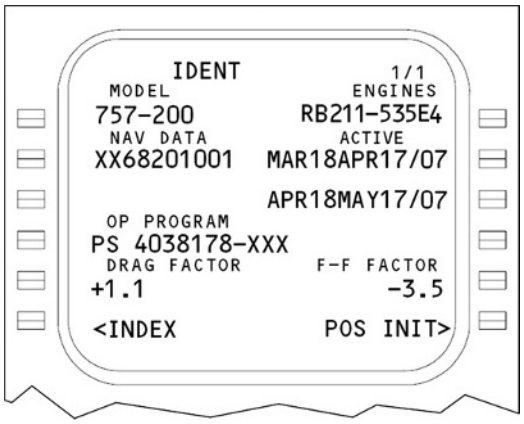


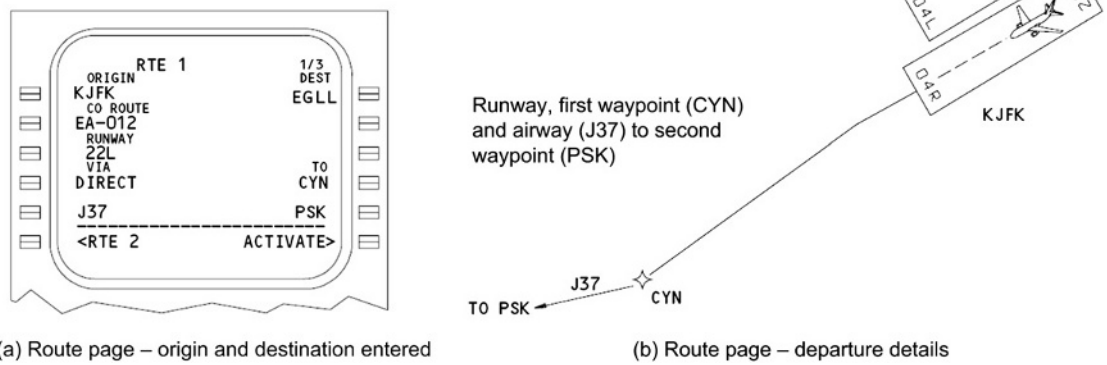
Figure 6.7-4 'IDENT' page—the system is prompting the selection of 'position initialisation' (POS INIT ON LSK-6R).

Figure 6.7-5 Position initialisation ('POS INIT') page.

Heathrow is EGLL, New York Kennedy international airport is coded KJFK. This system is used in preference to the International Air Transport Association (IATA) three-letter codes, e.g. LHR and JFK, since some of these three-letter codes are duplicated for some airfields. Note that most airlines have predetermined company routes, these are stored in the navigation database and can be entered (as a code) via LSK- 2L. There may be more than one route between the origin and destination; when the company route code is entered into an appropriate field, this will automatically enter the origin and destination together with all en route waypoints. Specific departure details, e.g. runway and initial departure fix, can also be contained within the company route as illustrated in Figure 6.7-6(b). Once the route is activated (LSK-6R), the bottom right field changes to PERF INIT for **performance initialisation**.

6.7.3.3 Performance Initialisation

The system requires gross weight (GW) or zero fuel weight (ZFW), reserve fuel, cost index and cruise altitude. Required entries are indicated as before with box prompts, see Figure 6.7-7. Note that since the total fuel onboard (52.3 tonnes in this example) is known by the FMC (via an input from the fuel quantity system), entering ZFW will automatically calculate GW and vice versa. Cost index can be entered manually, or it may be contained within the company route. All other entries on the page are optional; entry of data in these fields will enhance system performance. Once the performance initialisation details are confirmed, the system is ready for operation. Further refinement of the flight profile can be made by entering other details, e.g. take-off settings, standard instrument departures, wind forecasts, etc.



(a) Route page – origin and destination entered

(b) Route page – departure details

Figure 6.7-6 Route page details.

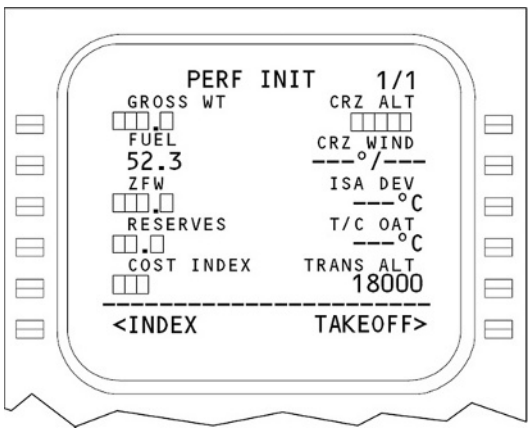


Figure 6.7-7 Performance initialisation ('PERF INIT') page.

6.7.4 FMCS Operation

The flight management computer system (FMCS) calculates key performance data and makes predictions for optimum operation of the aircraft based on the cost-index. We have already reviewed the system initialisation process and this will have given the reader an appreciation of how data is entered and displayed. The detailed operation of an FMS is beyond the scope of this book; however, the key features and benefits of the system will be reviewed via some typical CDU pages. Note that these are described in general terms; aircraft types vary and updated systems are introduced on a periodic basis. CDU pages can be accessed at any time as required by the crew; some pages can be accessed via the line-select keys as described in section 6.7.3; some pages are accessed via function/mode keys. The observant reader may have already noticed that in the top right of each CDU page is an indication of how many subpages are available per selected function.

Key Point

The FMS comprises the following subsystems: FMCS, AFCS and IRS.

Key Point

The page automatically displayed upon FMC power-up is the identification page; this confirms that the FMC has passed a sequence of self-tests.

Key Point

Required entries into the CDU are indicated by box prompts; optional entries are indicated by dashed line prompts.

Key Point

To define the destination airport on the FMC route page requires entry of the airfield's four-character identifier.

6.7.4.1 Identification Page ('IDENT')

This page is automatically displayed upon powerup; aside from displaying a familiar page each time the system is used, this also serves as confirmation that the FMC has passed a sequence of built-in test equipment (BITE) self-tests including: memory device checks, interface checks, program pin configuration, power supplies, software configurations and microprocessor operation. Information displayed on this page includes aircraft and engine types, navigation database references and the operational program number. By reference to the relevant aircraft documentation, one FMC part number could be fitted to a number of different aircraft types. Each aircraft type will have different aerodynamic characteristics and these differences will be stored in the FMC's memory. The FMC recognises specific aircraft types by program pins contained within the aircraft connector, see Figure 6.7-8. Given aircraft types can be operated with different engine models; these are recognised by using specific program pins. Furthermore, airlines have the option on the units used within the system, e.g. temperature in Centigrade or Fahrenheit, weights in kilograms or pounds, etc. These are also determined by program pins.

The navigation database (NDB) is identified by when it becomes effective, and when it expires. Referring to Figure 6.7-4, the active (current) database is adjacent to LSK-2R. The updated database is adjacent to LSK-3R; this is selected on the changeover date (April 18 in this illustration). A comprehensive range of information is contained in the NDB as detailed in Table 6.7-1; note that this is an indicative list since databases are usually customised for individual airlines. The synergy of integrated avionic systems can be demonstrated by FMC database information also being displayed on the EHSI (Figure 6.7-9 displays a number of airports contained in the database).

6.7.4.2 Progress Page

There are many pages available to the crew for managing and modifying data required by the system depending on circumstances. One of the pages used to monitor key flight information is the progress page, see Figure 6.7-10. By describing the information on the progress pages, the reader will gain an appreciation of the features and benefits of the flight management computer system.

Figure 6.7-8 Program pins located in the computer's connector.

Figure 6.7-9 Airports within the navigation database displayed on the EHSI.

The progress page can be accessed via the PROG key on the CDU. There are no entries required on this page; it is used for information only. The top line of the page displays details for the previous waypoint (CYN) in the active flight plan: name, altitude, actual time of arrival and fuel. The next three lines display details for the active waypoint (ENO), next waypoint (GVE) and final destination (KATL). Details include distance to go (DTG), estimated time of arrival (ETA) and predicted fuel. The fifth line gives selected speed, predicted time and distance to an altitude change point, e.g. top of climb (T/C) as illustrated in Figure 6.7-10. The last line of the page is providing navigation source information. In this case, the FMC selected inertial reference system (number 3) is being updated by two DME navigation aids ENO and MLC; these are being tuned manually and automatically as indicted by the letters M and A next to the navigation aid identifier.

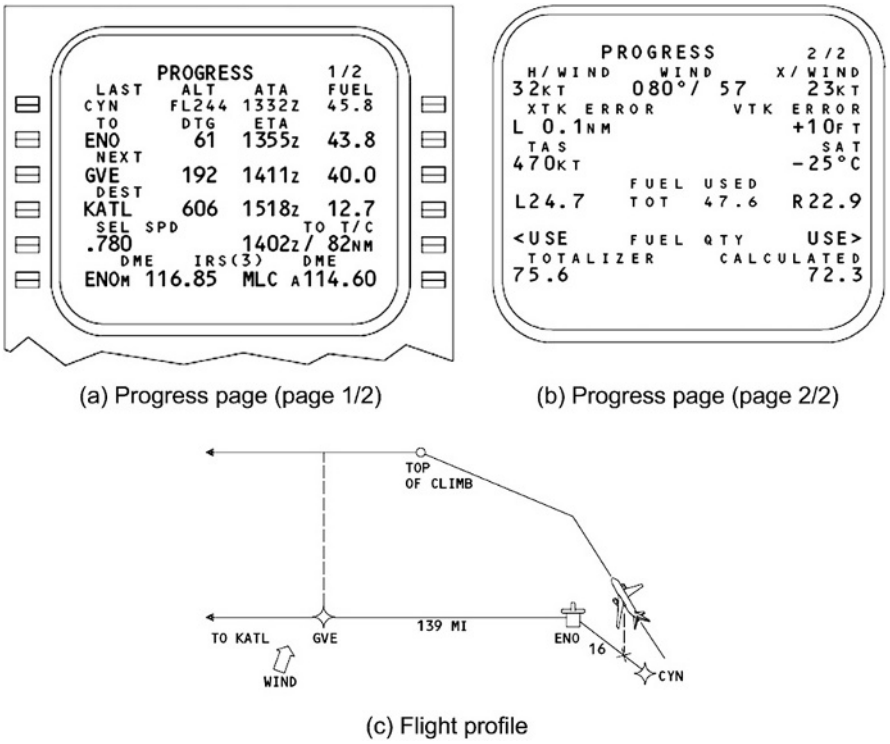


Figure 6.7-10 FMCS progress pages and flight profile.

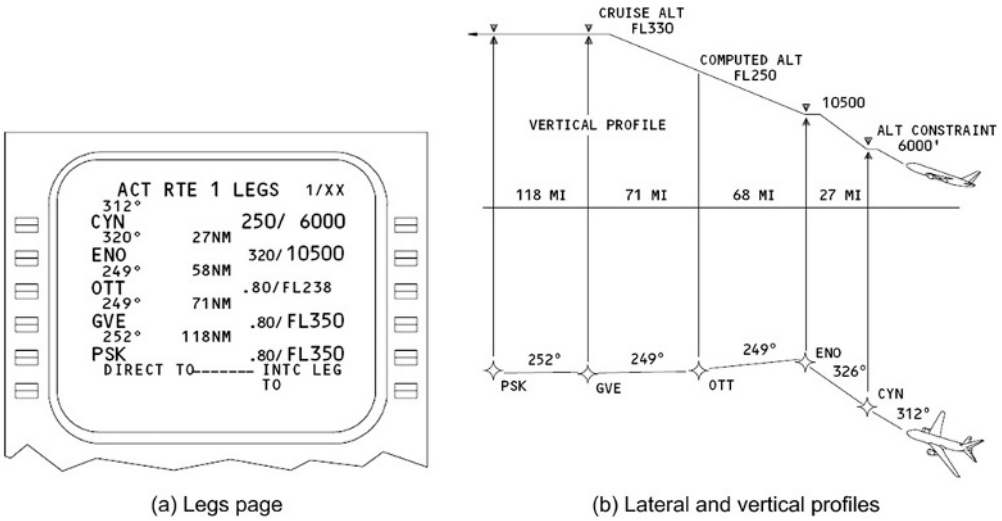


Figure 6.7-11 Legs page and associated flight profiles.

The second progress page contains a variety of useful information, e.g. wind speed and direction (displayed with associated head, tail and cross wind components), cross track (XTK) error, vertical track (VTK) error, true airspeed (TAS), static air temperature (SAT) and various fuel quantity indications.

6.7.4.3 Legs Page

Figure 6.7-11 provides an illustration of how en route lateral and vertical profiles are integrated within the FMC. In this example, the aircraft is flying towards waypoint CYN on a track of 312°. There is an altitude constraint of 6000 feet over CYN, climb speed is 250 knots. This combined lateral and vertical profile is depicted by the tracks,

distances, speeds and altitudes for each waypoint. This level of detail also applies for **standard instrument departures** (SIDs) and **standard terminal arrival procedures** (STARs), see Figures 6.7-12 and 6.7-13.

6.7.4.4 Other CDU Pages

A detailed review of every page available on the CDU is beyond the scope of this book; however, a summary of typical pages is provided in Table 6.7-2. Note that this table is provided for illustration purposes. Aircraft types vary together with the type and model of FMC installed.

Key Point

The highlighted waypoint on the progress page is the active waypoint.

Key Point

Alerting messages require attention from the crew before guided flight can be continued.

6.7.5 FMS Summary

As we have seen, the FMCS performs all the calculations and predictions required to determine the most economical flight profile, either for minimum fuel or for minimum time (or indeed some point in between depending on the operator's financial and commercial models).

When coupled to the automatic flight control system, with vertical and lateral navigation modes engaged, the flight crew act as managers monitoring and entering data as required. Much of the data presented on the CDU is also displayed on the primary flying displays; aircraft with electronic flight instruments have the advantage in that information is displayed with coloured symbols to identify key features of the flight plan, e.g. navigation aids, airfields and descent points.

Test your understanding 6.7.1

(a) What is the meaning of four-dimensional navigation?
(b) How can you confirm that the FMC has passed its power-up test?
(c) What is the significance of box and dash prompts on the CDU?

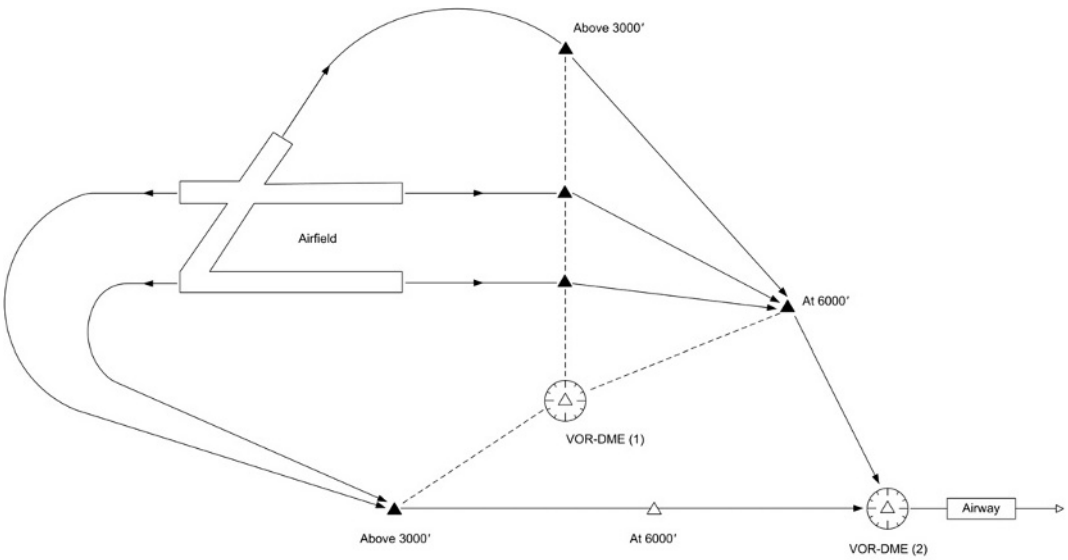
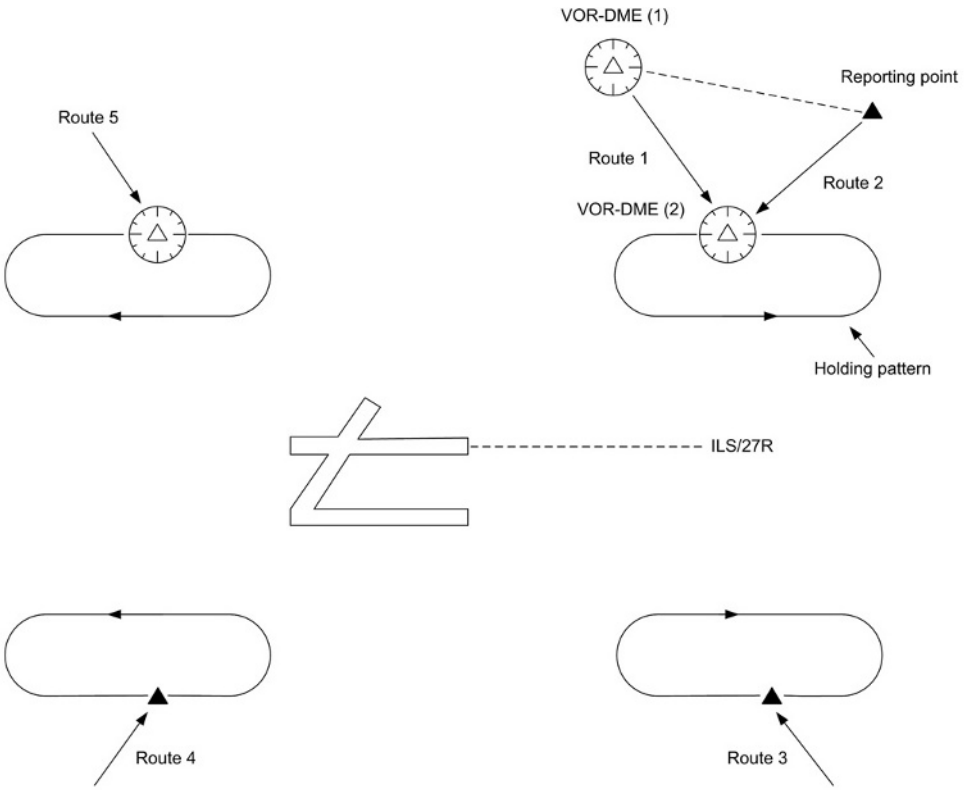


Figure 6.7-12 Standard instrument departure (SID).



Notes:

1. In this illustration, each of the three arrival routes is associated with a navigation aid (VOR-DME) and reporting point (solid triangles)
2. Each arrival route is normally allocated a holding pattern
3. Minimum sector altitudes are published for each route
4. When cleared by ATC, the aircraft would leave the holding pattern and be given a heading to join the ILS for the active runway, e.g. 27R

Figure 6.7-13 Standard terminal arrival routing (STAR).

Test your understanding 6.7.2

What is the purpose of program pins?

Test your understanding 6.7.3

Where would you confirm details of each of the following: navigation database, operational program, aircraft and engine type?

6.7.6 Multiple Choice Questions

1. To define the destination airport on the FMC route page requires entry of the airfield's:
   (a) three-character identifier
   (b) four-character identifier
   (c) latitude and longitude.

2. The page automatically displayed upon FMC power-up is the:
   (a) identification page
   (b) navigation database
   (c) position initialisation page.

Table 6.7-2 Typical CDU pages.

Page title	Full title	Purpose of page
IDENT	Identification	Verifies aircraft model, active database, operational program number, engine type(s)
POS INIT	Position initialisation	Present position required by entering data using one of three methods: latitude/longitude coordinates via the keypad, line selection of last position, line selection of departure gate coordinates
RTE	Route	Entry of route details, either by company route code, or manual construction
CLB	Climb	Selection of desired climb mode, e.g. economy, maximum rate, maximum angle, selected speed, engine out
CRZ	Cruise	Selection of desired cruise mode, e.g. economy, long-range cruise, engine out, selected speed
DES	Descent	Selection of desired descent mode, e.g. economy, selected speed
DIR INTC	Direct intercept	Used to select a waypoint that will be flown directly towards from the present position
RTE LEGS	Route legs	Used for confirming and modifying en route details, e.g. waypoint identification, course and distance to waypoints, speed and altitude constraints (see Figure 6.7-11)
DEP ARR	Departure/Arrival	Provides access to the navigation database for origin or destination SIDS, STARS and specific runways
RTE HOLD	Holding pattern	Review or revision of holding pattern details, e.g. fix point, turn direction, inbound course, leg time and target speed
PROG	Progress	In-flight status of progress along route (see separate notes provided)
N1 LIMIT	N1 limit	The N1 limit is automatically selected and controlled by the FMC. This page provides a range of manually selected N1 limit options including go-around, maximum continuous, climb and cruise
FIX INFO	Fix	Used to create fix points on the current flight leg from known waypoints using radials and distances from the waypoint

3. Program pins are defined by the:
   (a) FMC operational program
   (b) navigation database
   (c) aircraft wiring at the FMC connector.

4. Information entered into the CDU scratch pad is displayed in the:
   (a) lowest section of the display
   (b) box prompts
   (c) dash prompts.

5. Minimum flight time would be achieved with a cost-index of:
   (a) 0
   (b) 100
   (c) 50.

6. Aircraft and engine type can be confirmed on the:
   (a) progress page
   (b) identification page
   (c) position initialisation page.

7. The use of metric/imperial units is determined by:
   (a) the part number of the FMC
   (b) program pins
   (c) dashed line entries.

8. Required entries into the CDU are indicated by:
   (a) box prompts
   (b) dashed lines
   (c) highlighted text.

9. 'Not in database' is an example of an:
   (a) alert message
   (b) advisory message
   (c) active waypoint.

10. Display of the identification page after powerup confirms the:
    (a) IRS is aligned
    (b) navigation sources in use
    (c) FMC has passed its BITE check.

11. The FMC recognises specific aircraft types by:
    (a) CDU entry
    (b) program pins
    (c) the navigation database.

12. SIDs in the navigation database refer to:
    (a) arrivals
    (b) en route navigation
    (c) departures.

13. The EXEC key lights up when:
    (a) data is entered for initialisation/changes
    (b) advisory messages are displayed
    (c) incorrect data has been entered.

14. Alerting messages require attention from the crew:
    (a) before guided flight can be continued
    (b) when time is available
    (c) at the completion of the flight.

15. The highlighted waypoint on the progress page is the:
    (a) previous waypoint
    (b) active waypoint
    (c) destination.

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.8 Air Traffic Control System

Mike Tooley and David Wyatt

The purpose of the air traffic control (ATC) system is to enable ground controllers to maintain safe separation of aircraft, both on the ground and in the air. In addition to this, the controllers are managing the flow of traffic in a given airspace. The system is based on secondary surveillance radar (SSR) facilities located at strategic sites, on or near airfields. Ground controllers use the SSR system to identify individual aircraft on their screens. The basic system is referred to as the ATC radar beacon system (ATCRBS); the updated version, Mode S, improves surveillance with a high integrity digital data link. This chapter describes the various modes of ATC operation and concludes with a review of automatic dependent surveillance-broadcast (ADS-B) and communications, navigation and surveillance/air traffic management (CNS/ATM).

6.8.1 ATC Overview

We have seen examples of primary and secondary radar systems in previous chapters. To reiterate, with primary radar, high energy is directed via an antenna to **illuminate** a 'target'; this target could be an aircraft, the ground or water droplets in a cloud. In the case of ATC primary radar, the energy is reflected from the aircraft's body to provide range and azimuth measurements. ATC's primary radar system places the target(s) on a **plan position indicator** (PPI). Primary surveillance radar (PSR), see Figure 6.8-1, has its disadvantages, one of which is that the amount of energy being transmitted is very large compared with the amount of energy reflected from the target. Small targets, or those with poor relecting surfaces, could further reduce the reflected energy. Natural and man-made obstacles such as mountains and wind farms also shield the radar signals. SSR transmits a specific low energy signal (the interrogation) to a known target. This signal is analysed by a **transponder** and a new (or secondary) signal, i.e. not a reflected signal, is sent back (the reply) to the origin, see Figure 6.8-2. Secondary radar was developed during the Second World War to differentiate between friendly aircraft and ships: this system was called Identification Friend or Foe (IFF). In the ATC system, the primary and secondary radar antennas are mounted on the same rotating assembly, thereby providing a coordinated system. The complete system is illustrated in Figure 6.8-3.

The ATC system operates on two frequencies within the L-band of radar:

- **Interrogation** codes on a 1030 MHz carrier wave
- **Reply** codes on a 1090 MHz carrier wave.

The primary radar system provides a single icon per aircraft on the ATC controller's display; this means that each icon will look similar, depending upon the amount of reflected energy. As a consequence, an aircraft would have to change direction in order for it to be uniquely identified. By implementing the SSR transponder system, each icon can be identified via a unique **four-digit code** (allocated by ATC for each flight). Using SSR also means that the effects of clutter (from trees, buildings and hills, etc.) are not displayed on the controller's screen. With an uncluttered screen, and each aircraft readily identified, more aircraft can be allowed into the controlled airspace. The combined PSR/SSR system is illustrated in Figure 6.8-4. Developments of the ATC transponder system have provided additional functionality allowing details such as flight number and altitude to be displayed on the controller's screen. Emergency codes can be sent in the event of radio failure or hijacking. The reader will appreciate that it is essential for an aircraft operating in controlled airspace to be equipped with an ATC transponder.

6.8.2 ATC Transponder Modes

SSR systems have been developed for both military and commercial aircraft applications; a summary of commercial aircraft modes is as follows:

Mode A: In this transponder system, the pilot selects the four-digit code on the ATC control panel prior to each flight. The SSR system confirms this aircraft's azimuth on the controller's screen with an icon confirming that the aircraft is equipped with a transponder. If the controller needs to distinguish between two aircraft in close proximity an identity code will be requested; the pilot pushes a switch on his ATC control panel, and this highlights the icon on the controller's screen. Since each aircraft is allocated with a unique code, only one icon per aircraft will be

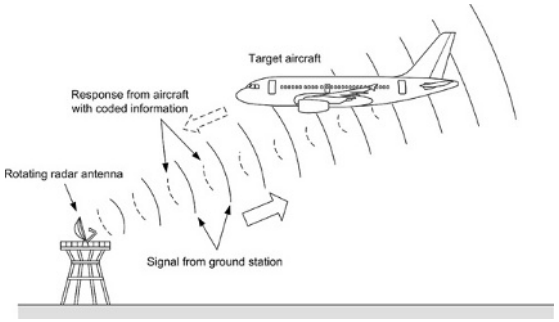


Figure 6.8-1 Primary surveillance radar (PSR).

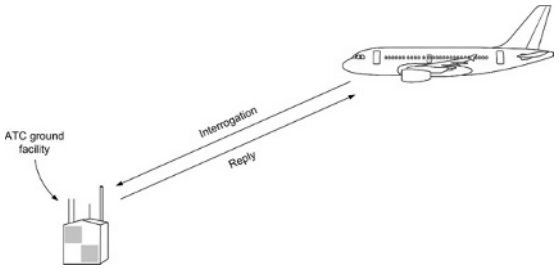


Figure 6.8-2 Secondary surveillance radar (SSR).

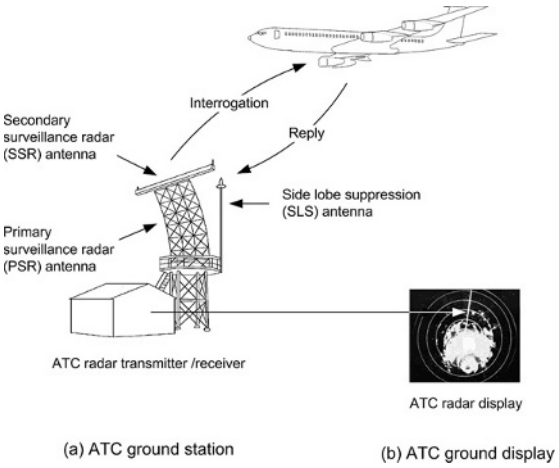


Figure 6.8-3 ATC system overview.

highlighted; this unique identification is referred to as a **squawk** code. Each of the four digits ranges from 0 to 7, these are then coded as octal numbers for use by the transponder. (This system is called Mode 3 for military users.)

Mode C: Azimuth is now augmented by pressure altitude; this is displayed on the controller's screen, adjacent to the aircraft icon thereby providing three-dimensional information. Altitude can be taken from the pilot's altimeter from an encoder that sends parallel data (in Gillham/Gray code) to the transponder. This coded data is in 100-foot increments. Aircraft with air data computers will send altitude to the transponder in serial data form, typically ARINC 429.

Mode S (select): In addition to the basic identification and altitude information, Mode S includes a data linking capability to provide a cooperative surveillance and communication system. Aircraft equipped with Mode S transponders allow specific aircraft to be interrogated; this increases the efficiency of the ATC resources. To illustrate this point, when aircraft equipped with Mode A or C transponders are interrogated, all aircraft with this type of

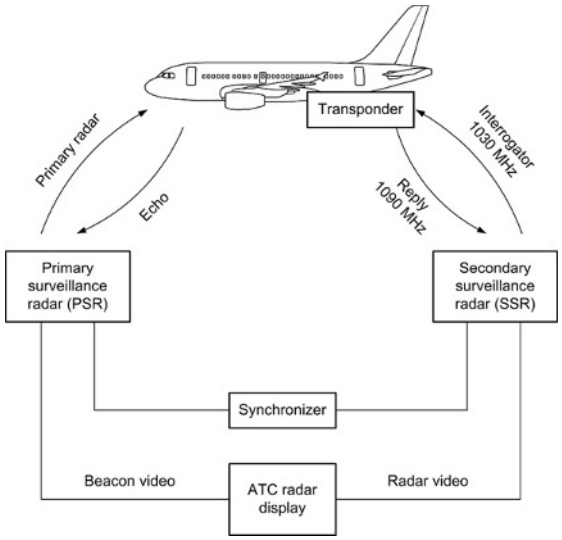


Figure 6.8-4 Combined PSR and SSR.

transponder will send replies back to the ground station. This exchange occurs each time an interrogation signal is transmitted. Imagine a room full of people; the question is asked: 'please state your name and location in the room'. The person asking the question could become overwhelmed with the replies. If the question was posed in a different way, i.e. on a **selective** basis: 'Mike, where are you?' followed by: 'David, where are you?', the replies are only given by the person being addressed. The Mode S system has a number of advantages:

- Increased traffic densities
- Higher data integrity
- Efficient use of the RF spectrum
- Reduced RF congestion
- Alleviation of Mode A and C code shortages
- Reduced workload for ground controllers
- Additional aircraft parameters available to the ground controller.

Mode S transponders only send a reply to the first interrogation signal; the ground station logs this aircraft's address code for future reference. Mode S provides additional surveillance capability into controlled airspace; this is being introduced on a progressive basis. Aircraft equipped with Mode S transponders are also able to communicate directly with the Mode S transponders fitted to other aircraft; this is the basis of the traffic alert and collision avoidance system (TCAS) and will be described in the next chapter. (Note that Modes B and D are not used by commercial aircraft.)

6.8.3 Airborne Equipment

Commercial transport aircraft are installed with two ATC antennas, a control panel and two transponders as illustrated in Figure 6.8-5. Since the ATC system and distance measuring equipment (DME) operates in the same frequency range, a mutual suppression circuit is utilised to prevent simultaneous transmissions.

6.8.3.1 Control Panel

This is often a combined ATC and traffic alert and collision avoidance system control (TCAS) panel, see Figure 6.8-6 (refer to the next chapter for detailed operation of TCAS). The four-digit aircraft identification code is selected by either rotary switches or push buttons, and displayed in a window. Altitude reporting for Mode C transponders can be selected on or off. When requested by ATC, a momentary make switch is pressed; this transmits the selected code for a period of approximately 15 to 20 seconds. Table 6.8-1 illustrates the codes that are used in emergency situations.

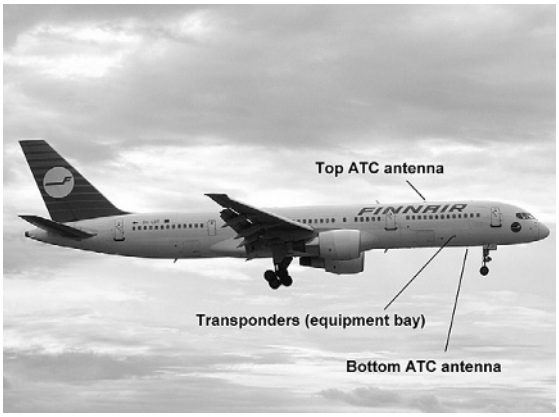


Figure 6.8-5 Airborne equipment.

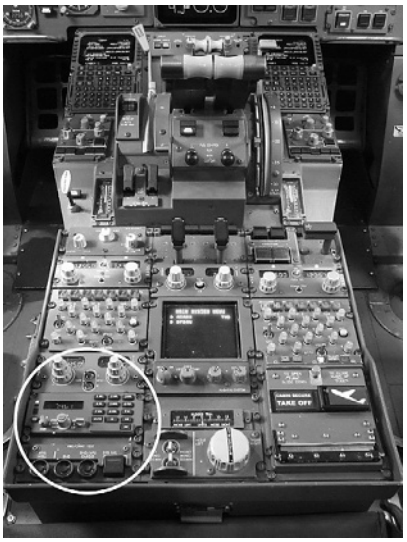


Figure 6.8-6(a) ATC control panel location.

Figure 6.8-6(b) ATC control panel.

6.8.3.2 Transponder

The aircraft transponder (Figure 6.8-7) provides the link between the aircraft and ground stations. General aviation products have a combined panel and transponder to save space and weight. These can be Mode S capable for IFR operations.

The ground station SSR antenna is mounted on the antenna of the primary radar surveillance system, thereby rotating synchronously with the primary returns. The airborne transponder receives interrogation codes on a 1030 MHz carrier wave from the ground station via one of two antennas located on the airframe. These signals are then amplified, demodulated and decoded in the transponder. The aircraft reply is coded, amplified and modulated as an

Table 6.8-1 Emergency ATC transponder codes.

Code	Meaning
7700	General air emergency
7600	Loss of radio
7500	Hijacking

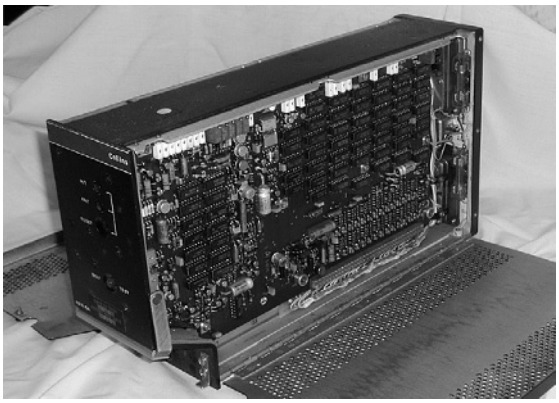


Figure 6.8-7 ATC transponder.

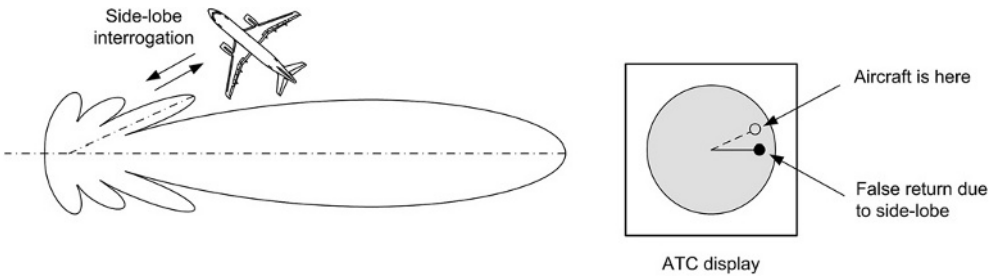


Figure 6.8-8 False returns from side-lobes.

RF transmission reply code on a 1090 MHz carrier wave. If the transponder is interrogated by a TCAS II-equipped aircraft, it will select the appropriate antenna to transmit the reply. This technique is called **antenna diversity**; this enhances visibility with TCAS-equipped aircraft flying above the host aircraft.

6.8.4 System Operation

Although SSR has many advantages over primary radar, the smaller antenna's radiation pattern contains substantial **side-lobes**. These side-lobes can generate false returns (Figure 6.8-8), and so a method of coding the interrogation signals via pulse techniques is employed. The solution is to superimpose an omnidirectional pattern from a second antenna onto the directional beam. Suppressing these side-lobes is discussed in the following sections on interrogations and replies.

6.8.4.1 Mode A and C Interrogation

Interrogation is based on a three-pulse format as illustrated in Figure 6.8-9; each pulse is 0.8 μs wide. Two pulses (P1 and P3) are transmitted on the rotating antenna thereby producing a **directional** signal. A third pulse (P2) is transmitted on the fixed antenna that radiates an **omnidirectional** signal. The purpose of the P2 pulse is described in the Mode A reply section. Referring to Figure 6.8-9, P1 and P2 have a 2 μs interval; P3 is transmitted at an interval of 8 μs for Mode A and 21 μs for Mode C interrogations (see Figure 6.8-10). This spacing between P1 and P3 therefore

determines the type of interrogation signal (Mode A or C). The pulse repetition frequency (PRF) of interrogation signals is unique to each ground station; a typical PRF is 1200 interrogation signals per second. Replies are sent by the aircraft at the same PRF.

6.8.4.2 Mode A Reply

A given aircraft's transponder will receive maximum signal strength each time the ground station's directional beam passes, i.e. once per revolution. Since P2 is transmitted from the fixed omnidirectional antenna, it is received with constant signal strength; but with lower amplitude than P1/P3. When the aircraft's transponder receives the maximum P1/P3 signal strength, i.e. when the rotating antenna is directed at the aircraft, they are received at higher amplitude than P2. Referring to Figure 6.8-11, an aircraft not within the main-lobe of the directional beam would receive a P2 pulse from the omnidirectional antenna at higher amplitude than the P1/P2 pulses. The transponder recognises this as a side-lobe signal and suppresses any replies until 25 to 45 μs after P2 is received. This is called **side-lobe suppression** (SLS), a technique ensuring that only the main lobe of the rotating antenna is being replied to and not a side-lobe. The physical arrangement and antenna patterns are illustrated in Figures 6.8-3 and 6.8-11 respectively.

The Mode A reply is the ATC code allocated to that flight, formed into a series of pulses. This reply is framed between two pulses (F1 and F2) that have a time interval of 20.3 μs as illustrated in Figure 6.8-12. Data to be transmitted is coded by 12 pulses (plus an unused 'spare' pulse in position X) at 1.45 μs intervals within F1 and F2. The 12 pulses are grouped into four groups of three; each group represents an octal code. Each of the four groups is labelled A, B, C and D; single pulses within the group carry a numerical weighting of 1, 2 and 4. For illustration purposes, Table 6.8-2 shows how group A pulses represent the octal values between 0 and 7.

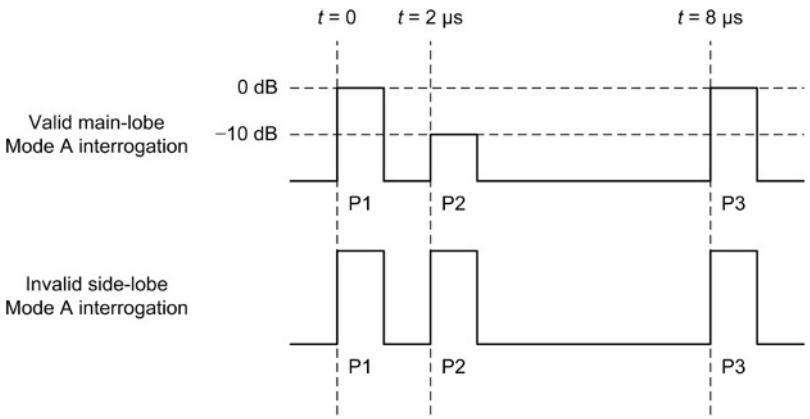


Figure 6.8-9 Interrogation signal validity.

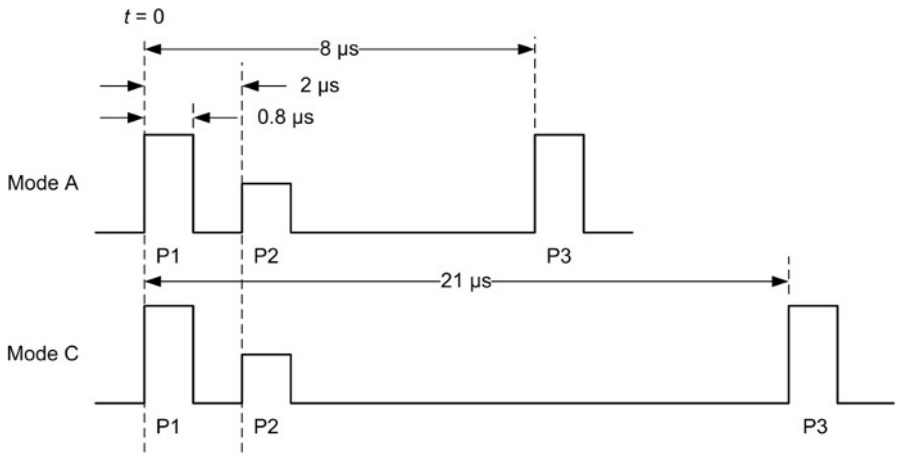


Figure 6.8-10 Pulse format for Mode A and Mode C interrogations.

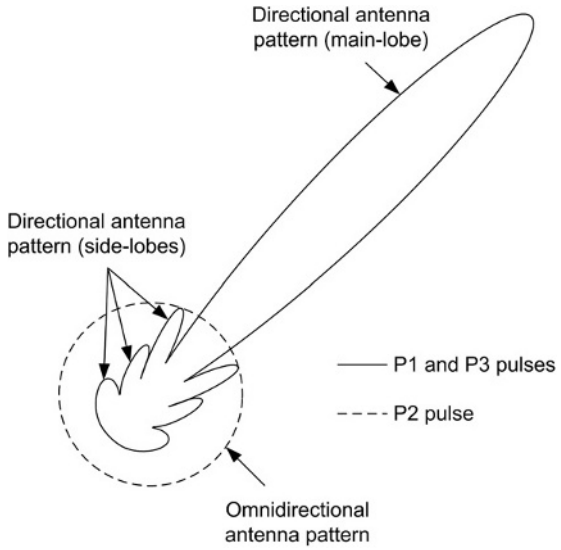
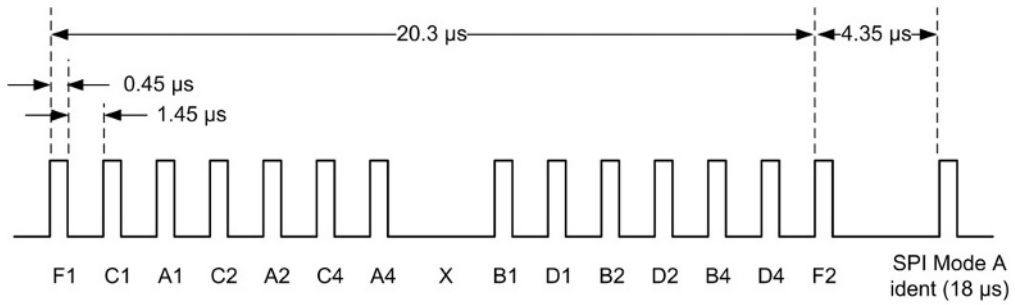


Figure 6.8-11 Side-lobe suppression (SLS).



(c) Reply pulse train

Figure 6.8-12 Mode A/C reply pulse train (ATC code or altitude).

Table 6.8-2 Illustration of Group A pulses.

A1	A2	A3	Octal value
0	0	0	0
0	0	1	1
0	1	0	2
0	1	1	3
1	0	0	4
1	0	1	5
1	1	0	6
1	1	1	7

Note: 0 = no pulse transmitted
1 = pulse transmitted

When a pulse occurs in group A, this represents the value 1, 2 or 4 depending on the position of the pulse. When a pulse is not transmitted in the allocated time frame, this represents a value of zero. With four groups of data, the **octal numbers** between 0000 and 7777_8 can be transmitted; this corresponds to the ATC code allocated to the flight and selected by the crew on their ATC control panel (4096 codes are possible using these four octal digits).

The final part of the aircraft reply is a pulse that is sent after F2; this pulse occurs 4.35 μs after F2, but it is only sent when an 'ident' is requested by ATC. The flight crew send this **special position identity** (SPI) pulse by pressing

a momentary make switch on the control panel. The SPI pulse is sent for a period of 15 to 20 seconds after the switch has been pressed; this highlights the aircraft icon on the controller's screen.

6.8.4.3 Mode C Reply

The aircraft's altitude is encoded by the transponder and transmitted as binary coded octal (BCO) (in 100 foot increments) as described for Mode A replies. The reply will also contain a code representing the aircraft's altitude; this is referenced to standard pressure, 1013.25 mB if the aircraft is above the transition altitude, see Figure 6.8-12. (The transponder sends Mode A and C replies on an alternating basis.) Figure 6.8-13 illustrates how an ATC identification code of 2703 is combined with an altitude code representing 1,900 feet.

Since all SSR transmissions are on the same frequencies (interrogation on 1030 MHz and replies on 1090 MHz), problems can occur when aircraft are within range of two or more ground stations. Several replies could be sent by an aircraft to each ground station that sends an interrogation signal; these undesired replies are known as non-synchronised garble, or **false replies from unsynchronised interrogator transmissions** (FRUITs). Note that **FRUIT** is sometimes written as false replies uncorrelated in time. When interrogations are received simultaneously, the transponder will reply to as many ground stations as possible.

If two or more aircraft are in close proximity, e.g. in a holding pattern, and within the ground station's directional antenna beam width, it is possible that their individual replies overlap at the ground station's computer, see Figure 6.8-14. The situation where replies are received from two or more interrogators answering the same interrogation is referred to as **synchronized garbling**. To resolve this, the controller can request the flight crew on a specific aircraft to provide an ident pulse. The problem is that ATCRBS for Modes A/C requires many interrogations to determine the position (**range and azimuth**) of an aircraft; this requires increased capacity of the ATCRBS infrastructure. The increasing density of aircraft within a given air space leads to false replies as the ground station saturates with garbling conditions. The solution to this is the Mode S (select) system.

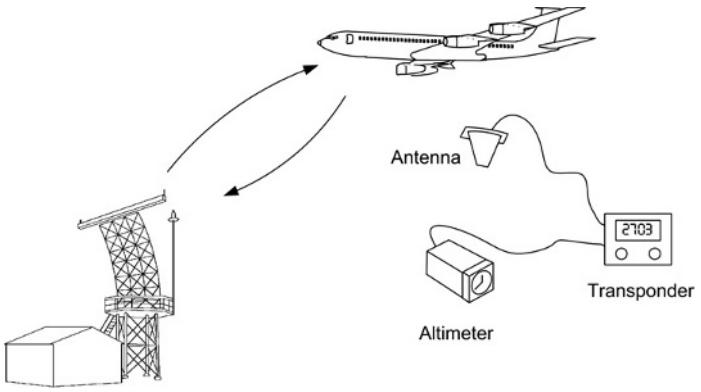


Figure 6.8-13(a) Mode C principles.

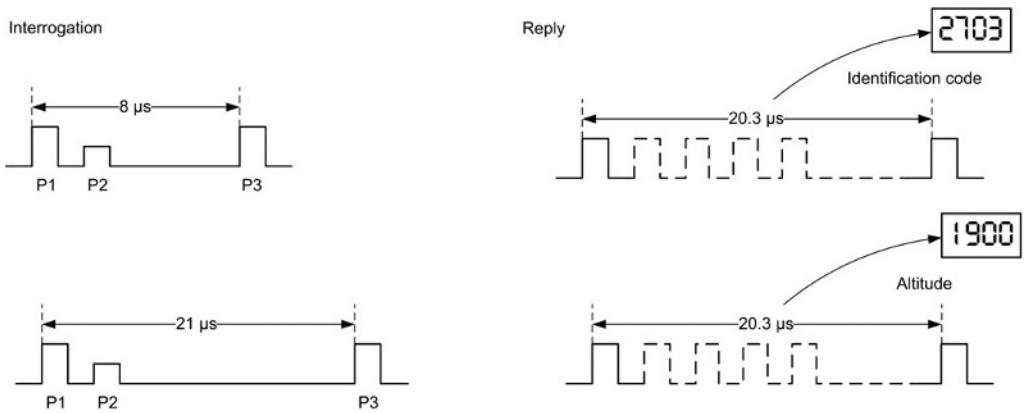


Figure 6.8-13(b) Mode C replies.

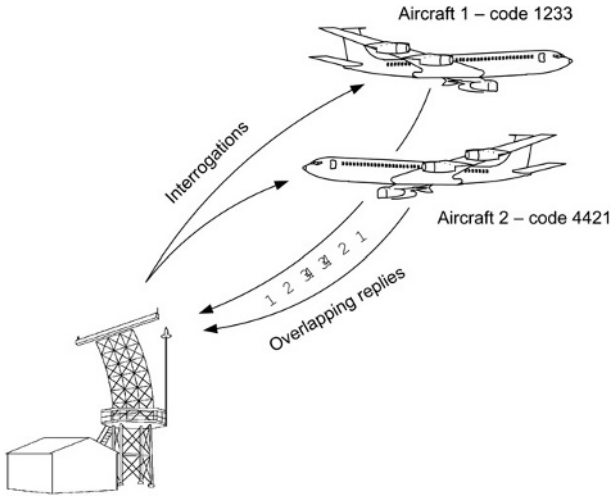


Figure 6.8-14 Synchronized garbling (from two aircraft within the antenna beamwidth).

6.8.4.4 Mode S Operation

Although Mode S communication is very different to that of Modes A/C, both types of equipment operate on the same frequencies. The system is two-way compatible in that aircraft equipped with Mode A/C transponders will respond with ident and altitude data if interrogated by a Mode S ground station. The principles of Mode S are illustrated in Figure 6.8-15. Individual interrogations are sent to specific aircraft; only the transponder on this aircraft sends a reply. This reply contains additional information, e.g. selected altitude and flight number. Directional and omnidirectional beam patterns are transmitted as illustrated in Figure 6.8-16. Unlike ATCRBS, Mode S uses a **monopulse** SSR; this reduces the number of interrogations required to track a target. In theory, monopulse radar only requires one reply to determine the target's azimuth (direction and range).

Two interrogation uplink formats (UFs) are transmitted as shown in Figure 6.8-17; these are the **all-call** and **roll-call** interrogations. The two interrogations are transmitted on an alternating basis and differentiated by the width of a P4 pulse; this is either 0.8 μs or 1.6 μs. The shorter pulse is used to solicit replies from Mode A/C transponders; they reply with their ATC code and altitude as before. Mode S transponders will not reply to this interrogation. When the P4 pulse is 1.6 μs, Mode S-equipped aircraft will reply with their unique address. These replies are stored by the Mode S system as unique identifiers for each specific aircraft, see Figure 6.8-18(b).

The Mode S discrete addressed interrogation UF is shown in Figure 6.8-17(b). Pulse P1 and P2 both have the same amplitude and are part of the directional antenna's mainlobe. This pair appears as suppression pulses to Mode A/C

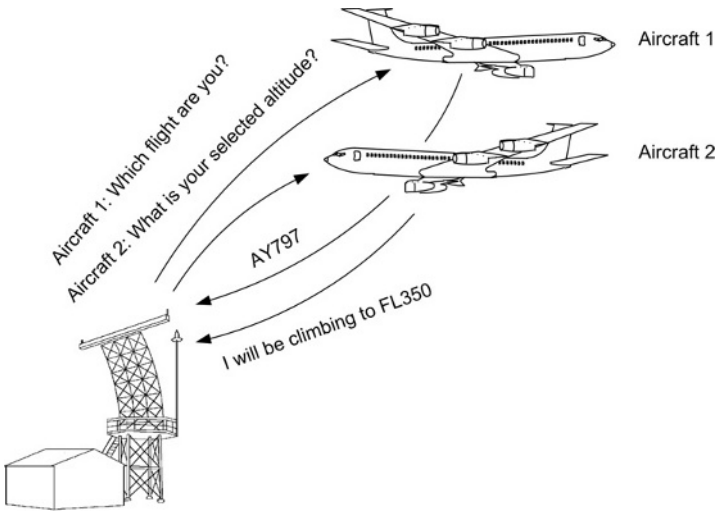


Figure 6.8-15 Principles of Mode S.

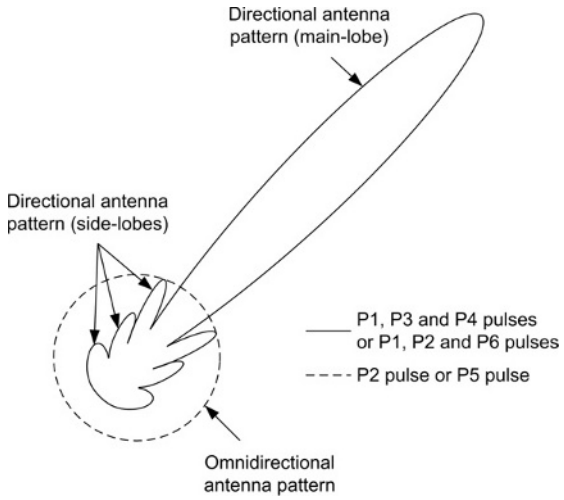


Figure 6.8-16 Mode S antenna pattern.

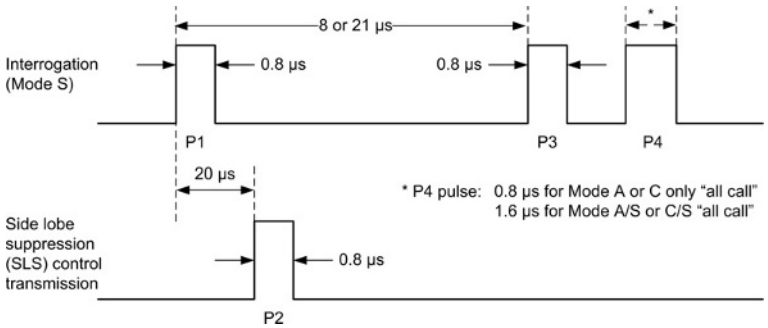


Figure 6.8-17(a) Mode S 'all-call' interrogation UF.

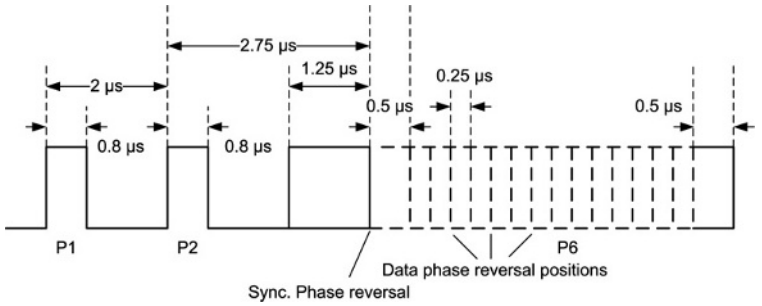


Figure 6.8-17(b) Mode S discrete interrogation signal UF.

transponders, so they do not reply. Mode S transponders then seek the start of the P6 data pulses; this is formed by a pattern of phase reversals that form a series of logic 1/0. **Phaseshift keying** (PSK) is a modulation technique that shifts the phase by −90 degrees for a logic 1, and +90 degrees for a logic 0. Each data pulse's duration is 0.25 µs; the pulse's phase is sampled at these intervals. A reference pulse of 1.25 µs duration is used to indicate the start of the data word. The word length of P6 (56 or 112 bits) depends on the transponder type.

The Mode S reply is sent via the 1090 MHz carrier wave, as illustrated in Figure 6.8-18(a); this contains a four-pulse preamble, starting with two pairs of synchronising pulses followed by a block of data pulses (either 56 or 112 bit blocks). Using **pulse position modulation** (PPM), each data bit is allocated a 1 µs time interval, divided into two halves. If the first half of this interval contains a pulse, this represents logic 1; if the second half of the interval contains a pulse, this represents logic 0. Note that both states are indicated by the presence of a pulse.

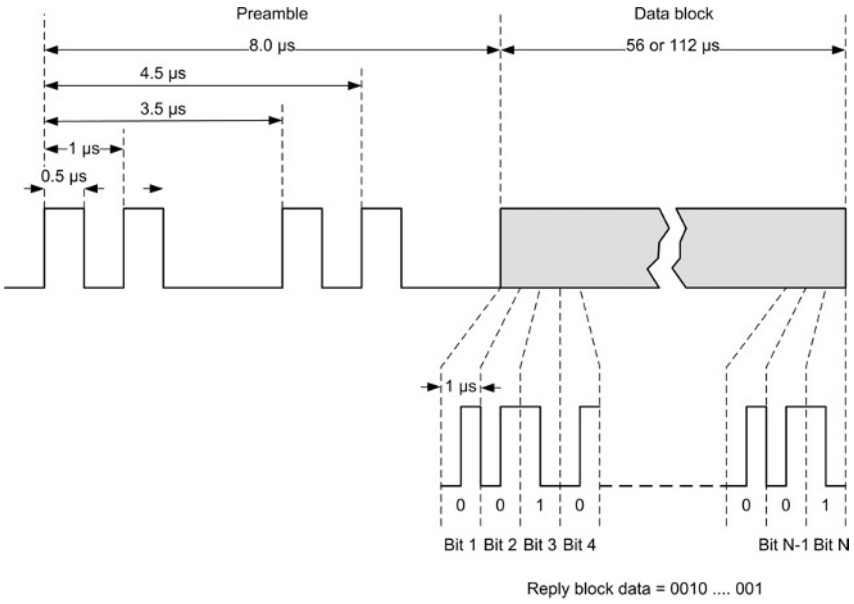


Figure 6.8-18(a) Mode S reply downlink format.

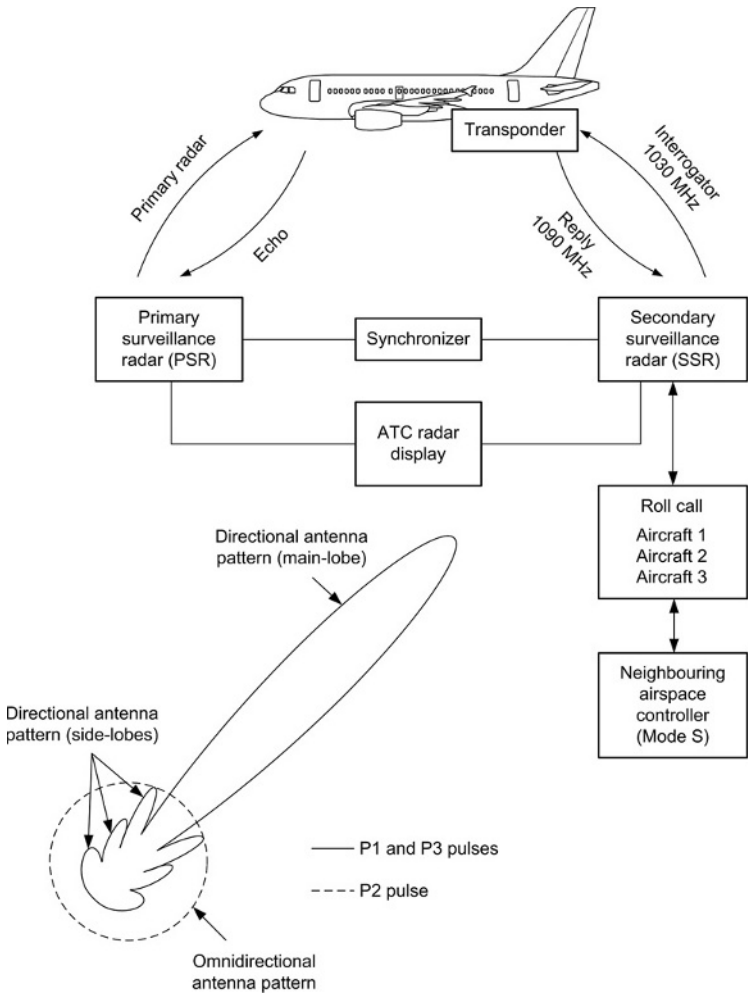


Figure 6.8-18(b) Mode S 'roll-call' function.

Each Mode S-equipped aircraft has a unique address allocated to it by ICAO via the individual national registration authorities; the **aircraft address** (AA) is a 24-bit code that cannot be changed. Each national authority allocates a header code within the 24 bits, e.g. the UK code is 01000. A 24-bit code of all zeros is not valid; all ones are used for the all-call interrogation.

Individual interrogators are also coded, and this is a key feature of Mode S. **Interrogator codes** (ICs) comprise 15 interrogator identifier (II) and 63 surveillance identifier (SI) codes. The use of these interrogator codes ensures unambiguous data exchange between interrogators and transponders. The Mode S all-call request acquires Mode S-equipped aircraft entering a given airspace. The aircraft transponder replies with its unique aircraft address (AA). Lock-out protocols are used to suppress further replies by aircraft to subsequent requests from the same interrogator. Transponders are thereby conditioned for replying to a specific interrogator identified by the ICs; the transponder will subsequently ignore requests from other interrogations. Following the all-call lock-out, the interrogator will address individual aircraft transponders on a **selective** basis. Only the selected aircraft transponder will reply; interrogators will ignore replies not intended for them.

Mode S is being introduced on a progressive basis via a transitional phase of equipment standards. The two standards are: **elementary surveillance** (ELS) and **enhanced surveillance** (EHS). Data sent by each of these two standards is shown in Tables 6.8-3 and 6.8-4 respectively.

6.8.4.5 Mode S Transponders

Aircraft are being equipped on a progressive basis with transponders that meet the necessary functionality for the category of operations. There are four levels of transponder; these levels specify the datalink capability as detailed in Table 6.8-5.

Mode S transponders are installed with 255 data registers, each containing 56 bits. These registers are automatically loaded with specific aircraft derived data. When interrogated, the transponder registers are extracted and sent as the reply messages. Referring to Table 6.8-5, the minimum level of transponder capability for ELS is Comm B.

Table 6.8-3 ELS transponder replies.

Data	Notes
24-bit aircraft address	Allocated to individual aircraft by ICAO via the national registration authority
SSR Mode 3/A	Range and azimuth measurements; selective addressing
SSR Mode C	25 foot altitude resolution (reduced from 100 foot)
Flight status	Ground/airborne. This includes the squawk ident function in the form of downlink aircraft communications (DAPs)
Data link capability report	This information is extracted when the transponder ID is first acquired by the interrogator, see Table 6.8-5
Common usage GICB report	Ground initiated Comm B (GICB), see Table 6.8-5
Aircraft identification	Call sign or registration number; selected by the flight crew (sometimes referred to as the flight ID). This will eventually replace the existing 4096 ATC codes
ACAS active resolution advisory report	Airborne collision avoidance system

Table 6.8-4 EHS transponder replies (in addition to ELS).

Data	Notes
Selected altitude	Typically from the autopilot mode control panel
Roll angle	
Track angle rate	True airspeed as an alternative
True track angle	
Ground speed	
Magnetic heading	
Indicated airspeed	Mach number as an alternative
Rate of climb/descent	

Table 6.8-5 Mode S transponder datalink capability levels.

Level	Type	Detail
1	Comm A	Mode A, C and S surveillance, without a datalink capability
2	Comm B	Level 1, plus standard length transmitting and receiving 112-bit messages
3	Comm C	Level 2, plus receiving of 16 linked 112 extended length messages (ELM)
4	Comm D	Level 3, plus transmitting and receiving 16 linked 112 ELM

The registers are referred to as Comm B data selectors (BDSs); each register contains either specific or common information. Data can be transmitted or received by level 2 (Comm B) transponders; this is via a downlink format (DF) or UF. Alphanumeric strings of data are sent as **downlink aircraft parameters** (DAPs). Certain BDS registers in the transponder are used for specific parameters, other registers have common usage. These registers are checked for timely updates. A report on the status of these updates is provided when requested by the ground station: ground initiated Comm B (GICB).

Test your understanding 6.8.1

On which frequencies are ATC interrogations and transmitted?

Test your understanding 6.8.2

Explain the term 'FRUIT' in the context of ATC.

Test your understanding 6.8.3

Explain the principles of PPM.

Test your understanding 6.8.4

What are the three emergency ATC codes?

Test your understanding 6.8.5

What are the differences between ATC transponder Modes A, C and S?

6.8.5 Automatic Dependent Surveillance-Broadcast (ADS-B)

6.8.5.1 Introduction to ADS-B

The development of Mode S has led onto two related systems: TCASs and ADS-B. Both systems can exchange data directly between aircraft. TCAS is a **surveillance** system that provides warnings directly to the crew when other navigation systems (including ATC) have failed to maintain safe separation of aircraft. ADS-B is an emerging technology for ATM that is intended to replace conventional SSR. TCAS is addressed in Chapter 6.9; ADS-B is described below.

ADS-B is intended eventually to replace conventional ground-based ATC radar systems. The system also provides surveillance in remote areas where ground radar coverage is not possible, e.g. over oceans. ADS-B forms part of the FAA's next generation air transportation system (NGATS). It will revolutionise how pilots obtain traffic and weather information. The intention is to increase air navigation safety by providing crews with real-time information about other traffic; this makes it possible for the crew to be responsible for their own aircraft's separation and collision avoidance. The system is **automatic** in that no interrogation is required to initiate a transponder broadcast from the aircraft; this type of unsolicited transmission is known as a **squitter**.

Key Point

SSR has the following advantages compared with primary radar:

- Low power transmitter
- Superior returns from the target in terms of signal strength and integrity
- Transmissions and returns can be coded to include data
- Smaller antennas.""

Key Point

Altimeters are used in some aircraft to provide an encoded digital output (barometric altitude to the transponder) in Gillham code. This is a modified form of Gray code, where two successive values differ in only one digit. This code prevents spurious outputs from the analog encoder. To illustrate how this code is used, a four-bit parallel output would count from zero to seven as follows: 0000, 0001, 0011, 0010, 0110, 0111, 0101, 0100.""

Key Point

The 100 foot resolution used in many Mode C transponders is being updated to 25 foot resolution for Mode S; serial data from the encoder will be required to achieve this.

ADS-B utilises conventional global navigation satellite system (GNSS) and onboard broadcast equipment for communication via satellites making it **dependent**. Air traffic coordination is thereby provided though **surveillance** between aircraft; the system has a range of approximately 150 nm.

A significant benefit of ADS-B is the estimated 90% cost saving compared with replacing ageing SSR system infrastructures. Other benefits include greater access to optimum routes and altitudes; this leads to reduced fuel consumption and greater utilisation of aircraft. The system provides **real-time data** for both flight crews and air traffic controllers. Data is exchanged between aircraft and can be independent of ground equipment. Since SSR is based on range and azimuth measurements, the accuracy of determining an aircraft's position reduces as a function of range from the antenna. Two aircraft in close proximity, but some range from an SSR ground station, can exchange data via ADS-B and calculate their relative positions more accurately. ADS-B is being proposed with three methods of exchanging data:

- Mode S transponder extended squitter (ES)
- VHF digital link (VDL)
- Universal access transceiver (UAT).

There are advantages and disadvantages for these methods; each technology is competing with the other, driven by technical, operational and political factors. A combination of the above is being introduced on a progressive basis to serve the needs of general aviation and commercial air transport. There are examples of where Mode S and VDL have been integrated into single ground stations. The reader is encouraged to monitor developments via the industry press.

6.8.5.2 Extended Squitter (ES)

The Mode S method extends the information already described above for EHS. ES messages include aircraft position and other status information. The advantage of using ES is that the infrastructure exists via Mode S ground stations and TCAS-equipped aircraft. Note that Mode S provides only unidirectional communications.

6.8.5.3 VHF Digital Link (VDL)

VDL utilises the existing aeronautical VHF frequencies to provide bi-directional communications; digital data is within a 25 kHz bandwidth. This protocol is based on a technique called '**s**elf organising **t**ime **d**ivision **m**ultiple **a**ccess' (STDMA). VDL is suited for short message transmissions from a large number of users over longitude range. The system utilises conventional GNSS to send messages of up to 32 bytes at 9.6 kbps. The system can manage 9,000 32-byte messages per minute. The system is self-organising; therefore, no master ground station is required.

6.8.5.4 Universal Access Transceivers (UATs)

In order to illustrate the principles of ADS-B, the UAT is described in more detail. UAT uses conventional GNSS technology and a relatively simple broadcast communications link, see Figure 6.8-19 for typical ADS-B architecture. The 978 MHz UAT receives inputs from a GNSS, combines this data with other parameters, e.g. airspeed, heading, altitude and aircraft identity, to facilitate the **air traffic management**, see Figure 6.8-20. Flight information services-broadcast (FIS-B), such as weather and other non-ADS-B radar traffic information services-broadcast (TIS-B), can also be uplinked. This data is transmitted to aircraft in the surrounding area, and to ground receivers that distribute the data in real time via existing communication infrastructures. The system allows operations in remote and/or mountainous areas not covered by ground radar. Trials have been conducted in the Yukon- Kuskokwim delta (Alaska), to provide radar-like surveillance (this is an area where secondary radar cannot be deployed).

Referring to Figure 6.8-21, aircraft A is in a remote area, and has flown beyond the range of a ground-based transceiver (GBT). This aircraft continues to broadcast its ADS-B data; however, no other aircraft is within air-to-air range. ATC and other ADS-B-equipped aircraft receive aircraft A's data via **satellite** link. Furthermore, aircraft A receives all other ADS-B aircraft positions (latitude and longitude) together with routine ATC data, e.g. weather updates via satellite link until it flies within range of a GBT. In Figure 6.8-21, aircraft B and C broadcast and receive data via the nearest GBT. (Data courtesy of ADS-B Technologies, LLC.)

6.8.6 Communications, Navigation and Surveillance/Air Traffic Management (CNS/ATM)

This subject is derived from the numerous disciplines and technologies required to enable aircraft to navigate, and ATC to manage, airspace. In many parts of the world, including the USA and Europe, dense traffic flows are currently being managed; other parts of the world are seeing continuing increases in air traffic. Today's ATC infrastructure, including operating methods and equipment, cannot possibly manage the predicted demands of air traffic management. It is vital that **global standards** are developed and implemented for the delivery of a safe, efficient and economic air navigation service provision.

Navigation is not always about flying great circle routes for the shortest distance between two points, e.g. tailwinds should be exploited and headwinds avoided. This requires real-time weather information for pilots and controllers. Close cooperation is also required with airports to ensure efficient arrivals and departures thereby minimising delays. ATM addresses traffic flow at the optimum speed, height and route to minimise fuel consumption.

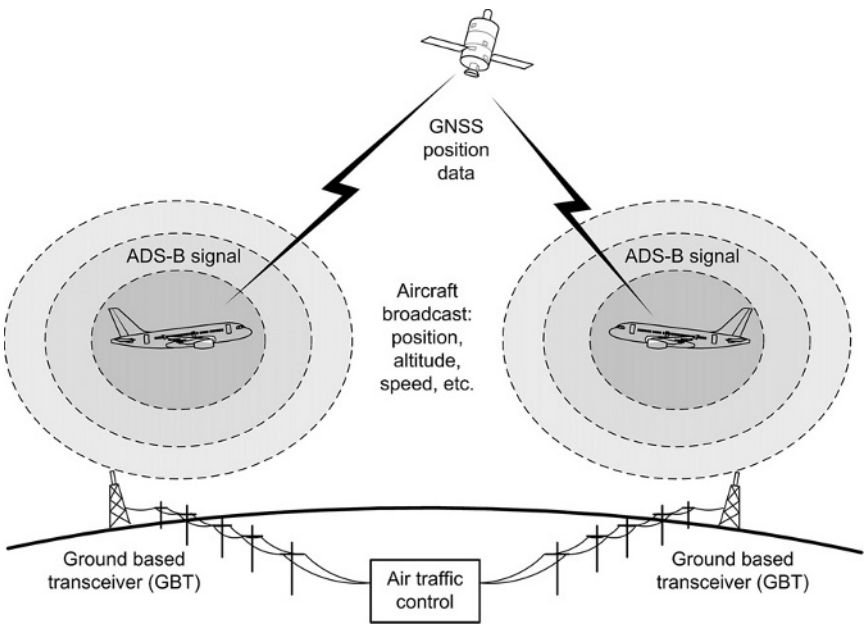


Figure 6.8-19 Typical ADS-B architecture.

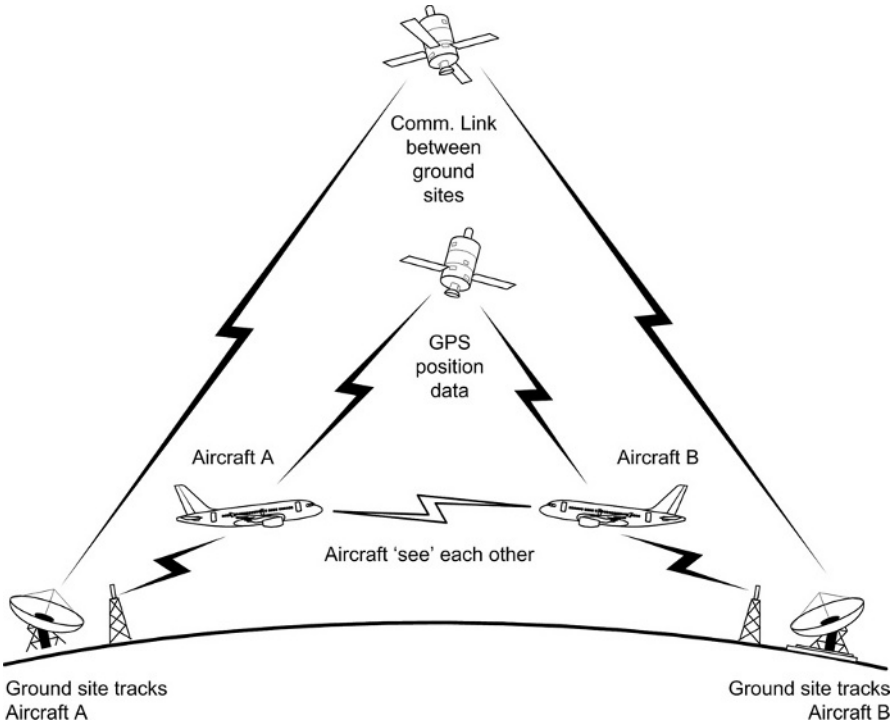


Figure 6.8-20 ADS-B used for air traffic management (ATM).

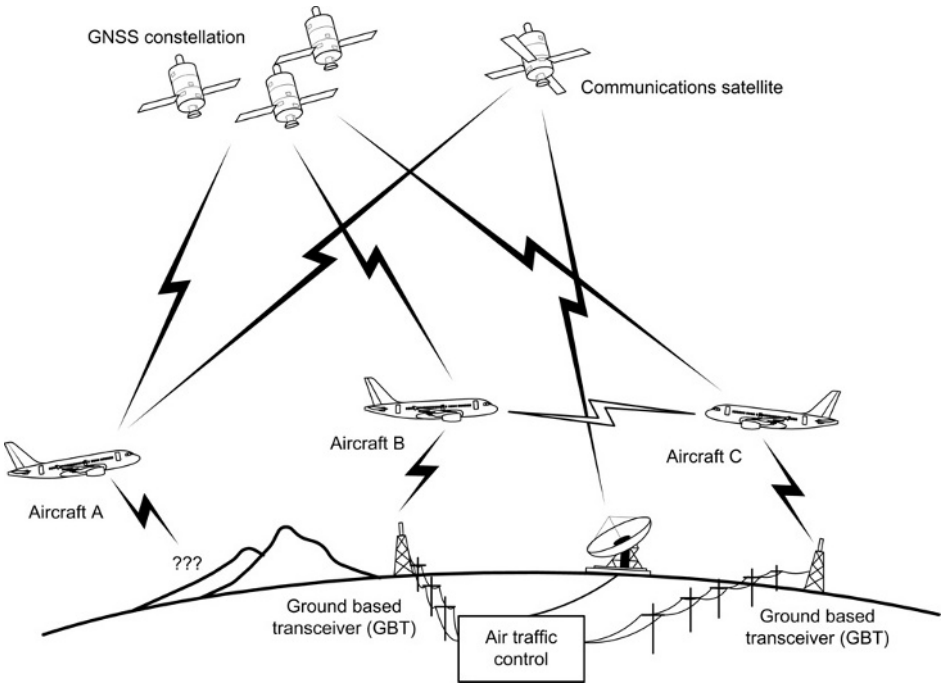


Figure 6.8-21 ADS-B operations in remote areas.

Numerous enabling factors for CNS/ATM will lead to higher navigation accuracy at lower cost (not just the cost of fuel, but also the impact of air travel on the **environment**). Area navigation already provides a flexible and efficient means of en route and terminal area operations in place of airway routings.

The FAA is proposing a new generation of VHF datalink mode 2 (VDL-2) equipment enabling **controller-pilot datalink communications** (CPDLCs). This will supplement company or engineering information currently handled by the lower capacity ACARS data link system. CPDLC features include fourdimensional navigation management,

en route clearances and traffic flow rerouting. The phased introduction of CPDLC will see existing voice communications for ATC purposes only used as a backup.

Key Point

ADS-B provides real-time data for both flight crews and air traffic controllers; data is exchanged between aircraft and is independent of ground equipment. The system allows operations in remote and/or mountainous areas not covered by ground radar.

Key Point

SLS is achieved by transmitting directional and omnidirectional pulses.

Key Point

Mode S eliminates synchronous garbling, increases the capacity of a given air space and improves surveillance accuracy.

Key Point

Enhanced Mode S provides data for the current state of motion of the aircraft, together with the aircraft's vertical intention, i.e. selected altitude.

Key Point

ATC and DME transponders operate in the same frequency band; their transmissions have to be coordinated.

Key Point

The phased introduction of CPDLC will see existing voice communications for ATC purposes only used as a backup.

Key Point

The demands on ATC resources are further reduced by Mode S since replies are only sent during the initial 'all-call' broadcast, and thereafter only when specifically requested.

Key Point

The Mode S reply is in a different form to that of Modes A or C, and is only sent to a Mode S interrogation.

Test your understanding 6.8.6

Who allocates the following: ATC identity codes, 24-bit aircraft address codes?

Test your understanding 6.8.7

What is the difference between ELS and EHS?

6.8.7 Multiple Choice Questions

1. Side-lobes are inherent in which part of the SSR interrogation?
   (a) omnidirectional antenna
   (b) rotating antenna
   (c) transponder reply.

2. The transponder always suppresses any reply when
   (a) P2 amplitude is =P1/P3
   (b) P2 amplitude is ≤P1/P3
   (c) P2 amplitude is ≥P1/P3.

3. Compared with primary radar, the transmission power used by SSR is
   (a) higher
   (b) the same
   (c) lower.

4. ATCRBS ident codes are formatted in which numbering system?
   (a) binary
   (b) octal
   (c) decimal.

5. Directional and omnidirectional pulses are transmitted by the SSR for
   (a) ident and altitude data
   (b) DME suppression
   (c) SLS.

6. Mode C (altitude information) is derived from an
   (a) altimeter or air data computer
   (b) ATC transponder
   (c) ATC control panel.

7. The SPI pulse is transmitted for a period of
   (a) indefinite time
   (b) 15 to 20 seconds
   (c) 8 to 21 μs.

8. SLS is achieved by transmitting:
   (a) directional (P1/P3) and omnidirectional (P2) pulses
   (b) directional (P2) and omnidirectional (P1/3) pulses
   (c) directional (P1) and omnidirectional (P2/3) pulses.

9. The transponder code of 7700 is used for
   (a) general air emergency
   (b) loss of radio
   (c) hijacking.

10. ATC interrogations and replies are transmitted on the following frequencies
    (a) interrogation on 1030 MHz, replies on 1090 MHz
    (b) interrogation on 1090 MHz, replies on 1030 MHz
    (c) interrogation on 1030 MHz, replies on 1030 MHz.

11. The 'Mode S all-call' interrogation will cause Mode A and C transponders to reply with their
    (a) ident and altitude data
    (b) aircraft address code
    (c) SPI pulse.

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.9 Traffic Alert and Collision Avoidance System

Mike Tooley and David Wyatt

With ever increasing air traffic congestion, and the subsequent demands on air traffic control (ATC) resources, the risk of a mid-air collision increases. This risk led to the concept of airborne collision avoidance systems (ACASs) being considered as early as the 1950s. Several ACAS technologies have been developed; this chapter focuses on the traffic alert and collision avoidance system (TCAS). TCAS is an automatic surveillance system that helps aircrews and ATC maintain safe separation of aircraft. It is an airborne system based on secondary radar that interrogates and replies directly between aircraft via a high-integrity data link. The system is functionally independent of ground stations, and alerts the crew if another aircraft comes within a predetermined time to a potential collision. Airborne collision avoidance is a complex task; it has taken years to develop. It is important to note that TCAS is a backup system, i.e. it provides warnings and guidance when other navigation systems (including ATC) have failed to maintain safe separation of aircraft. This chapter provides an overview of TCAS and describes how the system contributes to the safe operation of aircraft.

6.9.1 Airborne Collision Avoidance Systems (ACASs)

There are five different types of ACAS technology in use, or being planned:

Passive receivers. These units are intended for general aviation use; they monitor ATC transponder signals in the immediate area and provide visual or audible signals to warn of nearby traffic. They have a range of approximately 6 miles and monitor 2,500 feet above or below the host aircraft. The receiver is normally located on the aircraft's glareshield. It has an internal antenna, which can lead to intermittent coverage depending on how and where the unit is positioned. Passive systems provide approximations of where another aircraft is, and its position relative to the host aircraft. Some passive devices provide vertical trend information, e.g. indicating if the other aircraft is climbing or descending.

Traffic information system (TIS). This uses the host aircraft's Mode S transponder (refer to Chapter 6.8) to communicate with the ground-based secondary surveillance radar (SSR) network. Traffic information can be obtained within a 5-mile radius, 1,200 feet above or below the host aircraft. This traffic information is provided on a (near) real-time basis. The attraction of TIS is that aircraft hardware and software are minimal since the system 'feeds' off ground station computations. TIS is unavailable outside of areas covered by SSR and will be superseded by a system called automatic dependent surveillance-broadcast (ADS-B).

Traffic advisory system (TAS). The host aircraft's TAS actively monitors the airspace seeking nearby transponder-equipped aircraft and provides relevant traffic information via a display and audio warning. TAS uses active interrogation of nearby transponders to determine another aircraft's position and movement. The system can track up to 30 aircraft with a range of up to 21 nm, 10,000 feet above or below the host aircraft.

Traffic alert and collision avoidance system (TCAS). This is the industry standard system mandated for use by commercial transport aircraft, and the main subject of this chapter. Two types of TCAS are in operation, TCAS I and II. Both systems provide warnings known as 'advisories' to alert the crew of a potential collision. **TCAS I** assists the crew in visually locating and identifying an **intruder** aircraft by issuing a **traffic advisory** (TA) warning. **TCAS II** is a collision avoidance system and, in addition to traffic advisories, provides vertical flight manoeuvre guidance to the crew. This is in the form of a **resolution advisory** (RA) for **threat** traffic. An RA will either increase or maintain the existing vertical separation from an intruder aircraft. If two aircraft in close proximity are equipped with TCAS II, the flight manoeuvre guidance is coordinated between both aircraft. A third type of system (TCAS III) was intended to provide lateral guidance to the crew; however, this has been superseded by a new concept: ADS-B.

Automatic dependent surveillance-broadcast (ADS-B). This system is intended eventually to replace conventional ground-based ATC radar systems. The system also provides surveillance in remote areas where ground radar coverage is not possible, e.g. over oceans. ADS-B forms part of the FAA's next generation air transportation system (NGATS) and it will revolutionise how pilots obtain traffic and weather information. The intention is to increase air navigation safety by providing crews with real-time information about other traffic; this makes it possible for the crew to be responsible for their own aircraft's separation and collision avoidance. The system is **automatic** in that no interrogation is required to initiate a transponder broadcast from the aircraft; this

type of unsolicited transmission is known as a **squitter**. ADS-B utilises conventional global navigation satellite system (GNSS) and onboard broadcast equipment for communication via satellites making it **dependent**. Air traffic coordination is thereby provided through **surveillance** between aircraft. ADS-B is further described in Chapter 6.8 (ATC).

6.9.2 TCAS Overview

SSR transmits a specific low energy signal (the interrogation) to a known target. This signal is analysed and a new (or secondary) signal, i.e. not a reflected signal, is sent back (the reply) to the origin, see Figure 6.9-1. In the TCAS application, interrogations and replies are sent directly between the on-board ATC transponders, see Figure 6.9-2. The TCAS computer interfaces with the ATC transponder and calculates the time to a potential collision known as the **closest point of approach** (CPA).

TCAS creates a **protected volume of airspace** around the host aircraft, see Figure 6.9-3; this is based on altitude separation and a calculated time to the CPA. The Greek letter **tau** (T) is the symbol used for the approximate time (in seconds) to the CPA, or for the other aircraft reaching the same altitude. This protected volume of airspace is determined as a function of time (tau) for both range and vertical separation:

$$\text{Range tau} = \frac{3{,}600 \times \text{slant range (nm)}}{\text{closing speed (knots)}}$$

$$\text{Vertical tau} = \frac{\text{altitude separation (feet)} \times 60}{\text{combined vertical speed (fpm)}}$$

TCAS interrogates other aircraft within this protected airspace and obtains their flight path details, i.e. range, altitude and bearing. This data is analysed along with the host aircraft's flight path. If there is a potential conflict between flight paths, a visual and audible warning is given to the crew. This warning depends on the type of equipment installed in the host and other traffic as shown in Figure 6.9-4.

Tau is programmed for varying sensitivity levels determined by altitude bands as illustrated in Table 6.9-1. For each altitude band, there is a different sensitivity level and corresponding value of tau for traffic and resolution advisories (TA and RA respectively). Higher sensitivity levels provide a larger protected volume of airspace. If closure rates are low, modifying the range boundaries required to trigger a TA or RA provides a further refinement to the calculation of collision avoidance; this distance modification is known as DMOD.

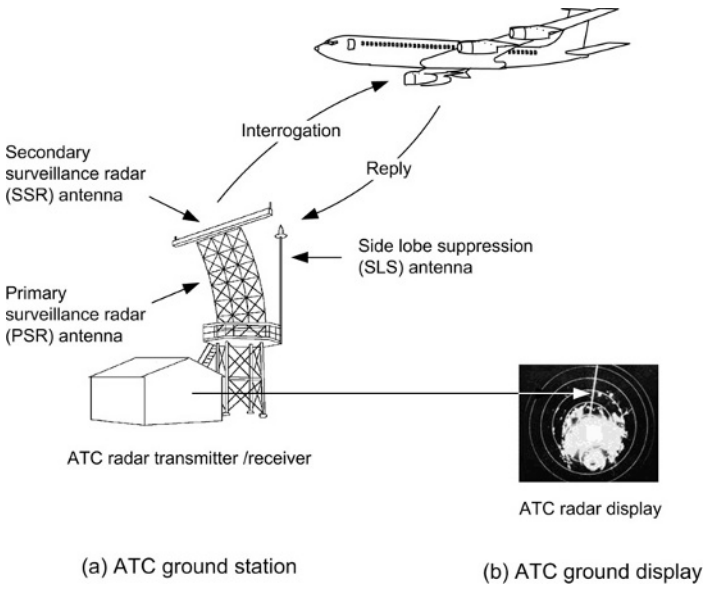


Figure 6.9-1 Secondary surveillance radar.

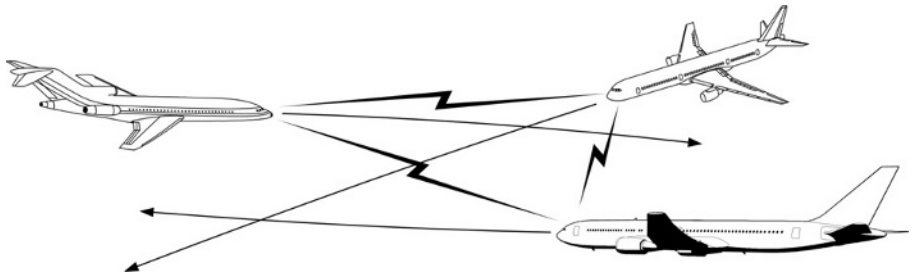


Figure 6.9-2 Airborne equipment—data link communication.

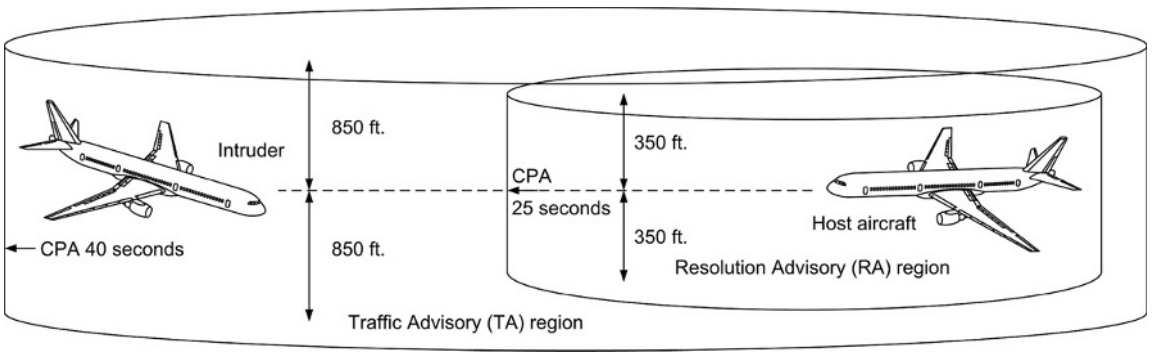


Figure 6.9-3 Protected airspace volume.

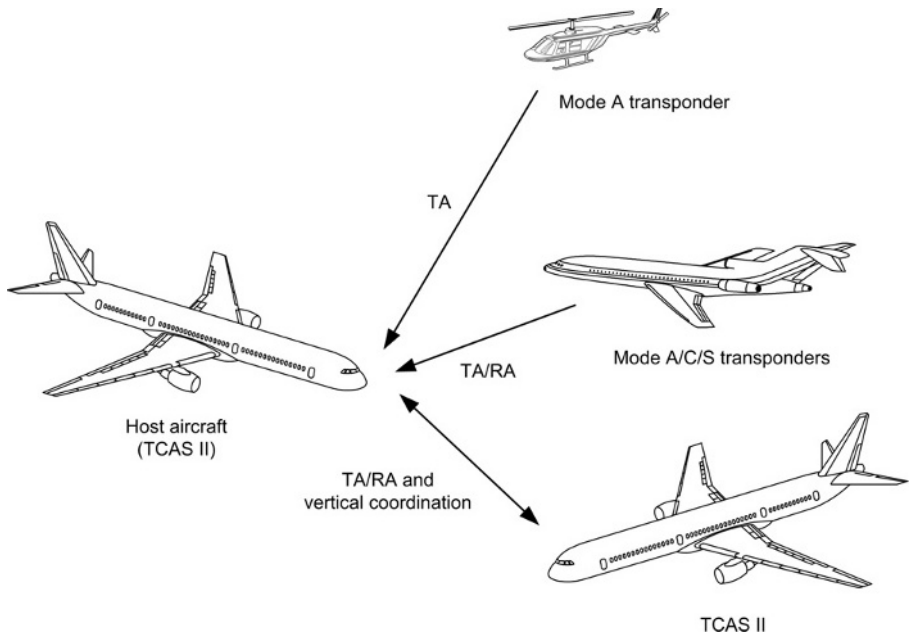


Figure 6.9-4 Variations of warning given to host aircraft depending on intruder's equipment.

Aircraft that are on the ground are filtered out of the collision avoidance algorithms to reduce nuisance warnings. Other Mode S-equipped aircraft are monitored if their altitude is less than 1750 feet above the ground. Referring to Figure 6.9-5, by using a combination of the host aircraft's barometric and radio altitude, together with the barometric altitude of the target aircraft, any target below 360 feet is deemed to be on the ground.

Table 6.9-1 TA/RA sensitivity levels.

Altitude (feet)*	Sensitivity levels	Tau (seconds)		DMOD (nm)		Altitude threshold (feet)	
		TA	RA	TA	RA	TA	RA
0 to 1,000	2	20	None	0.30	None	850	n/a
1,000 to 2,350	3	25	15	0.33	0.20	850	300
2,350 to 5,000	4	30	20	0.48	0.35	850	300
5,000 to 10,000	5	40	25	0.75	0.55	850	350
10,000 to 20,000	6	45	30	1.00	0.80	850	400
20,000 to 42,000	7	48	35	1.30	1.10	850	600
>42,000	7	48	35	1.30	1.10	1200	700

*Radio altitude for sensitivity levels 2/3, thereafter pressure altitude from the host aircraft's barometric altimeter (Data source: ARINC).

Key Point

TCAS is an airborne system based on secondary radar that interrogates and replies to other aircraft; the system utilises the aircraft's Mode S transponders and is functionally independent of the aircraft navigation systems and ground stations.

Key Point

The CPA is derived as a function of time, referred to as tau.

Test your understanding 6.9.1

What is the difference between TCAS I and TCAS II?

Test your understanding 6.9.2

What do the abbreviations TA and RA stand for?

6.9.3 TCAS Equipment

The system consists of one TCAS computer, two directional antennas, a control panel and two displays, see Figure 6.9-6. TCAS operates in conjunction with the Mode S surveillance system which includes two transponders,

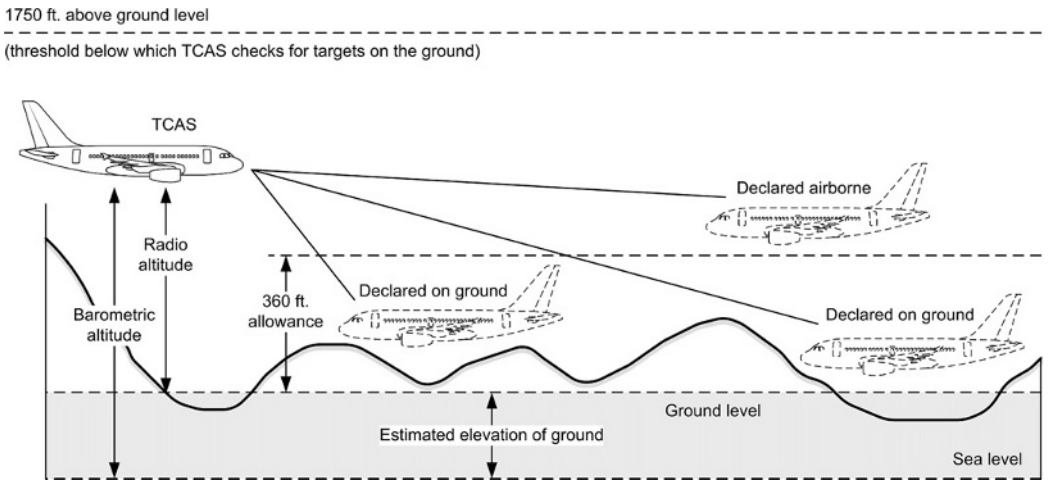


Figure 6.9-5 Aircraft deemed by TCAS as airborne/declared on the ground (courtesy of ARINC).

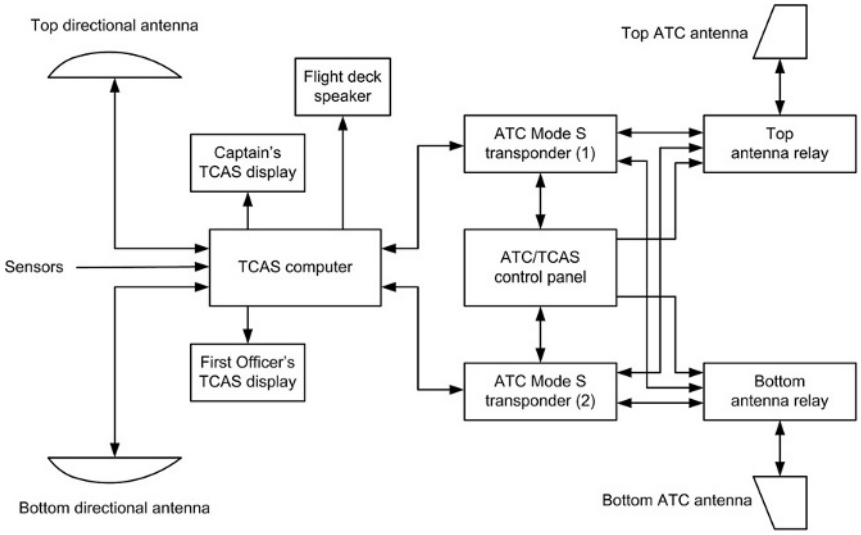


Figure 6.9-6 TCAS airborne equipment.

a control panel and antennas, as described in Chapter 6.8. Visual warnings can be displayed on the instantaneous vertical speed indicator (IVSI) or electronic flight instrument system (EFIS).

6.9.3.1 Antennas

These are located on the top and bottom of the fuselage as illustrated in Figure 6.9-7; this provides **antenna diversity**, a technique in which the datalink signal is transmitted along different propagation paths. The upper antenna is directional and is used for tracking targets above the host aircraft; the bottom antenna can either be **omnidirectional** or **directional** as an operator specified option. Interrogation codes are transmitted via the Mode S transponder on a 1030 MHz carrier wave; reply codes are transmitted on a 1090 MHz carrier wave. The phase array directional antennas are electronically steerable and transmit in four lateral segments at varying power levels. Note that two Mode S transponder antennas (Figure 6.9-7) are also required for TCAS operation. The latter is suppressed when the Mode S transponder is transmitting so that TCAS does not track the host aircraft.

6.9.3.2 Computer

This is a combined transmitter, receiver and processor that performs a number of functions including:

- **Monitoring** of the surveillance airspace volume for aircraft
- **Tracking** other aircraft
- **Monitoring** its own aircraft altitude

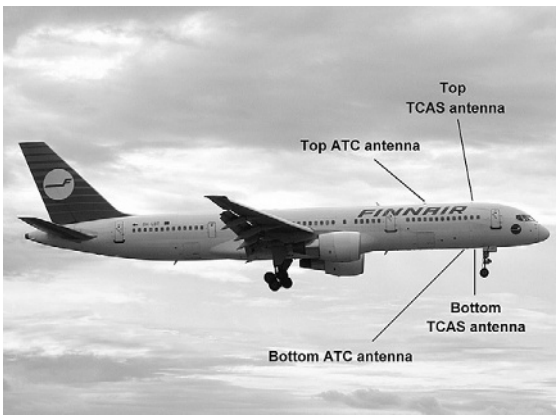


Figure 6.9-7 Location of ATC/TCAS antennas.

- Issuing **warnings** for potential flight path conflicts
- Providing recommended **manoeuvres** to avoid potential flight path conflicts.

Inputs to the computer include the host aircraft's heading, altitude and maximum airspeed. Other configuration inputs include landing gear lever position and weight on wheels sensor. Collision avoidance algorithms are used to interpret data from the host aircraft and proximate traffic.

6.9.3.3 Control Panel

This is a combined ATC/TCAS item, see Figure 6.9-8. The four-digit aircraft identification code is selected by either rotary switches or push buttons, and displayed in a window. (Refer to Chapter 6.8 for detailed ATC operation.) The mode select switch is used to disable TCAS surveillance, enable traffic advisories only, or both traffic and resolution advisories. With the system in standby, the TCAS transponder is powered but it will not interrogate or reply to interrogations; all surveillance and tracking functions are disabled. The above/below switch (ABV-N-BLW) allows the crew to select three bands of surveillance above or below the aircraft:

- ABV +7000 feet/−2700 feet
- N ±2700 feet
- BLW +2700 feet/−7000 feet.

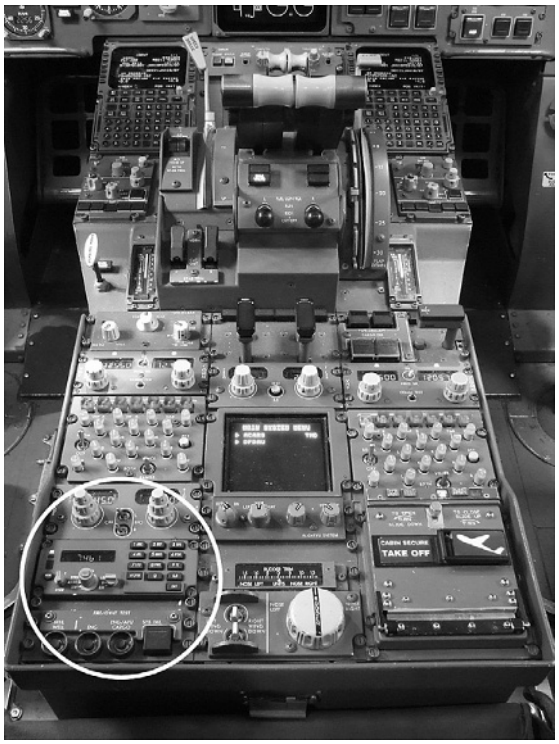


Figure 6.9-8(a) Location of ATC/TCAS control panel.

Figure 6.9-8(b) ATC/TCAS control panel.

6.9.3.4 Displays

The displays used for TCAS advisories vary between aircraft types. These include the IVSI and/or the EFIS. In either case, the advisory warnings are based on the same icons. Details of both IVSI and EFIS displays are provided in the description of TCAS system operation that follows in the next section.

6.9.4 System Operation

6.9.4.1 TCAS Compatibility

Referring to Figure 6.9-4, there will be a combination of aircraft systems in any given airspace. If the host aircraft is fitted with TCAS I equipment, then its computer will provide traffic advisories, regardless of the surrounding aircraft ATC transponder types. (Aircraft not fitted with a transponder are not tracked by TCAS.) When the host aircraft is fitted with TCAS II, but other aircraft have different transponder types, the advisories provided to the host crew will be as shown in Figure 6.9-4.

TCAS requires that the aircraft is equipped with a Mode S ATC transponder; the computers in TCAS II-equipped aircraft will coordinate their guidance commands such that they provide complementary manoeuvres. In the latest versions of TCAS, aircraft performance is taken into account when providing these commands.

6.9.4.2 Advisory Warnings

Traffic icons are shown relative to the host aircraft; these are colour coded to depict their threat level as shown in Figure 6.9-9. These icons are supplemented by altitude information for the other aircraft: relative altitude (± to depict if the other aircraft is above or below the host aircraft) and vertical manoeuvre (climbing or descending indicated by an arrow). These icons are displayed on the IVSI (Figure 6.9-10) or EHSI (Figure 6.9-11).

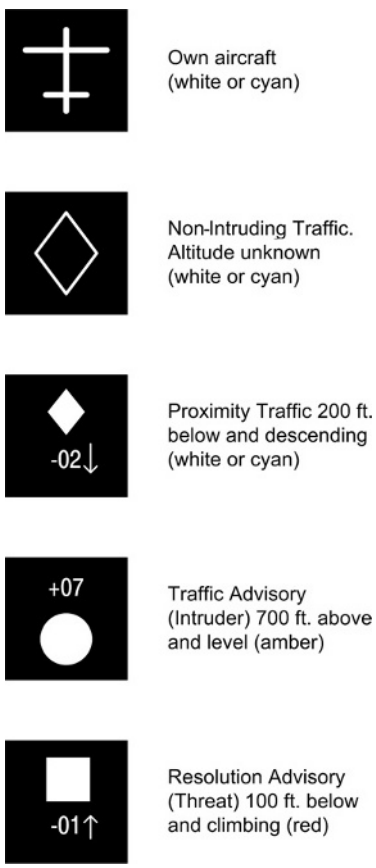


Figure 6.9-9 Traffic warning icons.

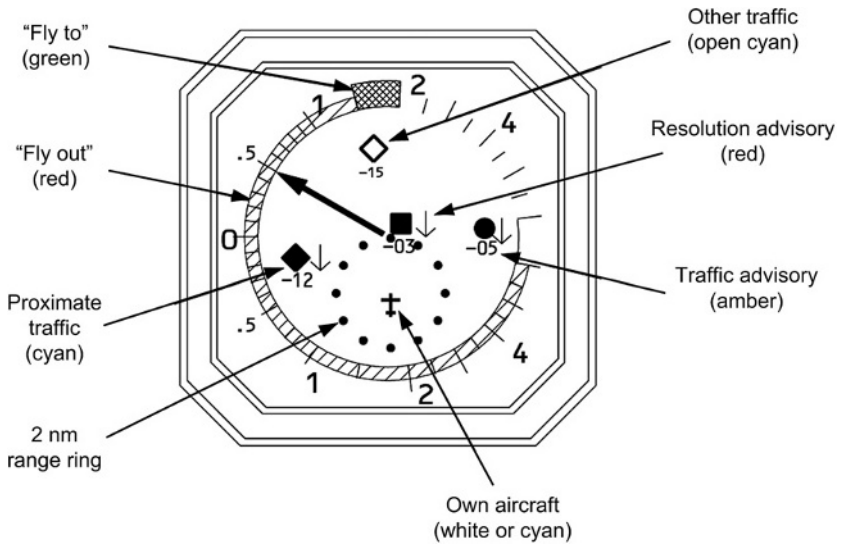


Figure 6.9-10 IVSI display with TCAS icons.

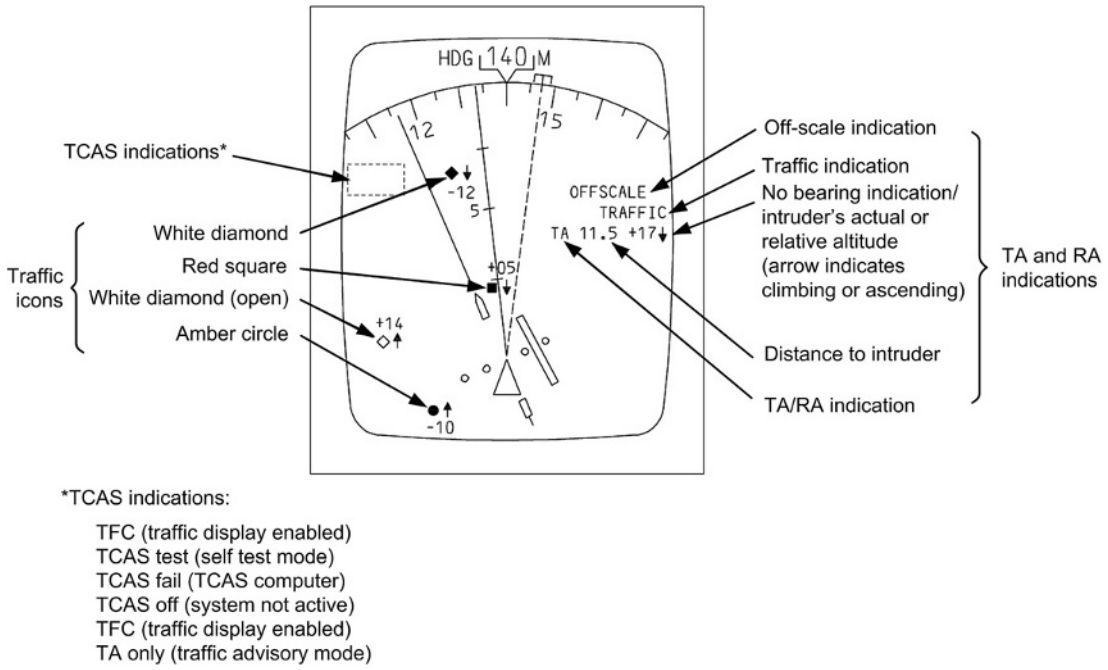


Figure 6.9-11 EHSI display with TCAS icons and messages.

Referring to Figure 6.9-11, the EHSI also has a display area for TCAS system indications (system off, self-testing, etc.), traffic icons and TA/RA indications:

- **Off-scale**: intruder aircraft is out of display range
- **Traffic**: TCAS has detected intruder within the protected airspace
- **No bearing**: TCAS cannot determine the bearing of an intruder
- **TA/RA**: TCAS has identified an intruder or threat aircraft.

6.9.4.3 TCAS Guidance

In the event of an RA, the IVSI will indicate red and green bands around the display to guide the pilot into a safe flight path, see Figure 6.9-10.

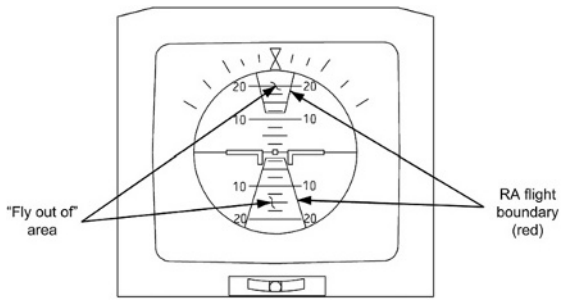


Figure 6.9-12 EADI vertical guidance.

Aircraft with an electronic attitude direction indicator (EADI) have 'fly-out-of' guidance (Figure 6.9-12) in the form of a red boundary. The pilot has to climb or descend, keeping the aircraft outside of these calculated boundaries until the RA is cleared. Note that the latest versions of TCAS take aircraft performance into account when issuing vertical guidance.

Key Point

TCAS interrogates aircraft within a surveillance volume of airspace and obtains their flight path details. This data is analysed along with the host aircraft's flight path. If there is a potential conflict between flight paths, a visual and audible warning is given to the crew.

Key Point

TCAS traffic and resolution advisories are displayed on either the IVSI or EFIS. If the host aircraft is fitted with TCAS II, but other traffic has different TCAS equipment or ATC transponders, the information provided to the host aircraft will vary.

6.9.4.4 TCAS Commands

Aural warnings are produced via dedicated TCAS speakers in the cockpit, or through the aircraft's audio system. TAs are announced by the words '*traffic*, *traffic*', stated once for each TA. RAs are announced as shown in Table 6.9-2; these are referenced to an IVSI for illustration purposes. Aural warnings are inhibited at altitudes less than 500 feet above ground level.

6.9.4.5 TCAS Surveillance

TCAS has an effective and reliable surveillance range of 14 nm. The host aircraft can simultaneously track at least 30 transponder-equipped aircraft within its surveillance range, with traffic densities of up to 24 aircraft within a 5 nm radius. The surveillance function transmits interrogations at 1030 MHz; transponders on nearby aircraft reply at 1090 MHz. These replies are decoded into range, altitude, and bearing by the TCAS computer and then analysed for potential conflicts by collision avoidance algorithms in the computer's software. If the ATC transponder is interrogated by a TCAS II-equipped aircraft, it will select the appropriate antenna to transmit the reply. This technique is called **antenna diversity**; this enhances visibility with TCAS-equipped aircraft flying above the host aircraft.

Key Point

TCAS I issues TAs these assist the crew in visually identifying intruder traffic.

Key Point

TCAS II issues TAs and RAs; the latter provides recommended manoeuvres needed to increase or maintain vertical separation.

Table 6.9-2 TCAS aural annunciations.

Aural warning	Notes
Climb, climb	Achieve climb rate in green arc.
Descend, descend	Achieve descent rate in green arc.
Monitor vertical speed	Check that vertical speed is out of the red arc.
Adjust vertical speed, adjust	Achieve vertical speed within the green arc.
Climb, crossing climb—climb, crossing climb	Achieve climb rate in green arc; safe separation is achieved by flying through the intruder's flight path.
Descend, crossing descend—descend, crossing descend	Achieve descent rate in green arc; safe separation is achieved by flying through the intruder's flight path.
Maintain vertical speed, maintain	Achieve climb or descent rate in green arc; safe separation is achieved by not changing the flight path.
Maintain vertical speed, crossing maintain	Achieve climb or descent rate in green arc; safe separation is achieved by not changing the flight path and by flying through the intruder's flight path.
Increase climb—increase climb	Achieve climb rate in green arc; aural warning has an increased sense of urgency.
Increase descent—increase descent	Achieve descent rate in green arc; aural warning has an increased sense of urgency.
Climb, climb now—climb, climb now	Received after a 'descend' RA has failed to reduce separation, i.e. a change of avoiding manoeuvre is required to achieve safe separation. Aural warning has an increased sense of urgency.
Descend, descend now—descend, descend now	Received after a 'climb' RA has failed to reduce separation, i.e. a change of avoiding manoeuvre is required to achieve safe separation. Aural warning has an increased sense of urgency.
Clear of conflict	Separation of aircraft is now adequate, i.e. the RA has now been removed

Mode S surveillance

TCAS surveillance of Mode S-equipped aircraft is relatively straightforward because of its inherent selective address feature. TCAS monitors spontaneous broadcast transmissions (referred to as **squitters**) which are generated once per second by the Mode S transponder. In addition to aircraft data, the squitter contains the unique Mode S address of the sending aircraft. Once the squitter message has been received and decoded, TCAS sends a Mode S interrogation to the Mode S address contained in the squitter. The Mode S transponder replies to this interrogation, and the information received is used by TCAS to determine the range, bearing and altitude of the Mode S aircraft.

Limiting the rate at which a Mode S aircraft is interrogated reduces congestion of the 1030/1090 MHz channels. At extended ranges, a target is interrogated at least once every five seconds. As the target aircraft approaches the protected airspace where a TA may be required, the interrogation rate increases to once per second.

Mode A/C surveillance

TCAS uses a Mode C only 'all-call' to interrogate nearby aircraft equipped with Mode A/C transponders. The nominal rate for these interrogations is once per second. Aircraft equipped with Mode A transponders reply to TCAS interrogations with no data contained in the altitude field of the reply. The replies from Mode C transponders are monitored and tracked in range, altitude and bearing.

Surveillance of Mode C transponders by TCAS-equipped aircraft is complicated by problems associated with synchronous and nonsynchronous garbling (in a similar way to that described for the ATC system, Chapter 6.8). The length of reply message from a Mode C transponder is 21 μs (refer to Chapter 6.8). Since the speed of radar pulse propagation is 3×10^{-8} m/s, the distance travelled in 21 μs will be 6300 metres, or 3.4 nm. All Mode C transponders within a range difference of ±1.7 nm from the host aircraft will send replies that overlap when received by TCAS; this is referred to as **synchronous garbling**, see Figure 6.9-13. A technique known as **whisper-shout** (WS) is used to overcome this. WS varies the power level of interrogations on a **progressive** basis; aircraft that are close to the host aircraft send their replies. The next interrogation suppresses the transponders that have already replied, see Figure 6.9-14, but seeks replies from aircraft that did not reply to the first interrogation. This process is repeated 24 times to ensure that all Mode C transponders in the given airspace provide a reply. Using directional signals as illustrated in Figure 6.9-15 further reduces the number of overlapping replies.

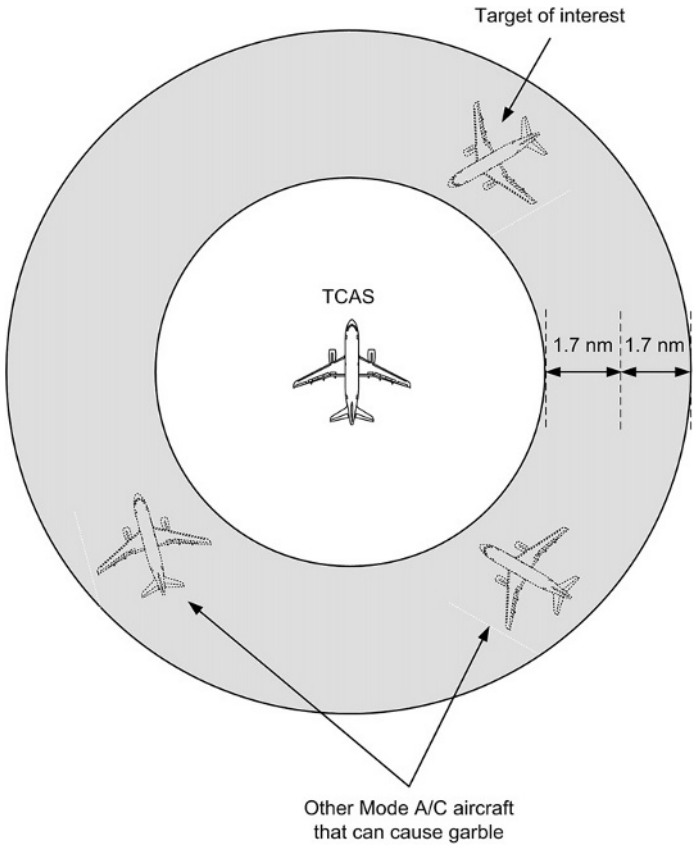


Figure 6.9-13 Synchronous garble area.

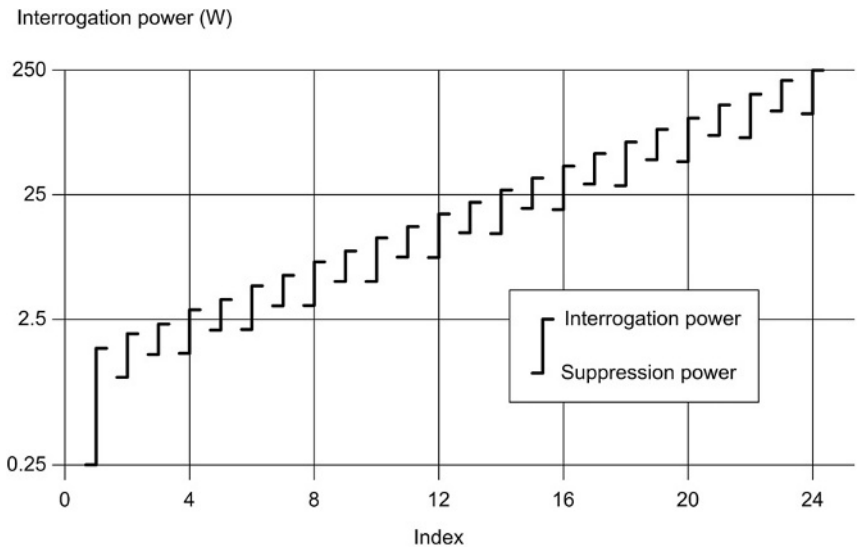


Figure 6.9-14 WS interrogation.

Test your understanding 6.9.3

What are the two types of TCAS advisory warnings provided to the flight crew in the event of a potential flight path conflict?

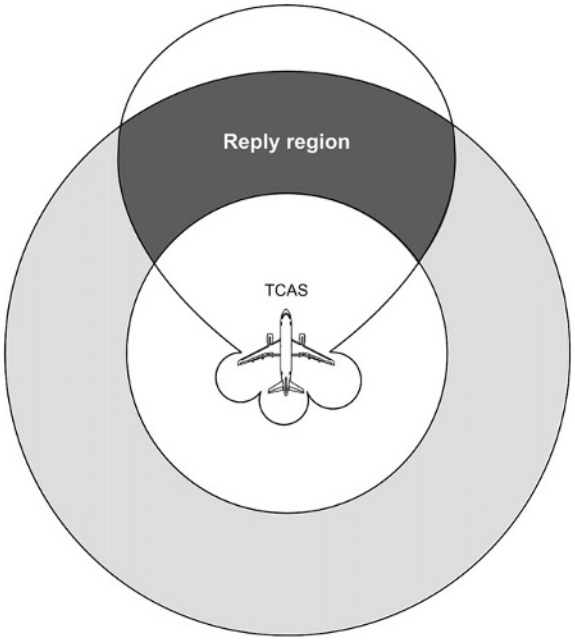


Figure 6.9-15 Directional transmission of TCAS and reply region.

Test your understanding 6.9.4

1. What symbols and colours are used for nonthreat and proximity traffic?
2. What symbols and colours are used for TA and RA warnings?

Test your understanding 6.9.5

What is the purpose of the 'WS' technique?

Non-synchronised garble or **false replies from unsynchronised interrogator transmissions** (FRUITs) are caused by undesired transponder replies that were generated in response to interrogations from ground stations or TCAS-equipped aircraft. (Note that **FRUIT** is sometimes written as false replies uncorrelated in time.) Since these replies are transitory, algorithms in the TCAS surveillance logic can discard them.

The final consideration in TCAS surveillance is the effect of **multi-path errors**; these are caused by more than one reply being received for a single interrogation. This is a reflected reply, and usually occurs over flat terrain. A technique known as dynamic minimum triggering level (DMTL) is used within the computer to discriminate against delayed and lower power level replies.

Test your understanding 6.9.6

What is the difference between these two TCAS warnings: *'climb, climb'* and *'climb, climb now'*?

6.9.5 Multiple Choice Questions

1. The CPA is determined as a function of
   (a) range
   (b) time
   (c) closing speed.

2. TCAS interrogation codes are transmitted and received on
   (a) 1030 MHz and 1090 MHz carrier waves
   (b) 1090 MHz and 1030 MHz carrier waves
   (c) 1030 MHz or 1090 MHz carrier waves.

3. What colour and shape of symbol is used for proximity traffic?
   (a) solid red square
   (b) solid cyan diamond
   (c) solid orange circle.

4. What TCAS equipment is required on host aircraft and threat traffic in order to provide coordinated manoeuvres?
   (a) TCAS I
   (b) Mode C transponder
   (c) TCAS II.

5. TCAS II requires which type of ATC transponder:
   (a) Mode S
   (b) Mode A
   (c) Mode C.

6. The version of TCAS that provides vertical flight manoeuvre guidance to the crew is
   (a) TAS
   (b) TCAS II
   (c) TCAS I.

7. The directional TCAS antennas transmit in:
   (a) four vertical segments at varying power levels
   (b) four lateral segments at fixed power levels
   (c) four lateral segments at varying power levels.

8. The WS technique:
   (a) varies the power level of interrogations on a decreasing basis
   (b) varies the power level of interrogations on an increasing basis
   (c) maintains the power level of interrogations on a progressive basis.

9. Recommended manoeuvres needed to increase or maintain vertical separation are provided by what type of TCAS warning?
   (a) resolution advisory
   (b) traffic advisory
   (c) non-threat traffic.

10. TAs:
    (a) assist the crew in visually searching and identifying an intruder
    (b) provide recommended manoeuvres needed to maintain vertical separation
    (c) provide lateral guidance to the crew.

11. TCAS interrogations and replies are sent:
    (a) directly between the onboard ATC transponders
    (b) via a ground link
    (c) directly to ATC ground controllers.

12. TCAS warning icons are shown relative to the
    (a) host aircraft
    (b) intruder aircraft
    (c) proximate traffic.

13. Range tau is based on the:
    (a) altitude separation and closing speed of traffic
    (b) slant range and closing speed of traffic
    (c) altitude separation and combined vertical speed.

14. TCAS is an airborne system based on
    (a) secondary radar
    (b) primary radar
    (c) DME.

15. An arrow and ± sign combined with a TCAS icon indicates the:
    (a) relative altitude information for the intruder aircraft
    (b) bearing of the intruder aircraft
    (c) recommended avoidance manoeuvre.

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.10 Data Buses

Mike Tooley

Aircraft data bus systems allow a wide variety of avionics equipment to communicate with one another and exchange data. In this section we shall take a brief look at the principles of aircraft data bus systems before introducing some of the systems that are commonly used in modern aircraft.

6.10.1 Introducing Bus Systems

The word 'bus' is a contraction of the Greek word 'omnibus' and the word simply means 'to all'. Thus, in the context of computers and digital systems, 'bus' refers to a system that permits interconnection and data exchange between the devices in a complex system. Note, however, that 'interconnection' involves more than just physical wiring, amongst other things it defines the voltage levels and rules (or **protocols**) that govern the transfer of data.

With such a large number of avionic systems, a modern aircraft requires a considerable amount of cabling. Furthermore, some of the cabling runs in a large aircraft can be quite lengthy, as shown in Figure 6.10-1. Aircraft cabling amounts to a significant proportion of the unladen weight of an aircraft and so minimising the amount of cabling and wiring present is an important consideration in the design of modern aircraft, both civil and military.

6.10.1.1 Bus Terminology

Bus systems can be either **unidirectional** (one way) or **bidirectional** (two way), as shown in Figure 6.10-2. They can also be **serial** (one bit of data transmitted at a time) or **parallel** (where often 8, 16 or 32 bits of data appear as a group on a number of data lines at the same time). Because of the constraints imposed by conductor length and weight, all practical aircraft bus systems are based on serial (rather than parallel) data transfer. Bus systems provide an efficient means of exchanging data between the diverse avionic systems found in a modern aircraft (see Fig. 6.10-3).

Individual **Line Replaceable Units (LRUs)**, such as the Engine Data Interface or Flap/Slat Electronics Units shown in Figure 6.10-3, are each connected to the bus by means of a dedicated **bus coupler** and **serial interface module** (not shown in Fig. 6.10-3).

Within the LRU, the dedicated digital logic and microprocessor systems that process data locally each make use of their own **local bus** system. These local bus systems invariably use parallel data transfer which is ideal for moving large amounts of data very quickly but only over short distances.

6.10.1.2 Bus Protocols

The notion of a protocol needs a little explaining so imagine for a moment that you are faced with the problem of organizing a discussion between a large number of people sitting around a table who are blindfolded and therefore cannot see one another. In order to ensure that they didn't all speak at once, you would need to establish some ground rules, including how the delegates would go about indicating that they had something to say and also establishing some priorities as to who should be allowed to speak in the event that several delegates indicate that they wish to speak at the same time. These (and other) considerations would form an agreed protocol between the delegates for conducting the discussion. The debate should proceed without too many problems provided that everybody in the room understands and is willing to accept the protocol that you have established. In computers and digital systems **communications protocols** are established to enable the efficient exchange of data between multiple devices connected to the same bus. A number of different standards are commonly used.

Key Point

Modern aircraft use multiple redundant bus systems for exchanging data between the various avionic systems and subsystems. These bus systems use serial data transfer because it minimises the size and weight of aircraft cabling.

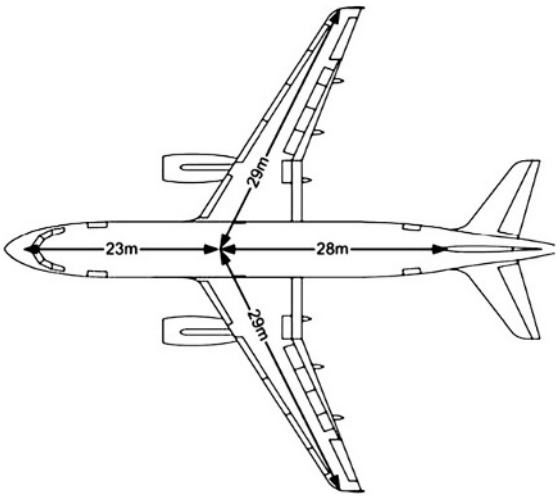


Figure 6.10-1 Typical dimensions of a modern passenger aircraft.

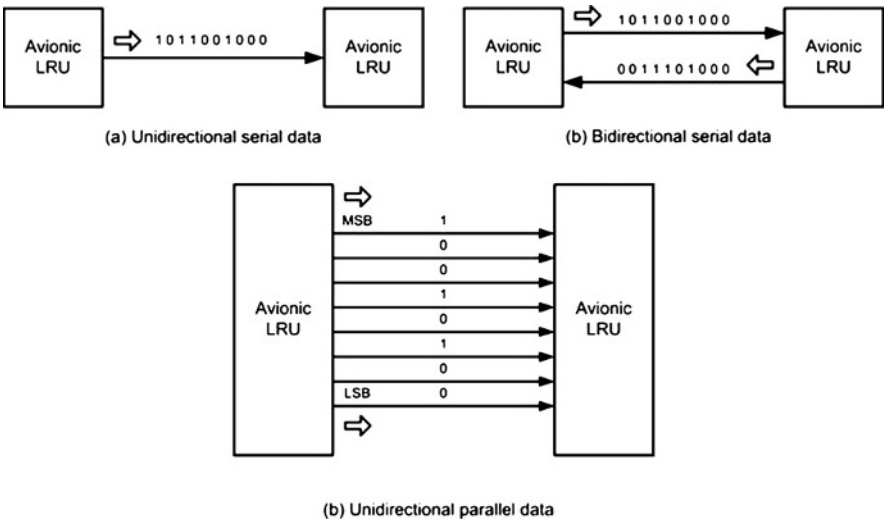


Figure 6.10-2 Unidirectional and bidirectional serial and parallel data.

6.10.1.3 Bus Architecture

Bus architecture is a general term that refers to the overall structure of a computer or other digital system that relies on a bus for its operation. The architecture is often described in the form of a block schematic diagram showing how the various system elements are interconnected and also how the data flow is organised between the elements. The architecture of a system based on the use of a unidirectional serial bus system is shown in Figure 6.10-4(a) whilst a comparable bidirectional bus system is shown in Figure 6.10-4(b). Note how the bidirectional system simplifies the interconnection of the LRUs and allows all of the devices to transmit and receive on the same bus.

Key Point

Communication protocols enable the efficient exchange of data between a number of devices connected to the same bus. Protocols consist of a set of rules and specifications governing, amongst other things, data format and physical connections.

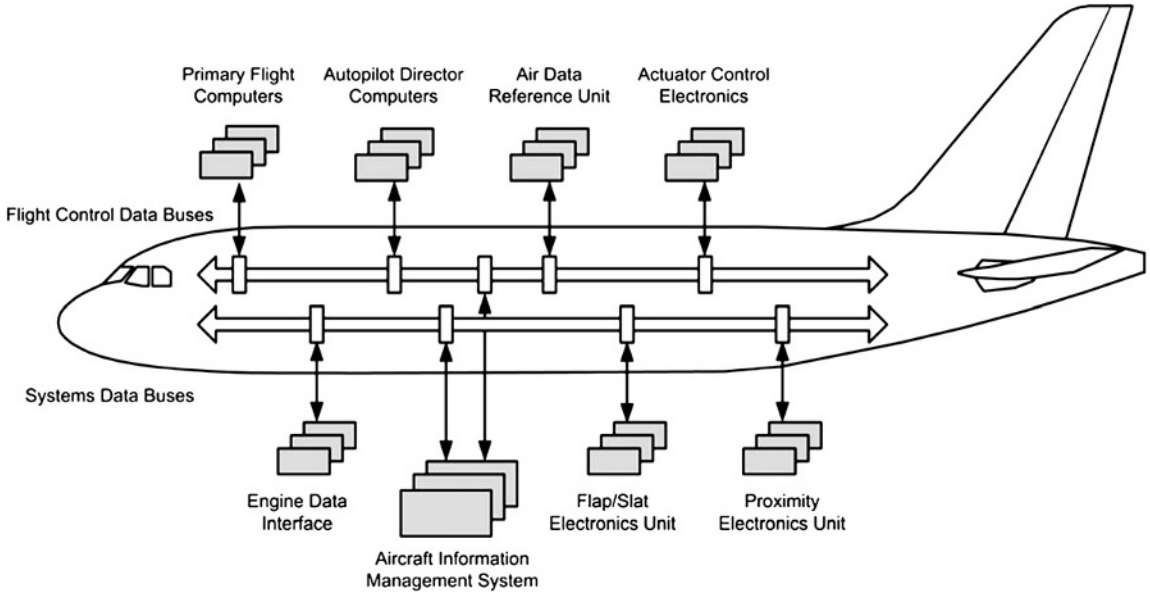


Figure 6.10-3 Multiple bus systems implemented on a modern passenger aircraft.

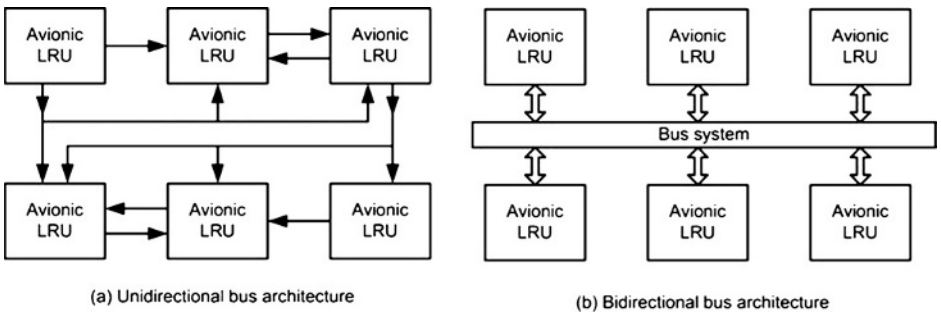


Figure 6.10-4 Bus architecture.

6.10.1.4 Serial Bus Principles

A simple system for serial data transfer between two LRUs each of which comprises an avionic system in its own right is shown in Figure 6.10-5. Within the LRU data is transferred using an internal parallel data bus (either 8, 16, 32 or 64 bits wide). The link between the two LRUs is made using a simple serial cable (often with only two, four or six conductors). The required parallel-to-serial and serial-to-parallel data conversion is carried out by a bus interface (often this is a single card or module within the LRU). The data to be transferred can be **synchronous** (using clock signals generated locally within each LRU) or it may be **asynchronous** (i.e. self-clocking).

The system shown in Figure 6.10-5 has the obvious limitation that data can only be exchanged between two devices. In practice we need to share the data between many LRU/avionic units. This can be achieved by the bus system illustrated in Figure 6.10-6. In this system, data is transferred using a **shielded twisted pair (STP) bus cable** with a number of **coupler panels** that are located at appropriate points in the aircraft (e.g. the flight deck, avionics bay, etc). Each coupler panel allows a number of avionic units to be connected to the bus using a **stub cable**. In order to optimise the speed of data transfer and minimise problems associated with reflection and mismatch, the bus cable must be terminated at each end using a matched **bus terminator**.

Bus couplers are produced as either **voltage mode** or **current mode** units depending upon whether they use voltage or current sensing devices. Within each LRU/avionics unit, an interface is provided that performs the required serial-to-parallel or parallel-to-serial data conversion, as shown in Figure 6.10-7.

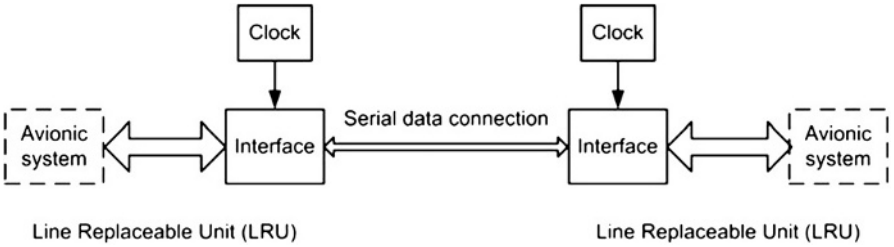


Figure 6.10-5 A simple system for serial data transfer between two avionic systems.

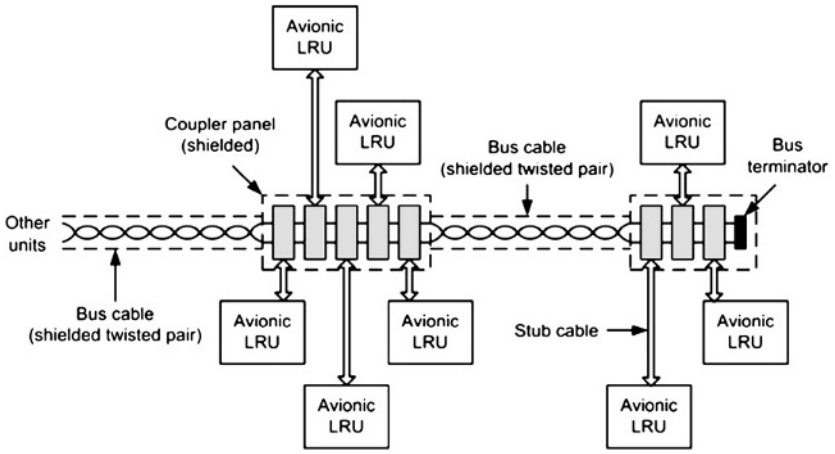


Figure 6.10-6 A practical aircraft data bus.

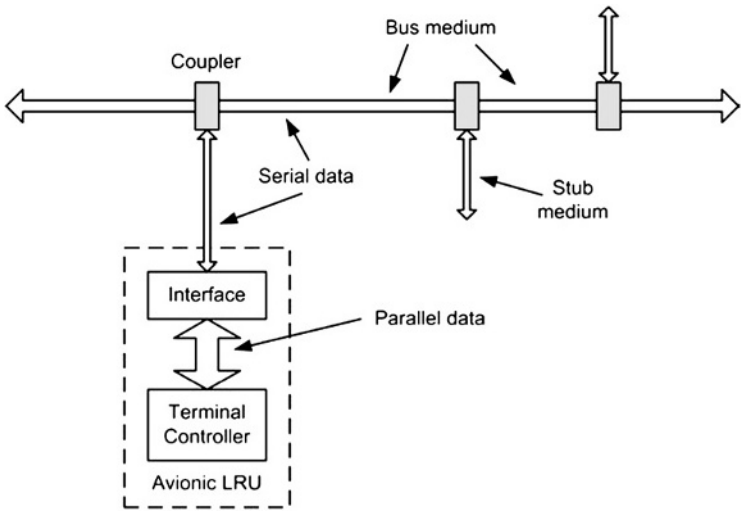


Figure 6.10-7 A basic bus interface.

In addition to providing an electrical interface (with appropriate voltage and current level shifting) the interface unit also converts the data formats (e.g. from serial analog doublets present in the stub cable to Manchester-encoded serial data required by the Terminal Controller) as shown in Figure 6.10-8.

In order to transmit data using the serial data bus information must be presented in a standard format. A typical format for serial data would use a word length of 32 bits. This word comprises several discrete fields including:

- Up to 20 bits for data (which may be further divided)
- An 8-bit label field which is used to identify the data type and any parameters that may be associated with it
- A source/destination identifier (SDI).

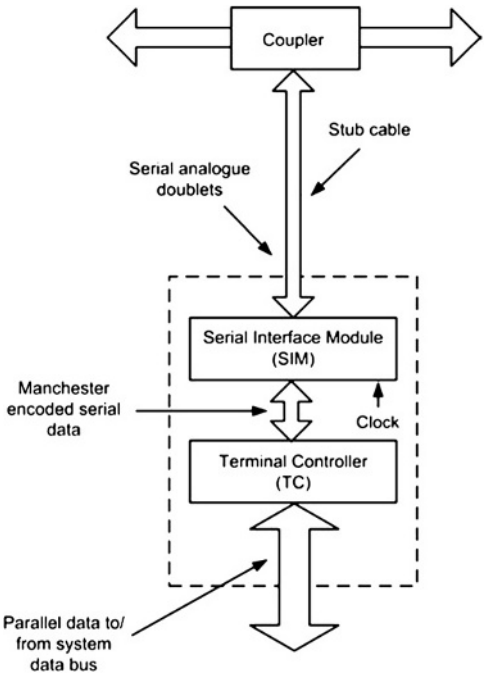


Figure 6.10-8 Data formats in a practical bus interface.

- A number of status bits used to provide information about the mode, hardware condition, or validity of data
- An added parity bit which provides a means of validating the data (i.e. determining whether or not it is free from error).

Key Point

A means of converting serial data to parallel data (and vice versa) is required whenever an LRU is to be interfaced to an aircraft bus system.

6.10.2 ARINC 429

The ARINC 429 data bus has proved to be one of the most popular bus standards used in commercial aircraft. The ARINC 429 specification defines the electrical and data characteristics and protocols that are used.

ARINC 429 employs a unidirectional data bus standard known as Mark 33 Digital Information Transfer System (DITS). Messages are transmitted in packets of 32-bits at a bit rate of either 12.5 or 100 kilobits per second (referred to as low and high bit rate respectively). Because the bus is unidirectional, separate ports, couplers and cables will be required when an LRU wishes to be able to both transmit and receive data. Note that a large number of bus connections may be required on an aircraft that uses sophisticated avionic systems.

ARINC 429 has been installed on a wide variety of commercial transport aircraft, including Airbus A310/A320 and A330/A340; Boeing 737, 747, 757, and 767; and McDonnell Douglas MD-11. More modern aircraft (e.g. Boeing 777 and Airbus A380) use significantly enhanced bus specifications in order to reduce the weight and size of cabling and to facilitate higher data rates than are possible with ARINC 429. Despite these moves to faster, bidirectional bus standards, the ARINC 429 standard has proved to be highly reliable and so is likely to remain in service for many years to come.

6.10.2.1 Electrical Characteristics

ARINC 429 is a two-wire differential bus which can connect a single transmitter or source to one or more receivers or sinks. Two speeds are available, 12.5 kbps (bits per second) and 100 kbps. The data bus uses two signal wires to

transmit 32-bit words. Transmission of sequential words is separated by at least four bit times of NULL (zero voltage). This eliminates the need for a separate clock signal and it makes the system **self-clocking**.

Key Point

Aeronautical Radio Inc. (ARINC) is an organization composed of major airlines and aircraft manufacturers which seeks to promote standardization within aircraft equipment. More information on ARINC and aircraft standards can be obtained from www.arinc.com

The ARINC 429 electrical characteristics are summarised below:

Voltage levels:	+5 V, 0 V, −5 V (each conductor with respect to ground) +10 V, 0 V, −10 V (conductor A with respect to conductor B)
Data encoding:	Bi-Polar Return to Zero
Word size:	32 bits
Bit rate (high):	100 Kbps
Bit rate (low):	12.5 Kbps
Slew rate (high):	1.5 ms (±0.5 ms)
Slew rate (low):	10 ms (±5 ms)

The nominal transmission voltage is 10 V ±1 V between wires (differential), with either a positive or negative polarity. Therefore, each signal leg ranges between +5 V and −5 V. If one conductor is at +5 V, the other is conductor is at −5 V and vice versa. One wire is called the 'A' (or '+'or 'HI') conductor and the other is called the 'B' (or '−' or 'LO') wire. The modulation employed is **bipolar return to zero** (BPRZ) modulation (see Fig. 6.10-9). The composite signal state may be at one of the following three levels (measured between the conductors):

- **HI** which should be within the range +7.25 V to 11 V (A to B)
- **NULL** which should be within the range +0.5 V to −0.5 V (A to B)
- **LO** which should be within the range −7.25 V to −11 V (A to B).

The received voltage on a serial bus depends on line length and the number of receivers connected to the bus. With ARINC 429, no more than 20 receivers should be connected to a single bus. Since each bus is unidirectional, a system needs to have its own transmit bus if it is required to respond to or to send messages. Hence, to achieve bidirectional data transfer it is necessary to have two separate bus connections.

6.10.2.2 Protocol

Since there can be only one transmitter on a twisted wire pair, ARINC 429 uses a very simple, point-to-point protocol. The transmitter is continuously sending 32-bit data words or is placed in the NULL state. Note that although there may only be one receiver on a particular bus cable, the ARINC specification supports up to 20.

6.10.2.3 Bit Timing and Slew Rate

The slew rate refers to the rise and fall time of the ARINC waveform. Specifically, it refers to the amount of time it takes the ARINC signal to rise from the 10% to the 90% voltage amplitude points on the leading and trailing edges of a pulse. The data shown in Table 6.10-1 applies to the high- and low-speed ARINC 429 systems.

Test your understanding 6.10.1

1. Explain the difference between serial and parallel methods of data transfer.
2. Explain why serial data transfer is used for aircraft bus systems.
3. Explain the function of each of the following bus system components:
   (a) Bus cable
   (b) Stub cable
   (c) Bus terminator
   (d) Coupler panel.
4. State the voltage levels present on the twisted pair conductors in an ARINC 429 data bus.

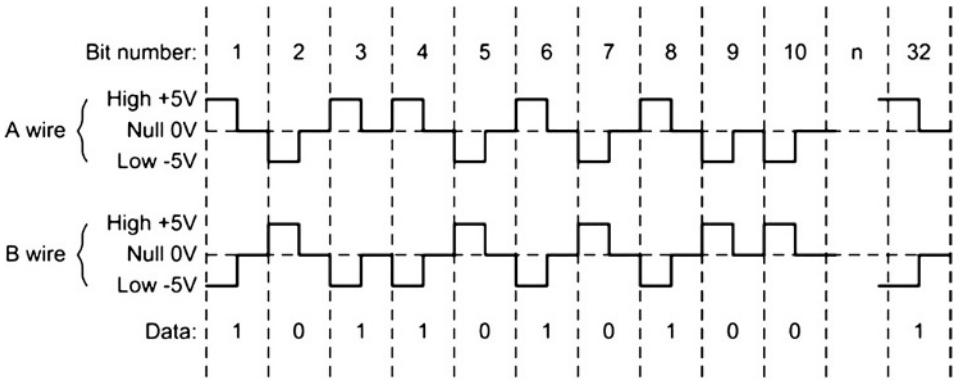


Figure 6.10-9 Signals present on the twisted pair conductors in the ARINC 429 aircraft data bus.

Table 6.10-1 ARINC 429 parameters.

Parameter	High-speed	Low-speed
Bit rate	100 Kbps	12.5 to 14.5 Kbps
Bit time (*Y*)	10 μs ± 5%	1/bit rate μs ± 5%
High time (*X*)	5 μs ± 5%	*Y*/2 μs ± 5% (see Fig. 6.10-10)
Rise time	1.5 μs ± 0.5μs	10 μs ± 5 μs
Fall time	1.5 μs ± 0.5μs	10 μs ± 5 μs

6.10.2.4 ARINC 429 Data Word Format

In most cases, an ARINC message consists of a single 32-bit data word (see Fig. 6.10-11). The 8-bit label field defines the type of data that is contained in the rest of the word. ARINC data words are always 32 bits and typically include five primary fields, namely Parity, SSM, Data, SDI, and Label. ARINC convention numbers the bits from 1 (LSB) to 32 (MSB). A number of different data formats is possible.

Bits are transmitted starting with bit 1 of the label and the final bit transmitted is the parity bit. The standard specifies the use of **odd parity** (the parity bit is set to 1 or reset to 0 in order to ensure that there is an odd number of 1s in each transmitted word). It is worth noting that the label is transmitted with the most significant bit (MSB) first while the data is transmitted least significant bit (LSB) first.

The **Label field** is an octal value that indicates the type of data (e.g. airspeed, altitude, etc) that is being transmitted.

The **SDI field** is used when a transmitter is connected to multiple receivers but not all data is intended for use by all the receivers. In this case each receiver will be assigned an SDI value and will look only at labels which match its SDI value. While the specification calls for SDI 00 to be universally accepted this may not actually be the case.

The **Data field** contains the actual data to be sent. The principal data formats defined in the specification are Binary Coded Decimal (BCD) which uses groups of four bits to contain a single decimal digit and BNR which is binary coding. For both of these data types, the specification defines the units, the resolution, the range, the number of bits used and how frequently the label should be sent.

The **SSM field** is used for information which assists the interpretation of the numeric value in the data field. Examples of SSM values might be North, East, South, West, Plus, Minus, Above or Below.

The **P field** is the parity bit. ARINC 429 uses odd parity. The parity bit is the last bit transmitted within the data word.

Some examples of data sent over an ARINC 429 bus are shown in Figures 6.10-12 and 6.10-13. In Figure 6.10-12 a **BCD** word is being transmitted whilst in Figure 6.10-13 the data is encoded in binary format. The ARINC 429 binary specification calls for the use of two's complement notation to indicate negative numbers and this binary format is known as **BNR**.

In Figure 6.10-13 the label (103) corresponds to Selected Airspeed and the indicated value is 268 knots (256 + 8 + 4). The zero in bit-29 position indicates a positive value and the data (presented in natural binary format) uses bit-28

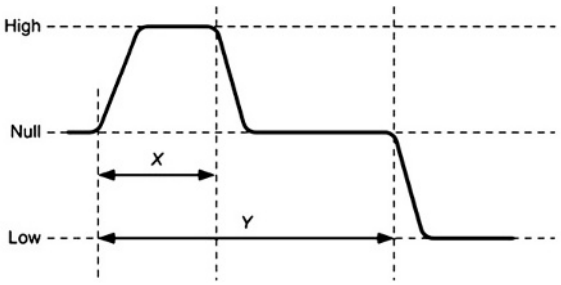


Figure 6.10-10 Timing diagram for logic level transitions in the ARINC 429 data bus.

32	31	30	29 ... 11	10	9	8 ... 1
P	SSM		MSB DATA LSB	SDI		LABEL

Figure 6.10-11 Basic ARINC 429 data word format (note the total length is 32 bits).

32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8 ... 1
P	SSM		CHAR 1			CHAR 2				CHAR 3				CHAR 4				CHAR 5				SDI		LABEL

(a) BCD word format

32	31	30	29	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10	9	8 ... 1
P	SSM		0	1	0	0	1	0	1	0	1	1	1	1	0	0	0	0	1	1	0	SDI		LABEL
	0	0	2			5				7				8				6						

(b) BCD word example

Figure 6.10-12 ARINC 429 BCD data word format.

(for the MSB) to bit-20 (for the LSB). The remaining bits are padded with zeros. In Figure 6.10-12 the BNR data conveys a value (this time expressed in BCD format) of 25786.

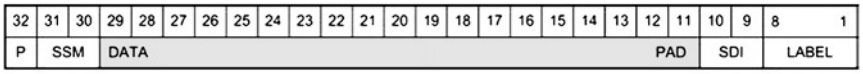
Tables 6.10-2 and 6.10-3 provide some examples of labels and equipment ID (SDI).

Test your understanding 6.10.2

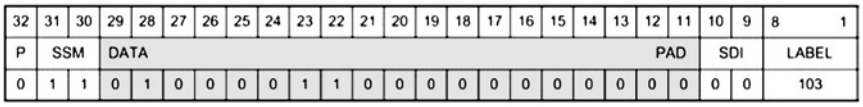
1. Explain the difference between BNR and BCD encoding.
2. Using BNR encoding, which bit in an ARINC 429 data word indicates whether the data is positive or negative?
3. What is the largest positive value that can be transmitted using a single ARINC 429 data word using:
   (a) BNR encoding
   (b) BCD encoding.
4. In relation to an ARINC 429 data word, explain the function of:
   (a) the SDI field
   (b) the SSM field.

6.10.3 Other Bus Standards

The following is a brief summary of some of the other aircraft data bus systems that have appeared over the last 40 years. It is important to note that these often describe enhancements to existing standards. However, in all cases, the main aim is that of ensuring that equipment manufacturers and operators are working to a common specification which ensures that hardware and software is both interoperable and upgradeable.



(a) BNR encoding



(b) BNR encoding example

Figure 6.10-13 Basic ARINC 429 BNR word format (note the total length is 32 bits).

ARINC 419

The ARINC 419 standard describes several digital transmission standards that predate ARINC 429. Some of these used 32-bit words similar to ARINC 429. Some standards were based on the use of a six-wire system whilst others used a shielded two-wire twisted pair (like ARINC 429) or a coaxial cable. Line voltage levels were either two state (HI/LO) or three state (HI/NULL/LO) with voltages ranging from 10 V to 18.5 V for the high state and from less than 1 V to less than 5 V for the NULL state (Fig. 6.10-14).

ARINC 561

ARINC 561 was based on a six-wire system involving three pairs that were used for DATA, SYNC, and CLOCK. Non return to zero (NRZ) encoding was employed with logic 1 represented by 12 V. Like ARINC 429, the word length was 32 bits with bits 32 and 31 comprising the SSM and no parity available. The remaining fields include an 8-bit label and six BCD fields, 5 of 4 bits and 1 of 2 bits. This system was widely used in aircraft manufactured prior to about 1970. ARINC 568 uses the same electrical interface specification as used in ARINC 561.

ARINC 573

ARINC 573 is the standard adopted for use with Flight Data Recorders (FDRs) which use a continuous data stream of Harvard Bi-Phase encoded 12-bit words. These words are encoded into which are encoded into **frames** which contain a snapshot of the data from each of the avionics subsystems on the aircraft. Each frame comprises four subframes. A unique synchronising word appears at the start of each subframe. **ARINC 717** supersedes ARINC 573 and caters for a number of different bit rates and frame sizes.

ARINC 575

Similar to ARINC 429, this standard is a low speed bus system that is based on a single twisted pair of wires. Due to the low data rate supported, this bus standard is now considered obsolete. Electrically, ARINC 575 is generally compatible with low speed ARINC 429. However, some variants of ARINC 575 use a bit rate that is significantly slower than ARINC 429 and they may not be compatible in terms of the electrical specification and data formats.

Table 6.10-2 Examples of ARINC 429 label codes.

Label (octal)	Transmitted code (binary)	Equip. ID (hex)	Parameter transmitted	Data format
140	01 100 000	001	Flight Director – Roll	BNR
		025	Flight Director – Roll	BNR
		029	Pre-cooler Output Temperature	BNR
		05A	Actual Fuel Quantity Display	BCD
141	01 100 001	001	Flight Director – Pitch	BNR
		025	Flight Director – Pitch	BNR
		029	Pre-cooler Input Temperature	BNR
		05A	Pre-selected Fuel Quantity Display	BCD
142	01 100 010	002	Flight Director – Fast/Slow	BNR
		003	Flight Director – Fast/Slow	BNR
		025	Flight Director – Fast/Slow	BNR
		05A	Left Wing Fuel Quantity Display	BCD

Table 6.10-3 Examples of equipment IDs.

Equipment ID (hex)	Equipment type
001	Flight Control Computer
002	Flight Management Computer
003	Thrust control computer
004	Inertial Reference System
005	Attitude and Heading System
006	Air Data System
007	Radio Altimeter
025	Electronic Flight Instruments
026	Flight Warning Computer
027	Microwave Landing System
029	ADDCS and EICAS
02A	Thrust Management Computer

ARINC 615

ARINC 615 is a **software protocol** that can be layered on top of ARINC 429 compatible systems. ARINC 615 supports high speed data transfer to and from on-board digital systems permitting, for example, reading and writing of 3½ inch disks.

ARINC 629

ARINC 629 was introduced in the mid-1990s and it supports a data rate of 2 Mbps (20 times faster than ARINC 429). The bus supports 120 connected devices and is currently used on the Boeing 777, Airbus A330 and A340 aircraft. A notable enhancement of the earlier ARINC 429 standard is that ARINC 629 is a bidirectional bus system (in other words, connected devices can transmit, receive or do both). Another advantage of ARINC 629 is that it achieves bidirectional bus communication without the need for a bus controller (BC; which could be a potential source of single-point failure). The physical bus medium is STP.

ARINC 708

ARINC 708 is used to transfer data from the airborne weather radar receiver to the aircraft's radar display. The bus is unidirectional and uses Manchester encoded data at a data rate of 1 Mbps. Data words are 1600 bits long and they are composed of one 64-bit status word and 512, 3-bit data words.

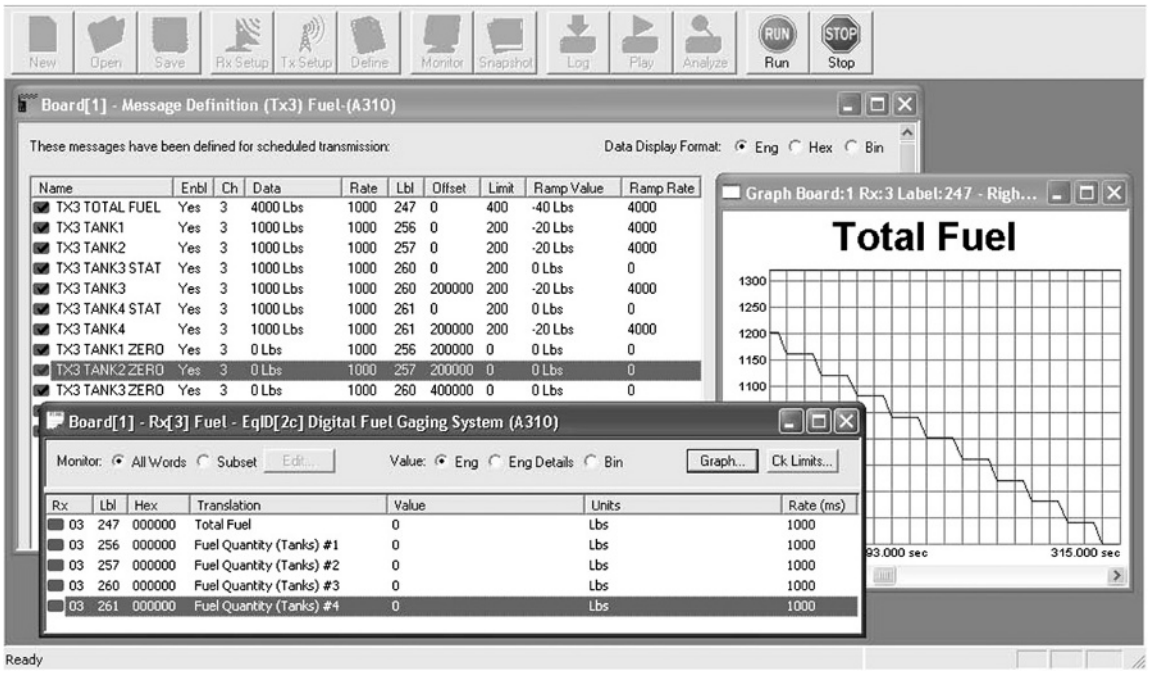


Figure 6.10-14 Using a bus analyzer (Bus Tools from Condor Engineering) to examine the data present on an ARINC 429 data bus and display the rate at which fuel is being used.

MIL-STD-1553B/1773B

Military standard 1533B is a bidirectional centrally controlled data bus designed for use in military aircraft. The standard uses a BC which can support up to 31 devices which are referred to as Remote Terminals (RTs). The standard supports a bit rate of 1 Mbps. MIL-STD-1773B is a fibre optic implementation of MIL-STD-1553B that provides significantly greater immunity to exposure to high intensity radiated electromagnetic fields (HIRFs).

CSDB and ASCB

The CSDB and ASCB standards are proprietary protocols from Collins and Honeywell respectively. These systems are often used in small business and private general aviation (GA) aircraft. CSDB is a unidirectional bus that permits the connection of up to 10 receivers and one transmitter. The standard supports data rates of 12.5 kbps and 50 kbps. ASCB is a centrally controlled bidirectional bus. A basic configuration comprises a single BC and two isolated buses, each of which can support up to 48 devices.

FDDI

The Fibre Distributed Data Interface (FDDI) was originally developed by Boeing for use on the Boeing 777 aircraft. FDDI is a local area network (LAN) based on a dual token ring topology. Data in each ring flows in opposite directions. The data rate is 100 Mbps and data is encoded into frames. CDDI and SDDI are similar network bus standards based on copper (Copper Distributed Data Interface) and shielded twisted pair (SDDI) as the physical media. The data format is NRZI (a data format similar to NRZ but where a change in the line voltage level indicates a logic 1 and no change indicates a logic 0). For reasons of cost and in order to reduce the number and complexity of network standards used in its aircraft, Boeing now plans to replace the system on the 777 with a less-expensive 10 Mbps copper Ethernet.

Test your understanding 6.10.3

Which of the listed bus standards would be most suitable for each of the following applications? Give reasons for your answers.

1. A low cost bus system for the simple avionics fitted to a small business aircraft.
2. Connecting a weather radar receiver to a radar display.
3. A bus system for linking the various avionics systems of a modern passenger aircraft to its FDR.
4. A bus system for the avionics of a military aircraft fitted with multiple radars and electronic counter measures (ECMs).

6.10.4 Multiple Choice Questions

1. A bus that supports the transfer of data in both directions is referred to as
   (a) universal
   (b) bidirectional
   (c) asynchronous.

2. The main advantage of using a serial bus in an aircraft is
   (a) there is no need for data conversion
   (b) it supports the highest possible data rates
   (c) reduction in the size and weight of cabling.

3. Which one of the following is used to minimise reflections present in a bus cable?
   (a) coupler panels
   (b) bus terminators
   (c) STP cables.

4. The data format in an ARINC 429 stub cable consists of
   (a) serial analog doublets
   (b) parallel data from the local bus
   (c) Manchester encoded serial data.

5. In order to represent negative data values, BNR data uses
   (a) two's complement binary
   (b) BCD data and a binary sign bit.
   (c) an extra parity bit to indicate the sign.

6. The physical bus media specified in ARINC 629 is
   (a) fiber optic
   (b) coaxial cable
   (c) STP.
7. The validity of an ARINC 429 data word is checked by using
   (a) a parity bit
   (b) a checksum
   (c) multiple redundancy.
8. The label field in an ARINC 429 data word consists of
   (a) five bits
   (b) three bits
   (c) eight bits.
9. The voltages present on an ARINC 429 data bus cable are
   (a) ±5 V
   (b) ±15 V
   (c) +5 V and +10 V
10. The maximum bit rate supported by ARINC 429 is
    (a) 12.5 kbps
    (b) 100 kbps
    (c) 1 Mbps.
11. An ARINC 429 NULL state is represented by
    (a) a voltage of −5 V
    (b) a voltage of 0 V
    (c) a voltage of +5 V.
12. The maximum data rate supported MIL-STD-1553 is
    (a) 12.5 kbps
    (b) 100 kbps
    (c) 1 Mbps.
13. A bus that is self-clocking is referred to as
    (a) universal
    (b) bidirectional
    (c) asynchronous.
14. The physical bus media specified in MIL-STD-1773B is
    (a) fibre optic
    (b) coaxial cable
    (c) STP.
15. The length of an ARINC 429 word is
    (a) 16 bits
    (b) 20 bits
    (c) 32 bits.
16. ARINC 573 is designed for use with
    (a) INS
    (b) FDR
    (c) weather radar.
17. The maximum data rate supported by the FDDI bus is
    (a) 1 Mbps
    (b) 10 Mbps
    (c) 100 Mbps.

Aerospace Engineering Desk Reference; ISBN 9781856175753

6.11 Software

Mike Tooley

Aircraft software is something that you can't see and you can't touch yet it must be treated with the same care and consideration as any other aircraft part. This chapter deals with aspects of the maintenance of the safety-critical software found in modern aircraft.

At the outset it is important to realise that software encompasses both the executable code (i.e. the programs) run on aircraft computers as well as the data that these programs use. The term also covers the operating systems (i.e. system software) embedded in aircraft computer systems. All of these software parts require periodic upgrading as well as modification to rectify problems and faults that may arise as a result of operational experience (see Fig. 6.11-1 for an example emphasising the importance of this).

The consequences of software failure can range from insignificant (no effect on aircraft performance) to catastrophic (e.g. major avionic system failure, engine faults, etc). Because of this it is important that you should have an understanding of the importance of following correct procedures for software modification and upgrading. Once again, this is an area of rapidly evolving technology which brings with it many new challenges.

6.11.1 Software Classification

Aircraft software can be divided into five levels according to the likely consequences of its failure, as shown in Table 6.11-1. The highest level of criticality (Level A) is that which would have catastrophic consequences whilst the lowest level of criticality is that which would have no significant impact on the operation of the aircraft. In between these levels the degree of criticality is expressed in terms of the additional workload imposed on the flight crew and, in particular, the ability of the flight crew to manage the aircraft without having access to the automatic control/or flight information that would have otherwise been provided by the failed software. Table 6.11-2 provides examples of software applications and level of software criticality associated with each.

6.11.2 Software Certification

The initial certification of an aircraft requires that the Design Organisation (DO) shall provide evidence that the software has been designed, tested and integrated with the associated hardware in a manner that satisfies standard DO-178B/ED-12B (or an agreed equivalent standard). In order to provide an effective means of software identification and change control, a software configuration management plan (CMP) (e.g. as defined in Part 7 of DO-178B/ED-12B) is required to be effective throughout the life of the equipment (the CMP must be devised and maintained by the relevant DO).

Post-certificate modification of equipment in the catastrophic, hazardous, or major categories (Levels A, B, C and D; see Tables 6.11-1 and 6.11-2) must not be made unless first approved by the DO. Hence all software upgrades and modifications are subject to the same approval procedures as are applied to hardware modifications. This is an important point that recognises the importance of software as an 'aircraft part'. Any modifications made to software must be identified and controlled in accordance with the CMP. Guidance material is provided DO-178B.

Test your understanding 6.11.1

Explain the purpose of a software CMP and specify the standard that applies to the development and management of safety critical aircraft software.

Key Point

Aircraft software can be taken to encompass both the executable code (i.e. programs) as well as the data that these programs use. The term also covers the operating systems (i.e. system software) embedded in aircraft computer systems. All of these software parts require periodic upgrading as well as modification to rectify problems and faults that may arise.

ATA 73—FUEL AND CONTROL
EEC SOFTWARE—MODIFICATION

There have been six events of in-service loss of engine parameters displayed in the aircraft cockpit, combined with freezing of engine power setting for the affected engine. Subsequent investigation has established the cause to be a fault in engine software. A safety assessment has identified that this could result in hazardous asymmetric thrust, if the event were to occur during a take-off roll and in response the crew attempted to abort the take-off. This Airworthiness Directive requires the introduction of revised engine software.

Figure 6.11-1 An extract from a recent Airworthiness Directive identifying the need for software modification.

Table 6.11-1 Levels of failure.

Level	Type of failure	Failure description	Probability	Likelihood of failure (per flight hour)
A	Catastrophic failure	Aircraft loss and/or fatalities	Extremely improbable	Less than 10^{-9}
B	Hazardous/severe major failure	Flight crew cannot perform their tasks; serious or fatal injuries to some occupants	Extremely remote	Between 10^{-7} and 10^{-9}
C	Major failure	Workload impairs flight crew efficiency; occupant discomfort including injuries	Remote	Between 10^{-5} and 10^{-7}
D	Minor failure	Workload within flight crew capabilities; some inconvenience to occupants	Probable	Greater than 10^{-5}
E	No effect	No effect	Not applicable	

Table 6.11-2 Examples of software levels.

Level	Typical aircraft applications
A	AHRS GPS/ILS/MLS/FLS SATNAV VOR ADF
B	TCAS ADSB Transponder Flight Displays
C	DME VHF voice communications
D	AHRS Automatic Levelling CMC/CFDIU Data Loader Weather Radar
E	In-flight entertainment

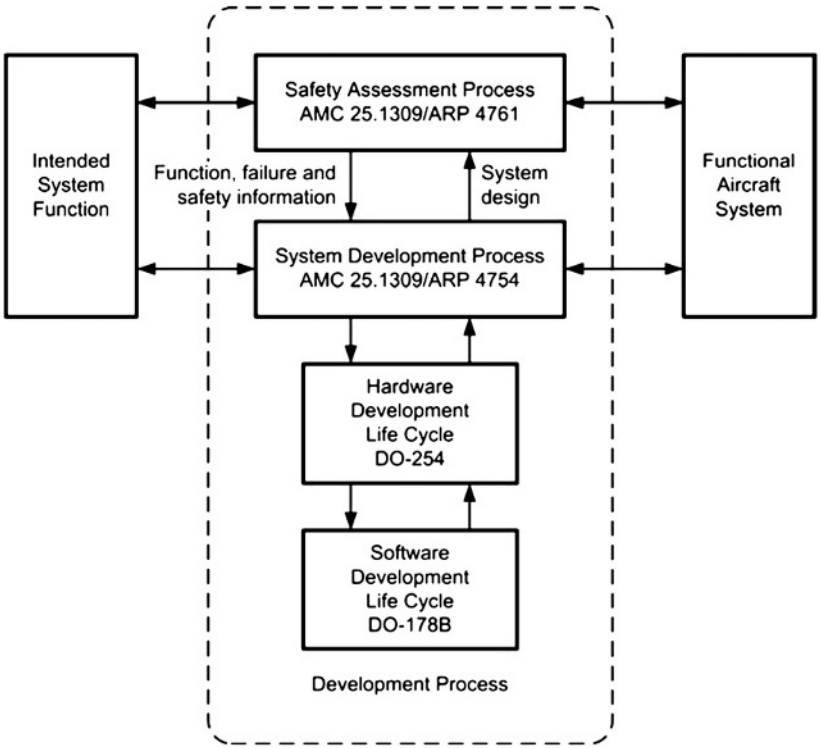


Figure 6.11-2 The software development process.

Key Point

DO-178B is the de-facto standard for the development and management of safety critical software used in aviation. DO-178B certification deals with five levels of criticality (A to E). Level A is the most critical (corresponding to the possibility of catastrophic failure) whilst Level E is the least critical (corresponding to a negligible impact on the aircraft and crew).

The relationship between the development of aircraft hardware and software is shown in Figure 6.11-2. Note that the two life cycles (hardware and software) are closely interrelated simply because a change in hardware configuration inevitably requires a corresponding change to the software configuration. The Safety Assessment Process (SAP) is a parallel activity to that of the System Development Process. It is important to be aware that changes to the system design and configuration will always necessitate a reappraisal of safety factors.

Testing takes place throughout the development process. Independent tests are usually carried out in order to ensure that the results of tests are valid. Testing generally also involves simulation of out-of-range inputs and abnormal situations such as recovery from power failure (ensuring that a system restart is accomplished without generating dangerous or out-of-range outputs).

Traceability of software is a key component of the DO-178B criteria. Planning documents and evidence of traceability help to ensure that not only are certification requirements met but also that the final code contains all of the required modules and that each module is the most recently updated version. Care is also needed to ensure that none of the final code will be detrimental to the overall operation of the system (for example, seeking data from sensors and transducers that may not be fitted in some configurations of a particular aircraft).

6.11.3 Software Upgrading

When considering software modifications and upgrades, it is important to distinguish between executable code (i.e. computer programs) and the data that is used by programs but is not, in itself, executable code. Both of these are commonly referred to as 'software' and both are likely to need modification and upgrading during the life of an aircraft.

6.11.3.1 Field Loadable Software

Field Loadable Software (FLS) is executable code (i.e. computer programs) that can be loaded into a computer system whilst the system is in place within the aircraft. FLS can be loaded onto an aircraft system by a maintenance mechanic/technician in accordance with defined maintenance manual procedures.

There are three main types of FLS: Loadable Software Aircraft Parts (LSAPs), User Modifiable Software (UMS), and Option Selectable Software (OSS).

Loadable software aircraft parts

An LSAP is software that is required to meet a specific airworthiness or operational requirement or regulation. LSAP is not considered to be part of the aircraft approved design and therefore it is an aircraft part requiring formal (controlled) release documentation.

Typical examples of target hardware for LSAP (FLS) include:

- Electronic Engine Controls (EECs)
- Digital Flight Data Acquisition Units (DFDAUs)
- Auxiliary Power Unit's Electronic Control Units (ECUs)
- Flight Guidance Computers (FGCs).

User modifiable software

UMS is declared by the aircraft Type Certificate holder's DO (or Supplementary Type Certificate holder's DO) as being intended for modification within the constraints established during certification. UMS can usually be upgraded by the aircraft operator, DO, or equipment manufacturer, without further review by the licensing authority. Typical examples of target hardware for UMS include:

- Aircraft Condition Monitoring Systems (ACMSs)
- In-Flight Entertainment Systems (IFEs).

Option selectable software

OSS is software that contains approved and validated components and combinations of components that may be activated or modified by the aircraft operator within boundaries defined by the Type Certificate or Supplementary Type Certificate holder. Typical examples of target hardware for OSS can be found in Integrated Modular Avionics (IMA) units.

6.11.3.2 Database Field Loadable Data

Database Field Loadable Data (DFLD) is data that is field loadable into target hardware databases. Note that it is important to be aware that the database itself is an embedded item that resides within the target hardware and is not, itself, field loadable and that the process of 'loading a database' is merely one of writing new data or over-writing old data from a supplied data file.

DFLD is usually modified or updated by overwriting it using data from a data file which is field loaded. The data file can contain data in various formats, including natural binary, binary coded decimal, or hexadecimal formats. Of these, natural binary code produces the most compact and efficient data files but the data is not readable by humans and, for this reason, binary coded decimal or hexadecimal formats are sometimes preferred.

It is important to note that the updating of an aircraft database will usually have an impact on aspects of the aircraft's performance. However, if the data used by a program is invalid or has become corrupt, this may result in erratic or out of range conditions. Because of this, it is necessary to treat DFLD in much the same manner as the executable code or LSAP that makes use of it. Hence a DFLD must be given its own unique part number and release documentation.

Typical examples of the target hardware with databases that can be field loaded with DFLD (and that need to be tracked in the same manner as other aircraft parts) include:

- Flight Management Computers (FMCs)
- Terrain Awareness Warning System (TAWS) Computers
- IMA units.

6.11.3.3 Distribution Methods

FLS and DFLD can be distributed by various methods including combinations of the following methods:

- Media distribution. A process whereby FLS or data files are moved from the production organisation or supplier to a remote site using storage media such as floppy disk, a PCMCIA (Personal Computer Memory Card International) card, a CD-ROM, or an Onboard Replaceable Module (OBRM)
- Electronic transfer. A process where a laptop, hand-held computer or Portable Data Loader (PDL) is used to transfer data using a serial data link or temporary bus connection
- Electronic distribution. A process whereby FLS or DFLD are moved from the producer or supplier to a remote site without the use of intermediate storage media, such as floppy disk or CD-ROM.

Note that the method of release is dependent upon whether the FLS or DFLD is required to meet a specific airworthiness or operational requirement, or certification specification. For FLS or DFLD that does not need to meet a specific airworthiness, operational or certification requirement, a Certificate of Conformity is normally sufficient. In other cases an EASA Form 1 or FAA 8130-3 should accompany any FLS (executable code) that is required to meet a specific airworthiness or operational requirement or regulation, or certification specification, i.e. LSAP. Examples of LSAP that would require such release would include EECs, DFDAUs, ECUs, FGCs, and IMA units.

An EASA Form 1 or FAA 8130-3 should accompany any DFLD that is required to meet a specific airworthiness or operational requirement or regulation, or certification specification. Examples of DFLD that require such release include FMCs, TAWS Computers, IMA units. A 'Letter of Acceptance' or equivalent should accompany the release of any navigational database's DFLD because an EASA Form 1 or FAA 8130-3 cannot be provided.

By virtue of the speed of distribution and the removal of the need for any physical transport media (which can be prone to data corruption), electronic distribution is increasingly being used to transfer FLS or DFLD from the supplier to an operator. Operators should maintain a register which provides the following information:

- The current version of the FLS and DFLD installed
- Which aircraft the FLS and DFLD are installed on
- The aircraft, systems and equipment that they are applicable to
- The functions that the recorded FLS or DFLD performs
- Where (including on or off aircraft location), and in what format it is stored (i.e. storage media type), the name of the person who is responsible for it, and the names of those who may have access to it
- Who can decide whether an upgrade is needed and then authorise that upgrade
- A record of all replicated FLS/DFLD, traceable to the original source.

When transferring FLS or a database, it is essential to ensure that:

- The FLS or DFLD has come from an appropriate source
- Effective configuration control processes are in place to ensure that only the correct data and/or executable code will be supplied
- The FLS or DFLD is accompanied by suitable release documentation and records are kept.
- Suitable controls are in place to prevent use of FLS and DFLD that have become corrupted during its existence in any 'open' environment, such as on the Internet or due to mishandling in transit
- Effective data validation and verification procedures are in place
- The FLS and DFLD as well as the mechanisms for transferring them (file transfer and file compaction utilities) are checked for unauthorised modification (for example, that caused by malicious software such as viruses and 'spyware').

Key Point

If there is any doubt about the status of an FLS or DFLD upgrade (for example, if there is a reason to suspect that the data may be corrupt or may not be the correct version for the particular aircraft configuration), no attempt should be made to transfer the data. Instead, it should be quarantined and held pending data integrity and/or version checks specified by the supplier.

Special precautions are needed when the FLS or DFLD is transferred or downloaded using the public Internet. These precautions include the use of an efficient firewall to prevent unauthorised access to local servers and storage media as well as accredited virus protection software. Further precautions are necessary when making backup copies of transferred or downloaded data. In particular, backup copies should be clearly labelled with version and date information and they must always be stored in a secure place.

Key Point

When an FLS or DFLD is transferred over an 'open environment' (such as the Internet) extra care will be needed to ensure that the software has not become corrupt or contaminated by viruses. The supplier will usually be able to provide guidance on approved methods of checking software and data integrity.

Test your understanding 6.11.2

1. Explain what is meant by an LSAP.
2. Give TWO examples of an LSAP.

Test your understanding 6.11.3

State FOUR precautions that must be taken when transferring FLS.

Test your understanding 6.11.4

Explain why DFLD should be treated with the same care and consideration as FLS.

A typical software loading procedure for the EEC of a modern passenger aircraft (see Fig. 6.11-3) is as follows:

1. Make sure the electrical power to the EEC is off
2. Connect one end of the PDL cable to the connector of the loader (see Fig. 6.11-4)
3. Disconnect the electrical connector from the EEC
4. Connect the other end of the PDL cable to the connector of the EEC
5. Connect the 28 VDC power source to the Brown (+), Black (−), and shield (chassis ground) terminals of the loader cable
6. Load the software using the following steps:

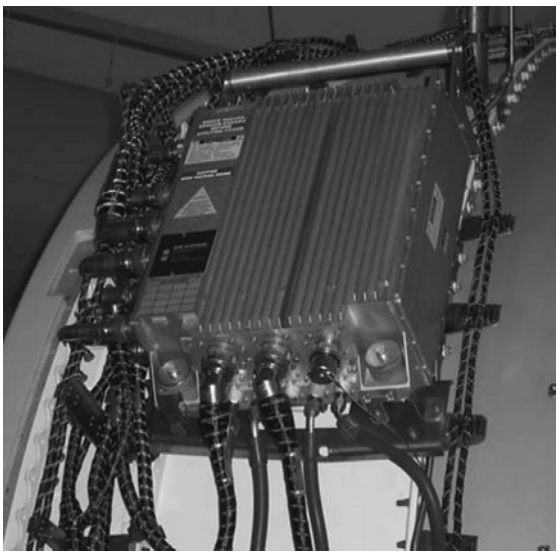


Figure 6.11-3 Electronic Engine Control (EEC) requiring Field Loadable Software (FLS).

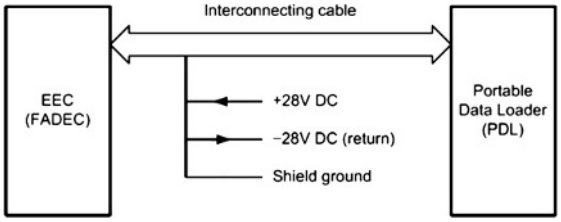


Figure 6.11-4 Typical Portable Data Loader (PDL) arrangement.

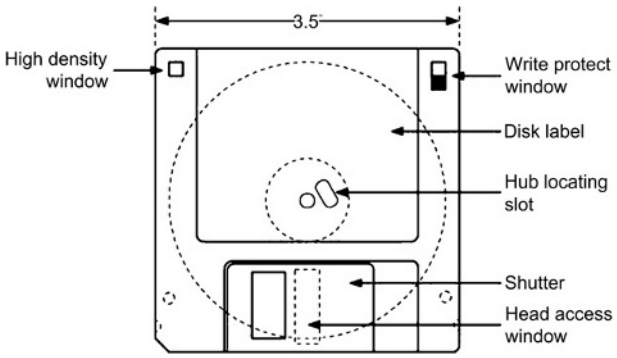


Figure 6.11-5 Floppy disk with write protect window (the window must be closed to write to the disk).

(a) Check that you have the correct disk that will be needed to load channels A and B with the non-volatile memory (NVM) and the operating program software
(b) Turn on the 28 VDC supply to the loader
(c) Turn on 115 VAC (400 Hz) to the EEC
(d) Turn on the loader without the disk inserted. The PDL display will show DISK NOT INSERTED after the PDL is initiated
(e) Make sure you use the correct EEC software version disk and also ensure that the same software version is used for both the left and right engines
(f) Check that the software disk is not write protected before you load the new software (the disk is write protected as shown in Fig. 6.11-5)
(g) Insert the appropriate disk into the PDL drive within 5 minutes after you turned on the power
(h) The PDL will automatically load the NVM and the operating program software to the two channels of the EEC. If it is successful, the PDL will display LOAD COMPLETE. It will take approximately 12 to 16 minutes to load the EEC. If the software loading has failed, it may take a few minutes before the transfer fail lamp illuminates. If the transfer fail lamp illuminates, make sure the cable connections are correct. Remove the power and repeat the software loading steps above. Note that no more than three attempts can be made if the failure reccurs
(i) If LOAD COMPLETE is displayed on the loader, turn off the loader and then turn off the 115 VAC or 28 VDC
(j) Remove the disk from the loader.

7. Remove all electrical power from the EEC and the loader
8. Remove the loader and reconnect the EEC cables
9. Use the appropriate memory verify program (included in the software reprogramming disks) to make sure that the EEC has been reprogrammed correctly. If the verify program displays PASS for the EEC part number as well as the checksums, the EEC has been programmed correctly. If FAIL is displayed because the comparison of the FADEC/EEC hardware part number did not agree with the software part number, make sure the FADEC/EEC hardware part number is correct on the nameplate. Use the PRINTSCREEN feature of the PC, and make a paper copy of the final results
10. Use a ball point pen to write the number of the appropriate software version, Service Bulletin number and date on the Service Bulletin plate (a metallic sticker) on the EEC (see Fig. 6.11-3).

6.11.4 Data Verification

Various techniques are used to check data files and executable code in order to detect errors. Common methods involve the use of **checksums** and **cyclic redundancy checks** (CRCs). Both of these methods will provide an indication that a file has become corrupt, but neither is completely foolproof.

Checksums involve adding the values of consecutive bytes or words in the file and then appending the generated result to the file. At some later time (for example when the file is prepared for loading) the checksum can be recalculated and compared with the stored result. If any difference is detected the file should not be used.

Cyclic redundancy checks involve dividing consecutive blocks of binary data in the file by a specified number. The remainder of the division is then appended to the file as a series of **check digits** (in much the same way as a checksum). If there is no remainder when the file is later checked by dividing by the same number, the file can be assumed to be free from errors.

Test your understanding 6.11.5

Classify each of the following applications in terms of level of software criticality:

1. Weather radar
2. VOR
3. In-flight entertainment.

Test your understanding 6.11.6

An FLS upgrade may have been corrupted during transfer. What action should be taken?

Test your understanding 6.11.7

Distinguish between UMS and OSS. Give a typical example of each.

Test your understanding 6.11.8

(a) Describe TWO methods of checking that a data file contains no errors.
(b) Describe the precautions that should be taken when making backups of FLS.

6.11.5 Multiple Choice Questions

1. A Level C software classification is one in which a failure could result in
   (a) aircraft loss
   (b) fatal injuries to passengers or crew
   (c) minor injuries to passengers or crew.

2. A Level B software classification is one in which the probability of failure is
   (a) extremely improbable
   (b) extremely remote
   (c) remote.

3. A software CMP must be created and maintained by
   (a) the CAA or FAA
   (b) the aircraft operator
   (c) the relevant DO.

4. Weather radar is an example of
   (a) Class B software
   (b) Class C software
   (c) Class D software.

5. EEC software is an example of
   (a) DFLD
   (b) LSAP
   (c) OSS.
6. OSS and UMS are specific classes of
   (a) FLS
   (b) LSAP
   (c) DFLD.
7. The final stage of loading EEC software is
   (a) disconnecting the PDL
   (b) verifying the loaded software
   (c) switching on and testing the system.

SECTION 7

Rotorcraft

This page intentionally left blank

Aerospace Engineering Desk Reference; ISBN 9781856175753

7.1 Introduction to Rotorcraft

John Watkinson

7.1.1 Applications of the Helicopter

Although helicopters must all follow the same laws of physics, the forms that practical machines take vary tremendously due to the range of tasks to which they can be put. To avoid confusion, this book takes the view that an aircraft is any man-made aerial machine capable of climbing out of ground effect. Thus a helicopter must be an aircraft. Neglecting aerostats (balloons etc.) any aircraft that is not a helicopter will be an aeroplane (USA: airplane).

It will be seen later in this book that the slower an aircraft goes, the more power it needs to maintain height. Hovering is the ultimate case of slow flight, suggesting that helicopters must have a high power to weight ratio. This will require heavy engines and a corresponding fuel capacity. These factors limit load carrying capability and range.

The mechanical complexity of the helicopter and the inevitable vibration demand a lot of maintenance. The airspeed of the true helicopter is forever restricted by fundamental limits. It will generally be more expensive to move a given load by helicopter than by almost any other means, and so if a suitable airstrip exists, a fixed-wing aircraft can do the job at lower cost, and generally at a higher airspeed. In most cases helicopters cannot compete economically with aeroplanes, and so their use is restricted to applications for which aeroplanes, or other forms of transport, are unsuitable.

In remote areas, there will be no airstrips; in wartime, runways are conspicuous targets and, with the exception of the aircraft carrier, are fixed. The helicopter's ability to hover means that it can land almost anywhere a fairly flat firm surface exists. Some are genuinely amphibious, landing with equal aplomb on water or land. If the ground is unsuitable (or if the waves are too high) many helicopters can transfer goods and passengers whilst in the hover.

This ability makes the helicopter the ideal rescue vehicle. Many lives have been saved because the helicopter can get to places that would otherwise be difficult or impossible to reach. War casualties, the victims of shipwrecks, mountain climbing accidents and natural disasters such as earthquake or flood today have significantly higher chances of survival because a helicopter can get them rapid treatment.

The accounts of helicopter rescues make more thrilling reading than novels because they are true. Helicopters have flown far out of range of the shore by taking fuel from oilrigs and ships, sometimes taking fuel in the hover if a landing was impossible. This would be remarkable in good weather, but emergencies occur in all weather conditions and the helicopter has evolved to handle the worst.

Despite their life-saving ability, most of today's helicopters were originally designed for military use. As military fixed-wing aircraft became faster and faster, they found it harder to attack ground targets or to support ground troops. During the Vietnam War, it was found that the helicopter had a major role. Transport helicopters excelled in inserting and withdrawing troops, delivering ammunition and food, evacuating the wounded, recovering the crews of downed planes and even the planes themselves.

Armed helicopters proved to be ideal for attacking ground targets. At first these were general-purpose machines hurriedly fitted with weapons, but later dedicated attack helicopters evolved, complete with armour plating and redundant systems to allow them to withstand ground fire. As their load carrying capability has increased, these machines have virtually rendered the tank obsolete.

Although ideal against ground targets, the helicopter is slow and vulnerable to attack from the air. Fixed-wing planes are necessary to provide the air superiority in which the helicopters operate. As an alternative, helicopters can operate under stealth conditions, avoiding detection by using terrain.

Large military helicopters are very expensive to operate, and armed forces found it worthwhile to have simpler machines specifically for training purposes. A small number of helicopters have been designed specifically for the civil market and these are popular with large companies for executive transport.

For the average private owner, the sheer cost of running helicopters precludes all but the smallest machines with aeroplane-derived piston engines. Virtually all other helicopters are now turbine powered.

7.1.2 A Short Technical Helicopter History

This is not a history book and this section must necessarily be brief. The reader interested in the US history of the helicopter is recommended to the comprehensive yet highly readable works of Jay Spenser.[1,2] The recent book by

Steve Coates, *Helicopters of the Third Reich*,[3] is essential reading to the historian as it shows how far ahead of the rest of the world German helicopter engineers were at that time. For those who read French, two more fascinating volumes are available. *L'Histoire de l' Hélicoptère* by Jean Boulet[4] contains the words of helicopter pioneers themselves. *Les Hélicoptères Florine 1920–1950* by Alphonse DuMoulin[5] recounts the pioneering work of Nicolas Florine.

The history of the helicopter has been very short indeed. In comparison with fixed-wing aircraft, helicopters need more power, have to withstand higher stresses, are harder to understand and control and have more moving parts. It is hardly surprising that the development of the helicopter took place well after that of the fixed-wing aircraft.

Early helicopters lacked enough power to fly. Once helicopters were powerful enough to leave the ground, they were found to be uncontrollable. Once the principles of control were understood, they were found to vibrate and to need a lot of maintenance and so on. Today's helicopters represent the sum of a tremendous number of achievements in overcoming one obstacle after another.

Before World War II helicopters were in an experimental phase. This was the heyday of the gyroplane, invented by the Spaniard Juan de la Cierva and technically refined by Raoul Hafner, an Austrian working in England who would later contribute much to the development of the helicopter.[6]

The first practical helicopter was the Focke-Wulf Fw-61 of 1938 (Figure 7.1-1), followed in the same year by the Weir W-5 (Figure 7.1-2) that flew two years before Sikorsky's VS-300 (Figure 7.1-3). The urgencies of war accelerated all technical development with the emergence of production helicopters, where the work of Anton Flettner (Figure 7.1-4) and Heinrich Focke was far in advance of anything taking place elsewhere.

The Focke-Achgelis Fa-223 shown in Figure 7.1-5 became, on 6 September 1945, the first helicopter to cross the English Channel when a captured machine was flown to England by Hans Gerstenhauer. This machine had a payload of 2000 kg.[7]

Helicopter development in the UK was halted by government order during World War II and, being an Austrian, Hafner was locked up, but fortunately not for long. In Germany, production was hampered by Allied bombing, whereas US helicopters were unrefined. The result was the same: helicopters made little contribution to the war itself.

After World War II great progress was made in the understanding of helicopter dynamics and stability. This led directly to machines that were less stressful to fly and correspondingly safer. The Bell 47 (Figure 7.1-6) based on the research of Arthur Young was in 1946 the first helicopter to be certified. The Sycamore of 1952 (Figure 7.1-7) designed by Hafner, was the first British helicopter to be certified and was noted not just for its performance, but also for its light control forces which needed no power assistance.[8]

Figure 7.1-1 The Focke-Wulf Fw-61 was the first helicopter to move beyond the experimental stage and was capable of extended flights. The hull was based on that of an Fw Stieglitz aeroplane. Note the vestigial airscrew that simply cools the engine. (Steve Coates.)

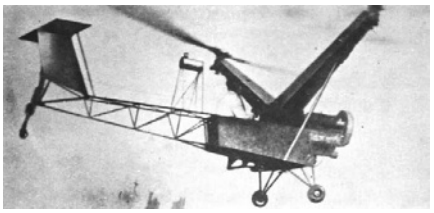


Figure 7.1-2 The Weir W-5 flew extensively, but development of all helicopters was suspended in the UK during World War II. (AugustaWestland.)

Figure 7.1-3 The Sikorsky VS-300 had an extended development period during most of which cyclic control was not understood and unwieldy auxiliary rotors were used instead. (Igor I. Sikorsky Historical Archives Inc.)

Advances in constructional techniques and materials continued to improve the service life of components, especially blades. Possibly the most significant single step was the introduction of the turbine engine which was much lighter than the piston engine for the same power, yet had fewer moving parts. This allowed greater payload and a reduction in maintenance.

The first turbine-powered helicopter to fly was a modified Kaman K-225 in 1951 and in 1954 the first twin-turbine machine, also a Kaman, flew. The Bell Huey first flew in 1956 (Figure 7.1-8). It was in the 1960s that the many disciplines in helicopter design were finally mastered allowing the machine to be considered as a system. Some elegant and definitive designs emerged during this period. Such was the validity of their basic concepts that they could accept a steady succession of upgrades that would allow them in some cases to remain in service to the present day. Sikorsky's S-61 and Boeing's Chinook are good examples of longevity.

Since that time, there have been few breakthroughs; instead there has been a steady process of refinement. The introduction of composite materials in blades, rotor heads and body parts has reduced weight and extended service

Figure 7.1-4 The Flettner Kolibri (Hummingbird) was the first synchropter and was an advanced and capable machine. Kaman developed the concept and produced the successful Husky and K-Max models. (Steve Coates.)

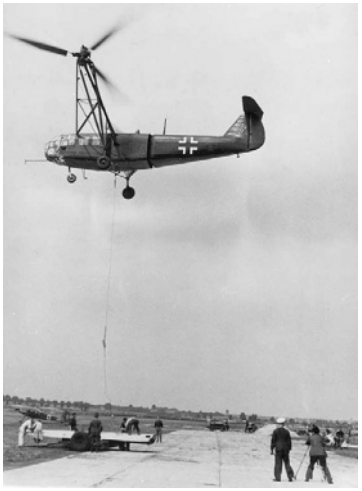


Figure 7.1-5 The Focke-Achgelis Fa-223 was a large and capable machine that was in production during World War II. Few were produced due to Allied bombing. The machine was far ahead of anything else in the world at the time. (Steve Coates.)

Figure 7.1-6 The legendary Bell 47 was based on Arthur Young's research and was the first helicopter to be certified in the USA. Larry Bell never liked the utilitarian appearance, but it outsold all of the more stylish models. The flybar stabilization system was adopted extensively in later Bell machines. (Bell Helicopter Textron.)

Figure 7.1-7 The attractive Sycamore was Hafner's masterpiece and the first helicopter to be certified in the UK. (AugustaWestland.)

life. Refinements in mechanical engineering have produced lighter engines and transmissions having longer life. Manufacturers have used production engineering to reduce the amount of labour needed to build machines.

Instead of a revolution, the employment of electronics and computers in helicopters has seen steady and relentless progress. There is no wear mechanism in electronics and complicated transfer functions can be realized in lightweight parts that use little power. Items such as turbine engine controllers and rotor rpm governors are ideal applications for electronics, along with stability augmentation systems. The flexible control systems that electronics make possible have enabled developments such as the tilt-rotor helicopter.

Helicopters can survive engine failures and the failure of a variety of parts, but there remain some parts such as gearboxes and rotor heads in which failure will be catastrophic. The technique of electronically monitoring critical

Figure 7.1-8 The Bell Huey was officially designated the Iroquois, but the Helicopter, Utility designation, HU-1, led to the nickname of Huey, which stuck. In Vietnam the Huey was used in enormous numbers. (Bell Helicopter Textron.)

components has made a great contribution to safety. Spontaneous failures are very rare. Generally there are symptoms such as a change in the characteristic of vibration or noise. These may be too slight to be heard by the crew, but a sensor mounted on the affected part in conjunction with a signal processor which knows what the normal sounds, or signature, from the part are, can give vital warning of a potential problem.

The life of a component may be reduced if it is subjected to higher stress. Modern electronic systems (HUMS: Health and Usage Monitoring Systems) can measure stress and the time for which it has been applied in order to calculate the safe remaining life in major components.

Much fundamental research into helicopters was done using models. De la Cierva's early work was with free flying model gyroplanes. Arthur Young used models extensively (Figure 7.1-9) and it was his demonstration of a model controlled by a trailing wire that convinced Larry Bell to enter the helicopter business. Irven Culver at Lockheed built what was probably the world's first radio-controlled model helicopter in the late 1950s.

Advances in radio control equipment in the 1960s made the necessary precision available at reasonable cost and this led to the availability of flying model helicopters for the model building enthusiast. Such models have become highly sophisticated even though, like their full-sized counterparts, they remain expensive to build and operate.

Conceptually somewhere between the model and the full-size helicopter is the UAV (unmanned autonomous vehicle). These machines are designed to perform surveillance tasks, typically carrying cameras and other sensors. As electronic and sensing devices have become smaller, useful equipment can be carried aloft at much lower cost if there is no need to carry a pilot. Unlike the model, which needs actively to be controlled by the pilot on the ground at all times, the UAV carries enough navigational equipment, automatic stabilization systems and processing power to be self-contained. Also unlike the hobbyist's machines, UAVs must be built and operated to professional standards.

7.1.3 Types of Rotorcraft

The definition of a rotorcraft is quite general, embracing any flying machine that produces lift from rotors turning in a plane that is normally close to the horizontal. This definition does not concern itself with the proportion of the machine's weight carried by the rotors, or whether that proportion changes at different stages of flight. Figure 7.1-10 shows the main classes of rotorcraft that will be defined here.

The pure helicopter obtains the great majority of its lift in all modes of flight from a power driven rotor or rotors, and any lift due to airflow around the hull is incidental. The majority of today's rotary wing machines are pure helicopters. The thrust from a rotor is closely aligned with a line drawn perpendicular to the tip path, and the pure helicopter propels itself by tilting the rotor forward to obtain a component of rotor thrust which balances the drag, as shown in Figure 7.1-10(a).

The gyroplane (also known by de la Cierva's trade name of autogyro) obtains lift from an undriven rotor that must be tilted *away* from the direction of flight to make air flow up through it. The rearward rotor-thrust component, along with the drag, is balanced by the forward thrust of a conventional airscrew as shown in Figure 7.1-10(b). As the rotor needs to be pulled through the air to maintain height, the autogyro cannot hover in

Figure 7.1-9 Arthur Young flies a wire-controlled model helicopter. Young was a philosopher who argued that designing a helicopter would teach him how to think. He was right. (Bell Helicopter Textron.)

still air, although it can give the illusion of hovering by flying into wind. Simple autogyros must taxi to spin up the rotor, but later machines could spin the rotors with engine power on the ground, and use the stored energy to perform a jump take-off.

Between the pure helicopter and the gyroplane is the gyrodyne, which obtains lift from a power driven rotor. Unlike the pure helicopter, the gyrodyne maintains the rotor disc parallel to the direction of flight, as in Figure 7.1-10(c) and propels itself with a conventional airscrew. The Fairey Gyrodyne (Figure 7.1-11) replaced the tail rotor with a side-mounted airscrew to cancel torque reaction when hovering, and also to provide thrust for forward flight. More recently the Lockheed Cheyenne (Figure 7.1-12) had both anti-torque and pusher rotors at the tail. The gyrodyne offers high speed potential, with the penalties of raised complexity, weight and difficulty of control. Some gyrodynes have wings in addition to the rotor.

The compound helicopter is one that hovers like a helicopter, but which may obtain supplementary lift from fixed wings during flight and may incorporate additional means of providing forward thrust.

The convertiplane is a more extreme example of the compound helicopter in that it reconfigures its means of providing lift and propulsion in different flight regimes. The Fairey (later Westland) Rotodyne was a convertiplane. Figure 7.1-13 shows that it consisted of a twin turboprop aircraft-like structure with a pylon-mounted rotor driven by tip jets. As a helicopter, the tip jet drive provided lift, and yaw control was obtained by differentially changing the pitch of the turboprops. Forward thrust from the airscrews would bring the machine to cruising speed, where much of the lift was developed by the wing, and the tip jets were turned off such that the rotor free-wheeled and the machine became a compound gyroplane.

Reconfiguring can also be done by tilting the whole wing–engine–rotor assembly (tilt wing) as shown in Figure 7.1-14(a) or by tilting the engine–rotor units on fixed wings (tilt rotor) as in Figure 7.1-14(b). The Bell-Boeing

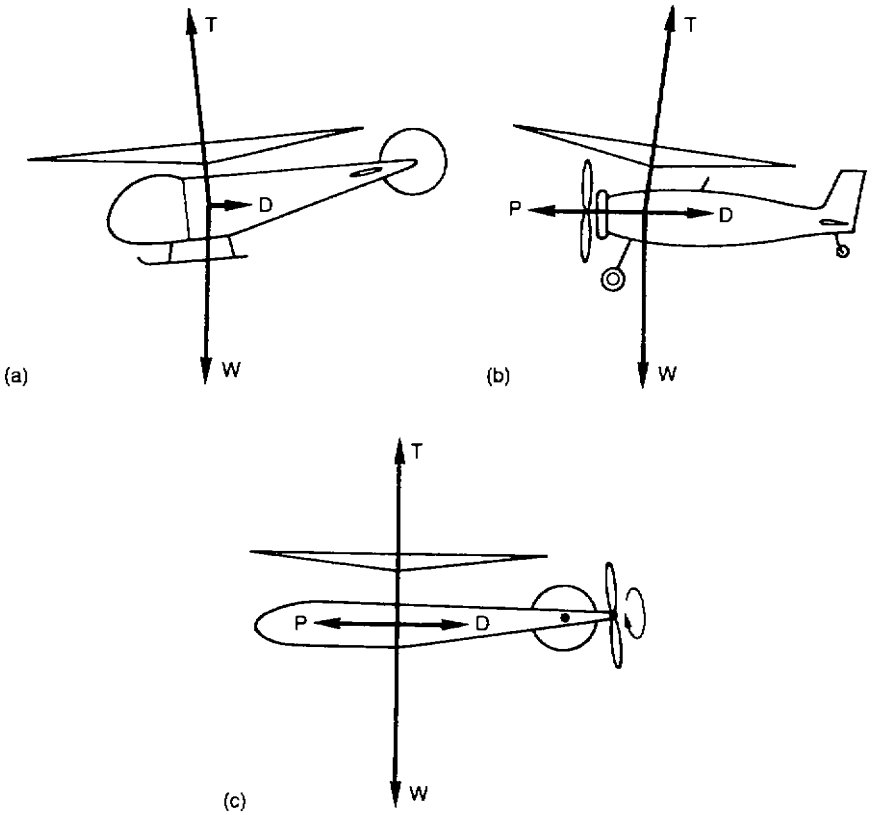


Figure 7.1-10 The main classes of rotorcraft. The conventional single main rotor machine predominates. See text for details.

Figure 7.1-11 The Fairey Gyrodyne had a side-mounted anti-torque rotor that became a tractor propeller in forward flight. (AugustaWestland.)

Osprey is a tilt rotor. As can be seen in Figure 7.1-14(c), the diameter of convertiplane rotors is usually such that the machine cannot land in the forward flight configuration, but must return to the hover.

The advantages of the convertiplane over the pure helicopter are that using the rotors as airscrews reduces vibration, it is much more efficient and allows a higher airspeed. This reduces fuel consumption and allows greater range.

Figure 7.1-12 The Lockheed Cheyenne with fixed-wing, anti-torque rotor and pusher propeller was a very fast, highly manoeuvrable machine. (Lockheed Martin.)

Figure 7.1-13 The Rotodyne hovered as a tip jet powered helicopter and cruised as a compound gyroplane. (AugustaWestland.)

The tilt rotor has its wing in the downwash which reduces hover performance, whereas in the tilt wing the wing is almost always working. However, the tilt wing needs a supplementary mechanism to control the pitch axis, such as a jet or rotor at the tail.

7.1.4 Rotor Configurations

The various configurations of the pure helicopter will now be considered. The most common configuration is the single main rotor and the anti-torque tail rotor. The remaining configurations, shown in Figure 7.1-15, handle torque

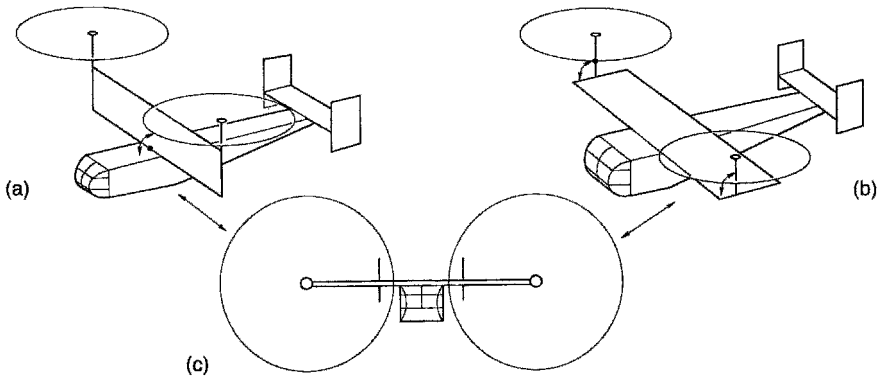


Figure 7.1-14 Types of convertiplane. At (a) the tilt wing machine moves the rotors and wing with respect to the hull. At (b) the tilt rotor machine has a fixed wing and the rotors alone tilt. (c) The machine cannot land with the rotors tilted forward.

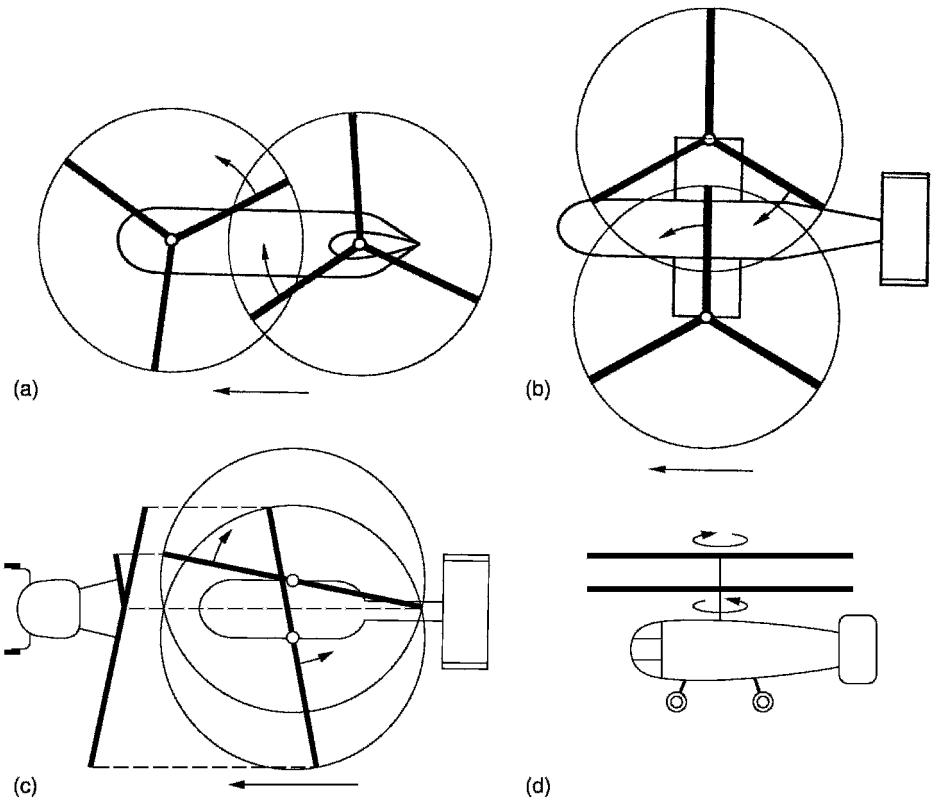


Figure 7.1-15 Types of contra-rotating helicopter. (a) The tandem. (b) The side-by-side. (c) The synchropter. (d) The coaxial. See text for details.

reaction by contra-rotation. Contra-rotating helicopters have no need for the tail rotor, but generally have a tail fin or fins for directional stability at speed.

In the tandem rotor helicopter, Figure 7.1-15(a), two rotors turn in opposite directions at the opposite ends of a long hull. The rotors are usually synchronized through a transmission system so that the shafts can be little more than a blade length apart. The Chinook in Figure 7.1-16 is a good example of the type, the large disc area offering good lifting ability, and the long cabin ample load space.

An alternative to the tandem rotor is the side-by-side twin rotor (Figure 7.1-15(b)). This has aerodynamic advantages in forward flight because the lifting area has a better aspect ratio, but it is difficult to avoid drag due to the structure needed to locate the rotors each side of the hull and to carry the transmission across the machine.

Figure 7.1-16 The Chinook is the definitive tandem rotor helicopter and has been produced in large numbers. (Boeing.)

See Figure 7.1-5. In large machines it is also difficult to make the structure stiff enough to avoid resonance and so the type is rare.

A relative of the side-by-side helicopter which has had more success is the synchropter (Figure 7.1-15(c)) in which the two rotors mesh so closely that the contra-rotating shafts can be driven by the same gearbox. The close meshing is achieved by tilting the shafts outwards so that the blades of one rotor can pass over the rotor head of the other. The German Kolibri of World War II, designed by Anton Flettner (Figure 7.1-4), was the first successful machine to use the idea. In the USA, Charles Kaman adopted the synchropter principle, and produced the famous H-43 Huskie which became the definitive crash rescue helicopter of its time (Figure 7.1-17). The synchropter is the easiest of all helicopters to fly, as the interactions and second-order effects of the conventional configuration are eliminated, but replaced by some interesting yaw characteristics.

Figure 7.1-17 The Kaman Huskie is the most successful synchropter design. Note the large fin area needed. (Kaman Aerospace.)

Figure 7.1-18 The contra-rotating coaxial principle is used extensively by Kamov. (Kamov.)

The final approach to contra-rotation is the coaxial helicopter. Stanley Hiller and Arthur Young both built such machines experimentally, but Nikolai Kamov in the USSR put the idea into production (Figure 7.1-18). The coaxial helicopter places both rotors one above the other on a common shaft, and drives them in opposite directions. Like the synchropter, control interactions are reduced, but yaw control remains an issue. The main advantage of synchropters and coaxial helicopters is that in the absence of a tail rotor, the machine can be much more compact, a crucial factor in naval aviation, where everything has to be squeezed into limited hangar space, although the height needed may increase. The alternative is to fit a folding tail on a conventional machine.

7.1.5 The Essential Elements

The helicopter contains a large number of systems and components, but these can generally be broken down into a smaller number of major areas. Figure 7.1-19 shows a cutaway drawing of a conventional tail rotor type helicopter. The main systems to consider in a helicopter are the hull or airframe, the engine and transmission, the fuel system, the landing gear, the rotors, the controls, electrical and hydraulic power, instrumentation and avionics. These subjects will briefly be introduced in the remaining sections of this chapter, and references will be made to more detailed treatments elsewhere in the book.

7.1.6 The Airframe

The fuselage or hull has a number of jobs to perform. One obvious task is to hold all of the components in the correct position and to transfer forces from the rotors, the tail surfaces, the landing gear and any internal or underslung payload. It also protects the occupants and the mechanisms from the elements whilst still allowing good visibility for the pilot. Space has to be found for fuel tanks close to the centre of mass (CM) as single rotor machines are sensitive to trim shifts as fuel burns off. It will also have a more or less aerodynamic shape, although the other requirements often combine to prevent this. In some cases the hull will also be designed to float in case the machine is forced down over water, whereas in others, amphibious operation is planned.

Fuselage construction varies considerably, but the materials and techniques are not much different from those used in any aircraft. Early machines such as the Bell 47 were no more than a steel tube lattice frame with a blown acrylic

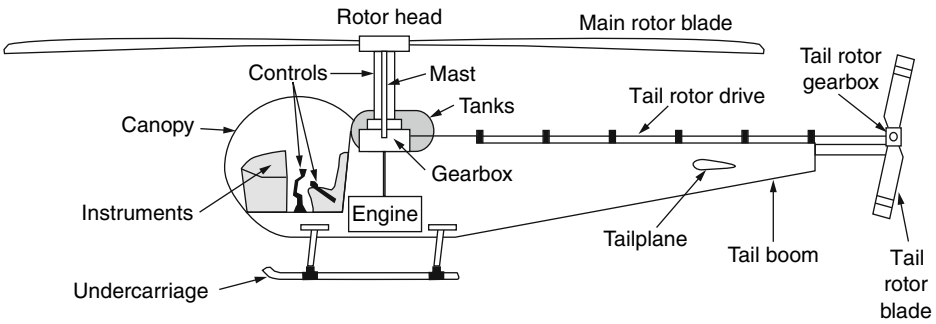


Figure 7.1-19 The major components of a conventional helicopter. Note the location of the fuel tanks close to the mast to minimize CM (centre of mass) shifts.

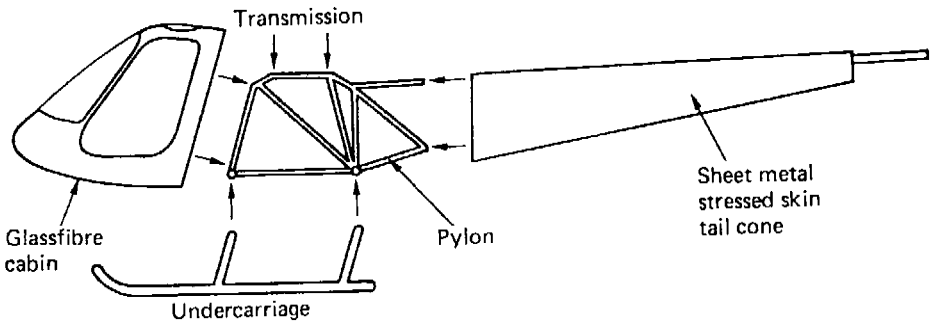


Figure 7.1-20 The structure of a light helicopter: Enstrom F-28. See text for details.

canopy for the crew. Aerodynamic improvements came when the hull was faired in. The tail cone is often a stressed skin structure, but the centre section has too many doors, windows and access hatches for the skin to carry all the loads, and alloy frames or steel tubes are often used beneath the skin. The main landing gear often shares the same framework as the engine and transmission so that the skin is not unduly stressed on landing.

Figure 7.1-20 shows the structure of an Enstrom F-28. The landing skids, engine, fuel tanks and transmission are all attached to a welded tube frame known as a pylon. A stressed skin tail cone is attached to the rear of the pylon structure, and the aluminium seat and cabin floor is attached to the front. The cabin is glass fibre with plastics glazing. The cabin lines are faired into the tail cone by unstressed panelling attached to the pylon, so the machine has a sleek outline. Much of the unstressed centre panelling can be removed for servicing.

Composite materials are ideal for airframe construction and are becoming increasingly important since they are less dense than metals and are inherently well damped, which helps to control the inevitable vibration that characterizes helicopters. They can also have an indefinite fatigue life.

7.1.7 Engine and Transmission

The engine or engines and transmission are generally close together. Piston engines are heavy, and they are almost always placed below the rotor head to balance the machine. Turbines are much smaller and lighter and are often built into the roof of the hull to maximize internal space. Both piston engines and turbines turn much faster than the rotor shaft, and the transmission must incorporate reduction gearing.

The gearbox will generate a good deal of heat on a large machine, and require an oil cooler. This will have its own air intake near the rotor shaft; often in the front of the pylon. The gearbox also drives the tail rotor shaft, which runs the length of the tail boom to the tail rotor gearbox. The tail rotor shaft is often mounted outside the tail cone for ease of inspection and maintenance.

The rotors may take some time to come to rest after the engines have been shut down, and this may be inconvenient. The civil user wants to disembark passengers with the rotors stopped. The naval user wants to put the machine on the elevator down to the hangar and the army user wants the rotors stopped quickly on a clandestine mission so that the machine can be camouflaged. This can be achieved by fitting a rotor brake on the transmission.

Helicopters have also been built with jets at the blade tips, and this has the advantages that no gearbox is required and there is no torque reaction. Unfortunately tip jets have short fatigue life, very high fuel consumption and noise level to match and are little used. They will not be considered further here.

7.1.8 The Fuel System

The fuel system can be simple or complex depending on the type of machine. In early piston engine machines using carburettors, the fuel system was little more than a pair of tanks that fed fuel by gravity through the pilot's cut-off valve and a filter to the engine. The tanks in a gravity fed system are mounted one each side of the mast so that lateral and longitudinal trim is unaffected as fuel is consumed. The tanks are clearly visible on the Bell 47 and the Hughes 300. Larger machines will use the space below the floor to put tanks close to the CM, in which case pumps will be required to feed the engines.

Piston engines burn AVGAS or aviation gasoline, which is basically similar to automobile fuel but made to tighter quality standards. Turbines burn AVTUR that has a similar relationship to kerosene. As a piston engine will stop if AVTUR reaches it, pilots like to know they have taken on the right fuel. AVGAS fillers are marked red and AVTUR fillers are marked black to help prevent a dangerous mistake.

Recently there have been significant advances in the development of Diesel engines, allowing a similar power to weight ratio to a gasoline engine to be obtained. The advantage of the aeroDiesel is that it can burn AVTUR and its improved fuel economy allows better payload or range.

7.1.9 The Landing Gear

The landing gear is subject to considerable variation. Utility and training helicopters are invariably fitted with skids to allow a landing on unprepared ground even with forward speed. Small wheels, known as ground handling wheels, can be fitted so the machine can be moved around. Some skids are broader than usual so they can act as skis for landing on snow or soft ground. Inflatable or rigid floats can be attached to the skids to permit operation over water, but there will be a drag and payload penalty. Larger machines invariably have wheels, as skids would make them too difficult to move. Naval helicopters need wheels to allow them to be moved below decks. In many cases the wheels can be locked so as to be tangential to a circle. The machine can be turned into the wind, but will not roll as the ship heels.

7.1.10 Oleos and Ground Resonance

Landing gears often incorporate a telescopic section containing oil and a compressed gas acting as a spring. These are formally known as oleo-pneumatic struts, invariably abbreviated to oleos. When the length of the oleo changes, the oil is forced through a small orifice to damp the movement. The struts that hold up automobile tailgates work on the same principle, but these are sealed units whereas the type of oil and the gas pressure in an oleo may be adjusted to give the correct spring rate and damping.

One obvious purpose of the oleo is to absorb the impact of landing, but a more important role is to control ground resonance. Ideally the rotor blades rotate with perfectly even spacing when run up on the ground, but it is possible for them to be disturbed from that condition. This results in the CM of the rotor moving away from the shaft axis and the rotor tries to whirl the top of the hull in a circular orbit. Under certain conditions this motion becomes uncontrollable unless there is damping to dissipate the energy. The rotor head may also need dampers to prevent the problem of ground resonance.

7.1.11 The Rotors

The main rotor takes the place of the wings of a conventional aircraft, and it is not unrealistic to think of a helicopter as being supported by the lift from wings that rotate instead of flying in a straight line. The main difference between rotor blades and wings is that the former are in comparison very thin and flexible, and the forces acting upon them are greater and more rapidly varying. The rotors have more to do than an aeroplane wing, because they are also the control system. Chapter 7.2 explains how the rotors produce lift and introduces the control of the machine.

7.1.12 The Control System

The control system cannot be treated in isolation, but must be integrated into the design of a machine from the outset. In the pure helicopter, control of the machine is achieved entirely by changing the pitch of the main and tail

rotor blades in various ways. This will then determine the amount of engine power needed. The rotors are generally designed to turn at constant speed and the throttle setting will have to be modified whenever the rotor power demand changes so that the speed does not change.

There are two main forces acting on a helicopter, the force due to gravity, which is always downwards, and the rotor thrust vector, which is always at right angles to the tip path plane, otherwise called the rotor disc. Chapter 7.2 shows how rotors develop thrust. The pilot can control the magnitude of the rotor thrust with the collective pitch lever held in his left hand, and the direction of the rotor thrust with the cyclic stick held in his right hand. The cyclic stick works in two dimensions: if the stick is pushed in any direction, the rotor thrust tilts the same way. These two fundamental controls are illustrated in Figure 7.1-21.

The blade movements necessary to produce lift and to achieve control will be outlined in Chapter 7.2.

The pilot controls the yaw axis by altering the pitch of the tail rotor blades using foot pedals.

It should be mentioned in passing that the No. 1 pilot in a helicopter conventionally occupies the right hand seat; the opposite of aeroplane practice. In some early machines, only one collective lever was provided, between the seats. The test pilots would sit in the left seat, as per aeroplane practice. Because of the difficulty of reversing the function of left and right hands, a pilot under training would be placed in the right seat and so the right seat became the conventional location for a helicopter pilot. A further factor is that in early helicopters, the forces fed back to the cyclic stick from the rotor were such that it was not safe to let go of the stick for an instant. However, the collective lever could be released temporarily. Given that most secondary controls such as the radio and instruments are centrally disposed so that pilot and co-pilot can both see them, it made sense to put the pilot in command on the right where his free left hand would be of most use.

7.1.13 Electrical and Hydraulic System

The power systems of a typical helicopter are not that much different from those of any aircraft. Electrical power is needed for instruments, radios, autopilots, lighting, engine starting, navigation and intercom systems, as well as a host of further avionics that might be needed for special purposes.

A light helicopter may have no power assistance, but as machines get larger the control forces may cause fatigue and in very large machines they will be beyond the strength of the pilot. Power-assisted controls then become essential. Power controls are also needed if some kind of automatic stabilization or autopilot is fitted so that the low powered electronic signals can operate the controls.

When powered controls operated by electrical signals have faultless reliability, the mechanical controls from the pilot can also be replaced by electrical controls, leading to the concept of fly-by-wire. The pilot operates controls having no mechanical connection to the rotors, but which instead produce electrical signals.

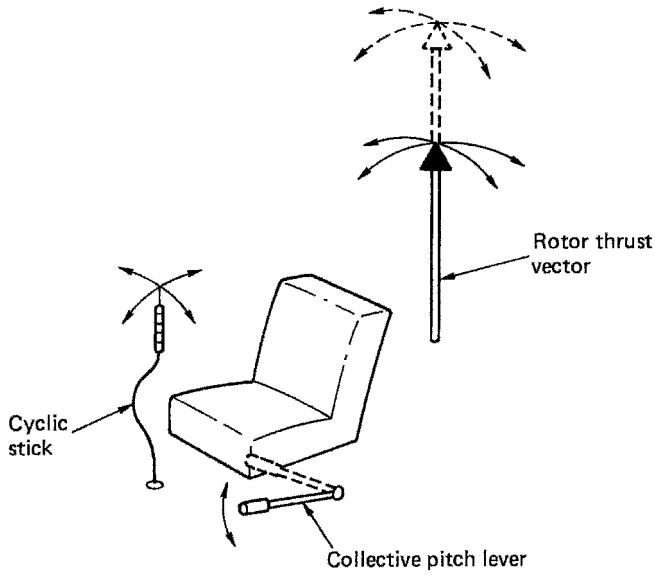


Figure 7.1-21 The fundamental rotor controls. The cyclic stick tilts the rotor in the direction it is moved, whereas the collective lever changes the magnitude of the thrust.

Electric motors are useful for low powered control purposes such as the trim mechanism, but hydraulics allows greater forces to be developed within small actuators, and so they will be used for powered flying controls.

Electrical and hydraulic power is vital to the safety of the machine, and the hydraulic pump and the generator may be driven from the rotor shaft so that power is still available even in the case of engine failure. In small machines the generator is driven from the engine, as battery capacity is enough to keep the electrical system working in the case of an engine failure. Larger machines may have two or more engines and each will have a generator so that electric power is still available in case of engine failure.

7.1.14 Instruments and Avionics

Many of the instruments fitted to helicopters are the same as those used in other aircraft, but in addition to the usual engine-related gauges the helicopter will also need a rotor tachometer. The response of a helicopter to control inputs depends on rotor speed, and this needs to be controlled to close limits.

References

1. Spenser, J.P., *Whirlybirds*, University of Washington Press, Seattle, ISBN 0-295-97699-3 (1998)
2. Spenser, J.P., *Vertical Challenge*, University of Washington Press, Seattle, ISBN 0-295-97203-3 (1998)
3. Coates, S., *Helicopters of the Third Reich*, Classic Publications, Hersham, ISBN 1-903223-24-5 (2002)
4. Boulet, J., *L' Histoire de l'Hélicoptère*, Editions France-Empire, Paris, ISBN 2-7048-0040-5 (1982)
5. DuMoulin, A., *Les Hélicoptères Florine 1920–1950*, Fonds National Alfred Renard, Brussels (1999)
6. Everett-Heath, J., *British Military Helicopters*, Arms and Armour Press, London, ISBN 0-85368-805-2 (1986)
7. Nowarra, H.J., *German Helicopters*, Schiffer, West Chester, ISBN 0-88740-289-5 (1990)
8. Dowling, J., *RAF Helicopters: the first twenty years*, HMSO, London, ISBN 0-11-772725-3 (1992)

Aerospace Engineering Desk Reference; ISBN 9781856175753

7.2 Helicopter Dynamics

John Watkinson

7.2.1 Creating and Controlling Lift

The basic principle of wings, rotors and airscrews is that they accelerate a mass of air and that the resultant lift or thrust is the Newtonian reaction to that acceleration. Producing sufficient lift to permit a helicopter to fly is only a matter of having enough power to accelerate the air mass without excessive weight. A practical helicopter must, however, be able to control that lift precisely or it would be dangerous.

Any structure placed in a flow of air for the purpose of generating lift is called an aerofoil or airfoil. Aeroplanes and helicopters can fly when there is a wind blowing and can climb or lose height in the process. In helicopters the rotor blades have airspeed due to their rotation. Thus there can be situations where the airfoil moves through the air, or where the air moves past the airfoil. All that matters from the point of view of generating lift is the relative velocity and the direction from which it appears to be approaching the airfoil. Figure 7.2-1(a) shows that this is known as a relative airflow (RAF). RAF is a vector quantity as it has speed and direction. Figure 7.2-1 further shows that a flat plate will produce lift if it is slightly inclined to the RAF at an *angle of attack*. When the airfoil changes the direction of air flowing by, this represents a change of velocity and so is classed as acceleration. The reaction to that acceleration will point in the opposite direction.

Figure 7.2-1(b) shows how the direction of the reaction is found. For the time being, the air is assumed to be inviscid (having no viscosity). The relative airspeed does not change, only the direction, so the velocity vector V is changed to V'. The acceleration to change V to V' must have been in the direction of the vector V_a so the reaction must be in the opposite direction. It can be seen that this is at right angles to the average airflow direction. The blade reaction is the only actual force present. In fixed-wing aircraft studies, the reaction is traditionally resolved into two components: that which is measured at right angles to the RAF and called the *lift*, and that which is measured along the direction of the RAF and called the *induced drag*. In cruise lift will be vertical and drag will be horizontal.

Figure 7.2-1(c) shows that the lift can be increased by increasing the angle of attack since this has the effect of increasing the change of velocity. In the aeroplane, the angle of attack is controlled by alteration of the attitude of the entire machine with the elevators. In a helicopter the blades are mounted on bearings in the rotor head that allows them to turn about a radial axis as can be seen in Figure 7.2-1(d). This mechanism is known as *feathering*.

When the relative velocity alone is doubled with respect to the case in (b), the vector V_a representing the change of velocity is doubled. As the air is moving twice as fast, the wing accelerates the air in half the previous time. Achieving double the velocity in half the time means that the reaction is increased by a factor of four. In other words, the reaction is proportional to the *square* of the velocity. In practice four times as much lift would be quite unnecessary to balance the weight of the machine and instead, as the relative velocity doubled, the angle of attack would be reduced to about one-quarter of the previous value. This is the mechanism by which aircraft obtain the same lift over a range of airspeeds.

In practice the air also has viscosity that resists movement as a function of the surface area of any body moved through it. The loss due to this effect is called *profile* drag. The profile drag tilts the blade reaction further to the rear as in Figure 7.2-2(a). There is a boundary layer between a moving object and the stationary air. At the leading edge, the speed difference across the boundary layer is maximal and the drag is greatest. Further back, the viscosity has been able to get the boundary layer moving and the drag is reduced. Thus the drag experienced by a wing or rotor blade is a function not just of the surface area but also of the chord. Very small objects possess disproportionately high profile drag, which is why the wind can blow sand but not rocks. Reynolds numbers are used to describe the relationship between dimensions and profile drag. It is important in model testing that the airspeed used is scaled correctly. Scale speed is not used. Instead the speed is chosen to produce the same Reynolds number in the model as in the full-sized device and then the drag of the model will be representative.

If an airfoil is not producing lift, the induced drag will be zero and only the profile drag will be observed. In a full-size helicopter in the hover roughly 30% of the rotor shaft power is lost to profile drag, the rest goes in accelerating the air downwards. In model helicopters and full-sized tail rotors profile power is proportionately greater as the small-chord blades operate with adverse Reynolds numbers.

Figure 7.2-2(b) shows the production of the same lift at a higher relative velocity. The profile drag has increased but the induced drag has reduced. This is not what one would expect from a true source of drag. The explanation is that the airfoil is an actuator or transducer, which accepts mechanical power as the product of the mechanical

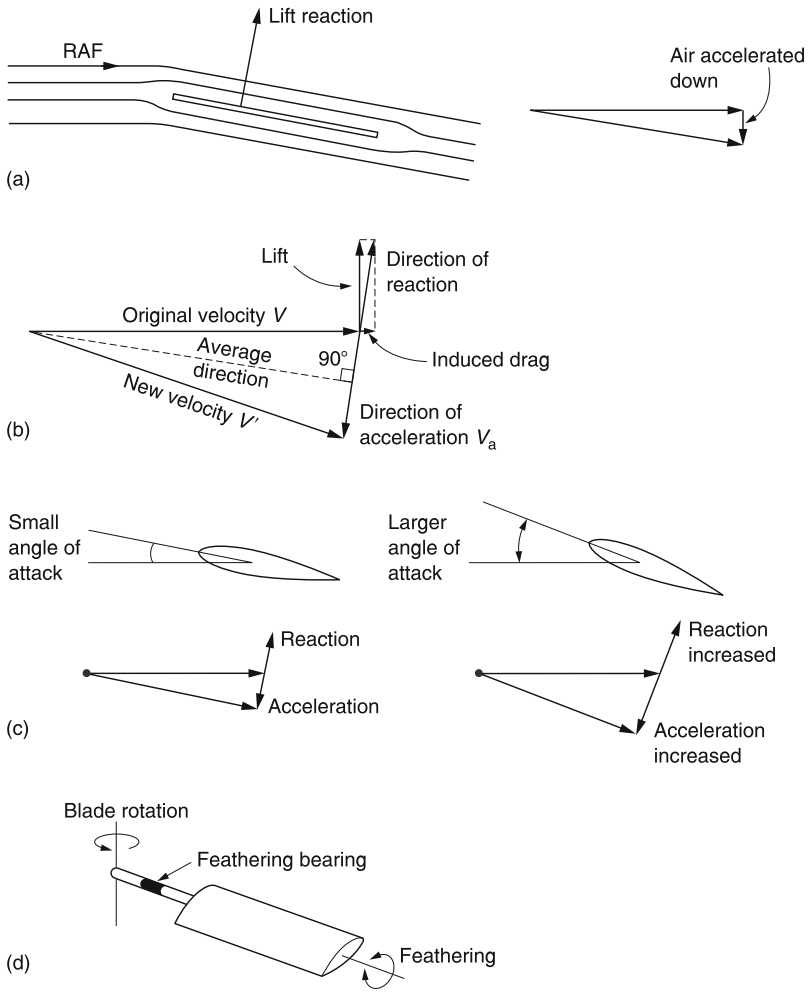


Figure 7.2-1 (a) Lift is the reaction to the acceleration of air into a new direction. (b) The direction of acceleration of the air is derived as shown. The reaction must be in the opposite direction. (c) An increase in the angle of attack increases the lift. (d) In the helicopter, lift is controlled through twisting the blade about a span-wise axis; a mechanism called feathering.

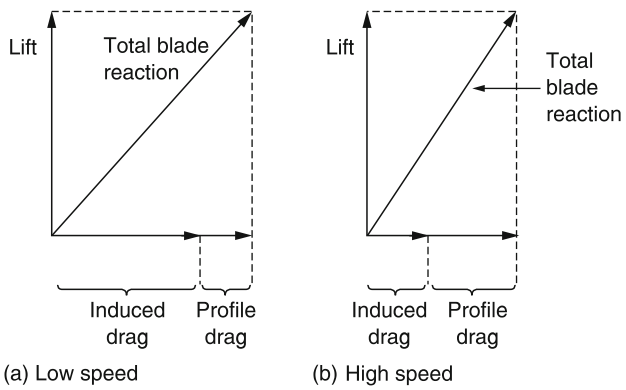


Figure 7.2-2 Profile drag must be included in the behaviour of any real airfoil. The induced drag and profile drag add to tilt the total blade reaction backwards. Conditions at low airspeed are shown in (a). At a higher speed (b), the profile drag is greater, but the induced drag is less. This is only possible because it is not really a drag.

impedance and the relative velocity in order to drive the air against aerodynamic impedance. Accelerating air downwards constantly provides lift. The actuator reflects the aerodynamic impedance back as mechanical impedance. The variation in angle of attack is analogous to changing the turns ratio of a transformer or the ratio of a gearbox. Clearly if the relative velocity goes up, to keep the power the same the impedance must fall.

Once the concept of an airfoil as an actuator is understood, it is a small step to appreciate that most types of actuator are reversible. Some helicopters use an electric motor to start the engine when that motor becomes a generator. In a turbine engine, the rotating blades in the compressor are putting energy into the air, whereas the rotating blades in the power turbine are taking energy out of the gas flow. Given this reversibility, we can immediately see how in a glider or in an autorotating helicopter the power flow through the actuator reverses. In descending through the air the wing or rotor delivers power extracted from loss of potential energy and this power is used to overcome profile drag. The glider can maintain airspeed and the helicopter can maintain rotor speed with no engine power.

The act of turning the airflow also results in rotational energy being imparted to the passing air. Conceptually, the airflow can be divided into two flows. One is the steady flow past the blade and the other is the rotational component that in the absence of the steady flow would rotate around the blade. This is known as *circulation* and it is proportional to the lift. In vortex theory, the circulation of an airfoil is calculated and the lift follows from that.

A further consequence of the change of direction is that the horizontal component of the relative velocity is reduced. It follows from the rearward inclination of the blade reaction that there must be a forward component of the acceleration imparted to the air. Air does not go straight down from a hovering rotor; it also revolves in the same direction the rotor turns, but much more slowly: a phenomenon known as swirl. Swirl is considered in section 7.2.14.

It is more efficient to accelerate the air down gradually rather than have it suddenly find an inclined plate. Air passing the top edge of a flat plate cannot change direction quickly enough and the flow separates and becomes turbulent. These are the reasons for curving or cambering an airfoil. Furthermore, practical wings must contain structural members such as spars, and the thickness will be increased to accommodate them. The thickness is contained within a streamlined shape to reduce drag.

Figure 7.2-3(a) shows a streamlined cambered section. The mean camber line is half way between the upper and lower surfaces. The chord line joins the ends of the camber line. The angle of attack is the angle between the chord line and the relative airflow. The camber of the airfoil can be optimized for the speed range of interest. For high lift at low speeds, the camber will be heavy in order to make V_a as large as possible. Clearly such a wing will be creating a lot of drag at the small angles of attack needed at high speeds. For high speed operation, the camber will need to be very small, but this kind of wing will be inefficient if used at large angles of attack at low speed, hence the use of flaps on fixed-wing aircraft. With a cambered section, air is still accelerated downwards leaving the section even when the angle of attack is zero, and so a slight negative angle of attack is necessary to obtain the zero lift condition.

Figure 7.2-3(b) shows a symmetrical section in which the camber line and the chord line are one and the same. A symmetrical section is a streamlined flat plate. The curvature prevents separation over the leading edge. At zero angle of attack, the airflow is also symmetrical, and no net air reaction results. There is thus no induced drag, only profile drag. If the angle of attack is made positive, air is accelerated down, and the reaction is upwards. If the angle of attack is made negative, air is accelerated up and the reaction is downwards.

The reader is cautioned against explanations of the origin of lift based on Bernouilli's theorem. Bernouilli made it quite clear that his theorem relates to conservation of energy in flowing air such that the sum of the static and dynamic pressures remains constant. Bernouilli's theorem only applies if no energy is put into the air. However, a wing or a rotor blade is an actuator that is exchanging energy with the air. This is clear from the presence of induced drag. When energy is put into the airflow, Bernouilli's theorem simply doesn't apply and the explanations based on it are flawed and should be disregarded.

7.2.2 The Centre of Pressure

The lift developed by the blades is distributed over the chord, but not uniformly. The centre of pressure is where a single force would act producing the same effect as the distributed lift. In real airfoil sections, the centre of pressure is ahead of the mid-chord point. If the blade is made of a material with uniform density, the centre of pressure will be ahead of the centre of mass and, as Figure 7.2-4(a) shows, a couple results. This couple would tend to twist the blade and increase the angle of attack, making the lift greater and increasing the twist further. In extreme cases the blade will flutter; a violent condition that will usually destroy any structure suffering from it.

In aircraft, wings can usually be made rigid enough to prevent flutter, but this cannot be done with the long thin blades of the helicopter. The solution universally adopted is to construct the blade so that the mass centroid of the blade is ahead of the local centre of pressure. As Figure 7.2-4(b) shows, this results in a stable blade, since an increase in angle of attack producing more lift tends to generate a couple *reducing* the angle of attack.

The airfoil section selected for a helicopter blade will be a compromise to satisfy a number of conflicting requirements. One of these is a minimal migration of the centre of pressure over the normal range of operating conditions so that excessive feathering couples are not fed back into the controls.

In a cambered airfoil, the centre of pressure moves fore and aft with changes in angle of attack. The downward twist at high speed was enough to twist the blades against the pilot's efforts on early machines and caused some crashes. For some time, helicopters used little or no camber in the blade section, but subsequently cambered sections having reflex trailing edges were developed which reduced the centre of pressure movement. This along with the development of structures with greater torsional stiffness allowed cambered blades to return to use, although almost invariably in conjunction with powered controls.

7.2.3 The Coefficient of Lift

Lift is equal to the rate of change of momentum induced downwards into the surrounding air. For unit airfoil area, it is proportional to the air density, the square of the relative air velocity and, as was seen in section 7.2.1, it is a function of the angle of attack. When an airfoil section is tested, the reaction is resolved into horizontal and vertical directions with respect to RAF so that two factors of proportionality, called the coefficient of lift C_L and the coefficient of drag C_D are measured. Figure 7.2-5(a) shows how these coefficients change as a function of the angle of attack. For small angles C_L is proportional as it was seen in Figure 7.2-1(c) that the acceleration of the airflow is proportional to the angle through which it is deflected. As the angle of attack increases the reaction rotates back and the lift component of the reaction will be smaller than the reaction. The graph curves away from proportionality. At a large angle of attack, separation takes place, and the C_L drops sharply, accompanied by a sharp increase in C_D. When this occurs the airfoil is said to be *stalled*.

An airfoil does not stall immediately the angle of attack is raised, and if the rate of angular change is sufficiently high the lift momentarily available might be double the amount available in steady-state conditions. This is variously called the Warren effect or *dynamic overshoot* and is shown in Figure 7.2-5(b). In the case of a helicopter the angle of attack is changing at rotor speed and dynamic overshoot has a considerable effect. In practice it means that helicopter rotors may not lose significant amounts of lift through stall.

In this respect the helicopter has the advantage over the fixed-wing aircraft. The latter has to be flown at all times with the probability of a stall in mind. Even when the pilot has sufficient airspeed, a sudden gust due to wind shear or a microburst can reduce that airspeed and cause a stall. It was this concern that led Juan de la Cierva to develop the gyroplane. He correctly argued that a rotary wing aircraft that could not stall would be safer.

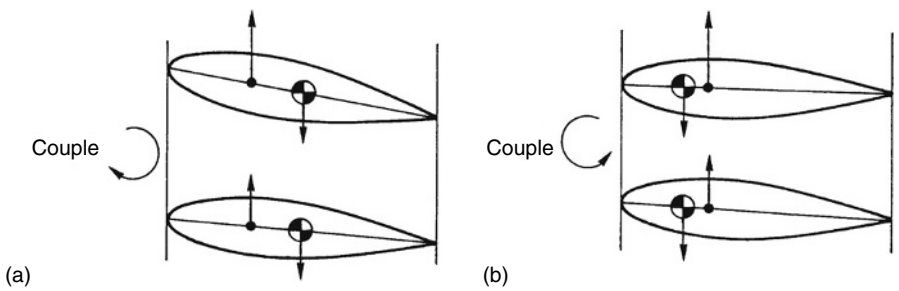


Figure 7.2-4 (a) If the centre of lift is ahead of the CM, lift tends to increase the angle of attack. (b) If the centre of lift is behind the CM, lift tends to reduce the angle of attack: a stable condition.

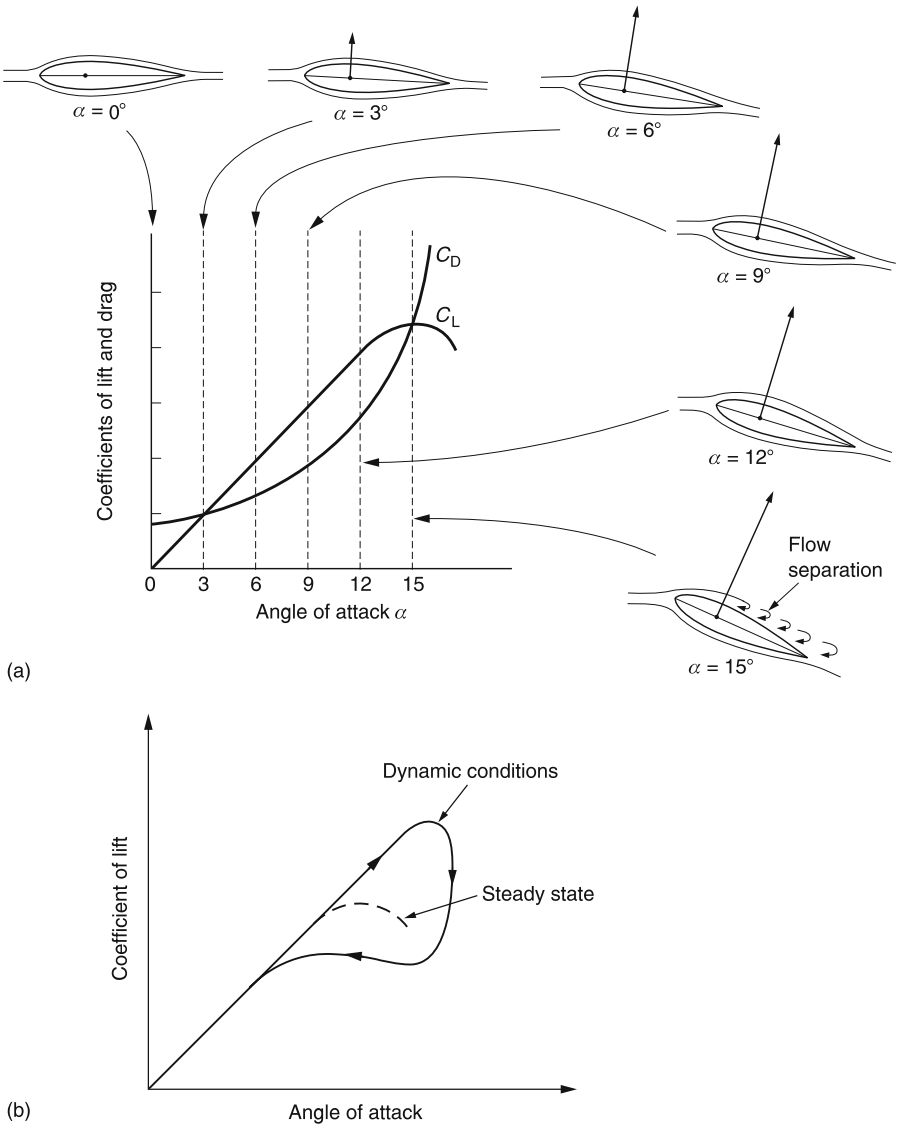


Figure 7.2-5 (a) How lift and drag change with angle of attack. Initially lift is almost proportional whereas drag changes with the trend shown. As there is profile drag at all angles of attack, the curve does not start from zero. The most efficient use of the airfoil will be where the *L/D* ratio is greatest. At steady large angles of attack the airfoil stalls. (b) Stall does not occur instantaneously and for a short time after a large angle of attack is applied, lift increases. With their rapidly feathering blades, helicopters can exploit this phenomenon.

The coefficient of lift cannot be controlled directly; in order to obtain a certain C_L the airfoil must be set to the appropriate angle of attack. The drag is also a function of angle of attack. It is more useful to know the ratio of lift to drag, since the peak value of this, $L/D_{max.}$, indicates the most efficient mode in which the airfoil can be used.

For a given engine power, the greatest lift in the hover would be obtained by operating the blades at the maximum lift over drag ratio. However, a practical helicopter cannot operate in this condition because it would be unsafe in forward flight. Many manoeuvres, such as a banked turn, require the rotor to produce a thrust which exceeds the static weight of the machine. The ratio of thrust to static weight is called the *load factor*. A rotor hovering at $L/D_{max.}$ has a load factor of unity and would stall in any real manoeuvre. It is thus a characteristic of helicopters that the rotor must operate below $L/D_{max.}$ of the blade section in the hover in order to allow a reasonable load factor for forward flight.

Turning at a given rate (number of degrees per second) requires a lateral acceleration proportional to the square of the airspeed. As aeroplanes tend to fly faster than helicopters, they will need high load factors to obtain reasonable rates of turn. Aeroplanes automatically have high load factors at speed because available lift increases as the square of

the airspeed. In helicopters lift is dominated by rotor speed, not airspeed and so load factors in helicopters tend to be small although helicopters can still easily out-turn most aeroplanes.

As the relative airflow in helicopters is dominated by the rotor speed, the gust response of helicopters is much reduced in comparison to that of aeroplanes. This is one reason why helicopters can operate in bad weather. A further small luxury is that the flexibility of the rotor gives in gusts a decoupling effect similar to that given by the suspension of a car.

7.2.4 Collective Control

In a hovering helicopter, the only source of lift is the rotor. Obtaining sufficient lift is only a matter of providing a suitable combination of power and efficiency, and is much easier than controlling lift. As the lift is proportional to the square of the speed, it is in principle possible to control lift by changing the rotor speed. Some early machines did just that. Unfortunately, the inertia of the rotor means that speed changes cannot rapidly be accomplished. For practical reasons a constant rotor speed is much to be preferred.

The proportionality between the coefficient of lift and the angle of attack is the solution. As was shown in Figure 7.2-1, the rotor blades are mounted on feathering bearings. Figure 7.1-21 showed that the pilot holds in his left hand a collective pitch lever pivoted near the back of his seat. The lever gets its name because by lifting it, the pitch angle of all of the rotor blades is increased by the same amount, and the rotor lift increases immediately. This feature of the helicopter contributes enormously to its safety. When a fixed-wing aircraft flies slowly, sudden loss of lift can only be countered by raising the airspeed and this takes time. In the helicopter the airspeed is always present due to the rotating blades and as this speed is many times higher than any normal windspeed, the airspeed seen by the blades is always adequate.

In addition to instantaneous response to changing lift conditions, the helicopter rotor stores in kinetic energy the equivalent of full engine power applied for several seconds. The pilot can transiently increase the available power by using so much collective pitch that the rotor slows down and converts its kinetic energy into lift.

7.2.5 In the Hover

Figure 7.2-6(a) shows a helicopter with flight rotor RPM (RRPM), but with the collective lever lowered. The thrust vector is small, so gravity keeps the machine firmly on the ground. If the collective lever is raised, eventually the rotor thrust vector will exceed the weight, and the machine will rise. By adjusting the collective lever, the machine can be made to hover with the thrust exactly balancing the weight as in (b). If the thrust is directly upwards, there is no resultant and the machine stays still. In the still-air hover, the inflow through the rotor is axial, and if, for simplicity, it is assumed that the centre of gravity of the machine is at the mast, the forces on the blades will not vary as they turn.

The collective control can be used to make the rotor thrust greater than or less than the weight of the machine. The force imbalance causes the machine to accelerate up or down. However, this acceleration does not continue indefinitely. Figure 7.2-7(a) shows that as the machine rises, the angle of relative airflow experienced by the blades is reduced by the vertical velocity. This tends to reduce the angle of attack and hence the rotor thrust. As a result the vertical velocity reaches a value where the rotor thrust once more balances the weight and the velocity then remains constant.

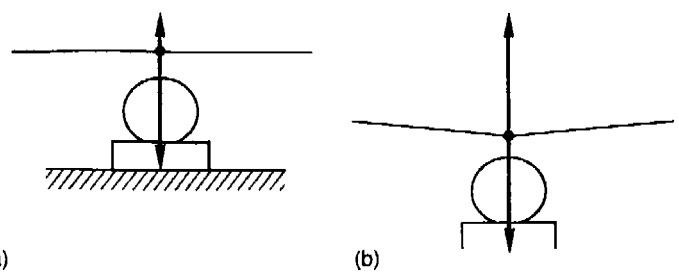


Figure 7.2-6 (a) With the collective lever lowered, the rotor thrust is less than the weight and the machine remains on the ground. (b) With an appropriate setting of the collective lever, the rotor thrust can be made exactly equal to the weight.

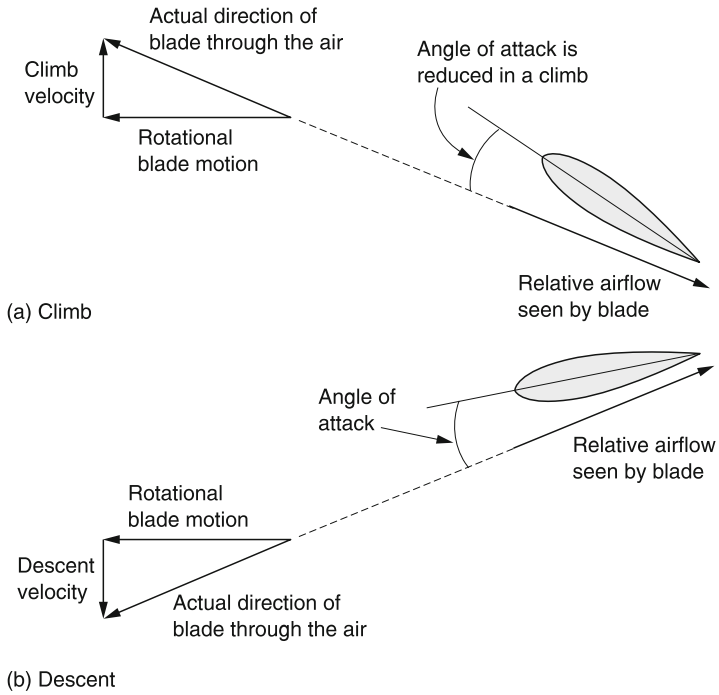


Figure 7.2-7 (a) With increased collective, the machine rises, but in doing so changes the angle at which the air approaches the blades and reduces the angle of attack. (b) With reduced collective, the opposite happens. Thus the collective lever primarily controls vertical velocity.

Figure 7.2-7(b) shows that if the collective control is lowered, the machine accelerates downwards, but the change in relative airflow is such that the rotor thrust is increased. The vertical velocity again stabilizes at a value where the thrust balances the weight. Thus the effect of the collective control in the hover is primarily to control the vertical velocity.

In most powered flight conditions, but particularly when hovering, the rotor is working in the descending airflow it has itself caused. This self-inflicted problem is called *inflow*. When inflow is present, there is a vertical component of airflow vectorially to be added to the local blade velocity in order to find the direction of the RAF.

The pilot controls the pitch angle using the collective lever and this, along with the velocity of inflow through the rotor, determines the angle of attack. As a result there is no one correct position for the collective pitch lever and the pilot will need constantly to adapt the collective pitch setting for changes in speed, fuel burned and altitude. As there is no correct position for the collective lever, there is no spring centring. Instead there may be a friction mechanism that can be adjusted to hold the lever at the last position without unduly adding to the force needed to move it. This allows the pilot to release the lever for short periods in order to operate secondary controls.

7.2.6 Forces on the Blades

In a helicopter having a hingeless rotor, the displacement between the area where lift is generated and the hub where the load is applied causes upward bending stresses that increase dramatically towards the rotor head. The underside of the blade is placed in tension, and the upper surface is placed in compression. The rotation of the blades produces a tension along their length that also increases towards the rotor head. The stress caused by this tension is of the same order as the stresses due to bending, and the stress on the underside of the blade is roughly doubled, although the compression on the upper surface is relieved. The drag of the blades causes a rearward bending stress that is constant in the hover.

There is a tendency for the blade rotation to force the blade to twist to a pitch angle of zero. Figure 7.2-8(a) shows that when the blade is set to zero pitch, the leading and trailing edges of the blade are further from the feathering axis than when pitch is applied. In order to set the blades to a working pitch, some of the mass of the blade must be brought closer to the shaft axis, and work will need to be done against the rotational forces that will be reflected as a reaction seen by the control system.

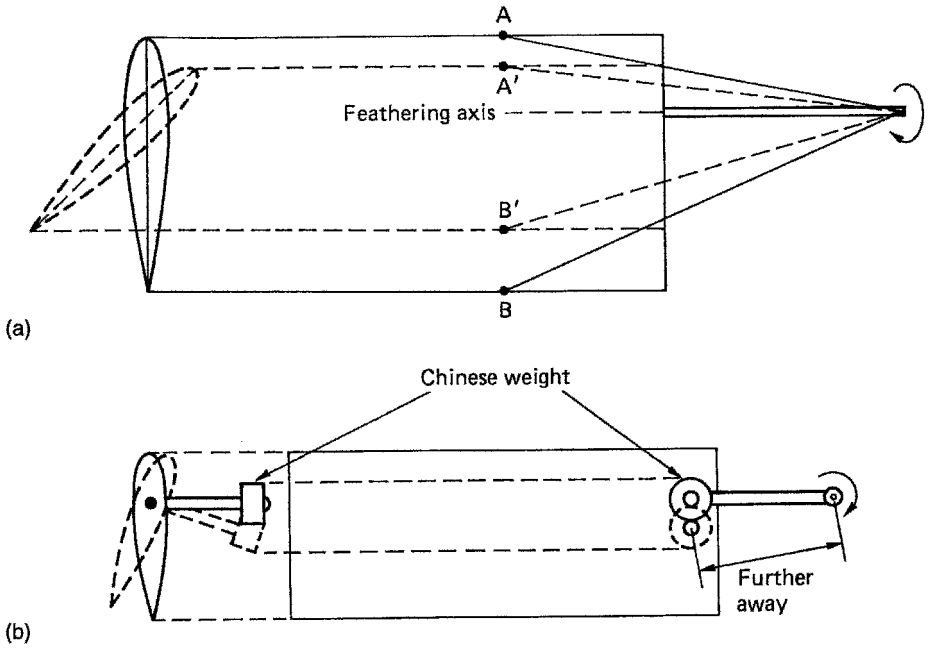


Figure 7.2-8 (a) When the blade pitch is increased, parts of the blade are brought closer to the axis of rotation. The blades will tend to be thrown to flat pitch when the rotor is turning. (b) So-called Chinese weights are mounted so that they oppose the effect of (a) by moving away from the axis of rotation as pitch is increased.

The effect can be reduced by the use of so-called Chinese weights that are placed on a rod at right angles to the chord line as shown in Figure 7.2-8(b). These will tend to move away from the rotor shaft axis as a result of rotation. The term comes from the tales told by sailors to the gullible alleging the orthogonality of certain parts of the anatomy of the oriental female at a time long before the development of either the helicopter or political correctness.

A less colourful alternative is to incorporate a spring into the collective linkage that is adjusted to supply a steady force to hold the blades at a typical positive pitch, relieving the pilot of the need to apply a constant force to the collective lever. When the blades are not turning this spring may be powerful enough to force the blades to maximum pitch and so the collective lever is fitted with a lock that the pilot can apply prior to shutting down the rotor. Another possibility is for the rotor head to incorporate centrifugal weights that only apply an upload to the collective when the rotor is turning.

7.2.7 Rotor Coning

Anyone who has examined a parked helicopter will immediately notice how droopy the blades are. A considerable deflection can be obtained by lifting the tip with one finger. How, then, can the blades lift the weight of the helicopter? The answer is that the rapid rotation in conjunction with the mass of the blades causes them to be pulled out straight; a phenomenon called *centrifugal stiffening*. Rotating blades are not in static equilibrium and they must be made to accelerate towards the shaft if they are to follow a circular path, and this requires a considerable inward or centripetal force, which will always be an order of magnitude more than the weight of the helicopter.

Figure 7.2-9(a) shows that if the blade bends upwards, the downward component of the centripetal force will balance the lift at some *coning angle*, and the resultant will be a horizontal force only so no further bending takes place. The reaction of the blades at the rotor head is shown in Figure 7.2-9(b). The force from the coning blade has an upward component, which is the lift, and an outward component due to the rotation.

In the hover, if the blades are properly balanced and all have the same coning angle, the horizontal forces cancel in the rotor head and only lift results. Adjusting all of the blades to the same angle is achieved using the process of tracking which ensures that the collective control applies exactly the same pitch to each blade. Clearly, if the blades are not balanced or not tracking, vibration will result. Accurate blade balancing and tracking is important.

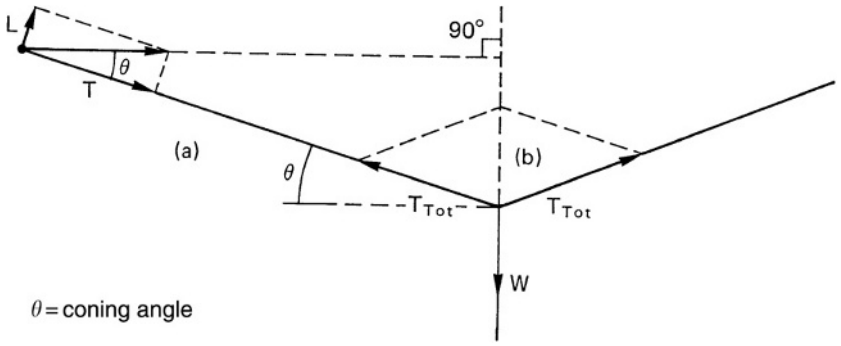


Figure 7.2-9 The blades must be accelerated inwards to make them rotate. At (a) the blades cone upwards until the resultant of the lift and the blade tension is perfectly horizontal. At (b) the force balance at the rotor head is shown. The tension in the upwardly coned blades cancels in the horizontal direction leaving only a vertical component to balance the weight of the machine.

7.2.8 Torque and Thrust in Rotors

Whilst resolving the airfoil reaction into lift and drag is useful for fixed wings, it is less useful for helicopters because of the feathering action of the blades. Figure 7.2-10 shows how the reaction on a rotor blade can be resolved in a more useful fashion. As the only real force on the airfoil is the reaction, resolving the force into components is only taking place in our imagination, so we can resolve into whatever directions we find useful. In helicopters, whatever the mechanical pitch angle of the blade, and whatever angle of attack results, it is more convenient if the reaction is resolved into the rotor thrust acting substantially at right angles to the tip path plane or rotor disc, and the rotor drag acting in the tip path plane.

The angle of attack is the angle by which the pitch angle exceeds the angle of the RAF. The reaction of the airfoil is tilted further back in the presence of inflow, so that a larger component of the reaction is opposing the thrust delivered by the engine, and power is consumed simply driving against it. The rotor thrust is also slightly reduced by the tilt of the reaction. The hover out of ground effect (HOGE) is a particularly power consuming exercise. Inflow tilts back the blade reaction to oppose the engine, so that more torque is required to drive the rotor and more tail rotor power will be needed to balance the torque. HOGE is also one of the worst situations in which to lose power as the high rotor drag will reduce RRPM more before the pilot reacts and lowers the collective lever.

Clearly increasing the collective pitch will increase the induced drag, and so more power will be required to maintain rotor speed. In simple piston engine helicopters, the end of the collective lever is fitted with a twist-grip connected to the throttle. Maintaining RRPM is then the responsibility of the pilot who will need to monitor the rev. counter. In more sophisticated piston engine machines and in all turbine machines automatic rotor speed governing systems or throttle correlators are used. The pilot just changes the collective pitch and the system computes the engine power for itself.

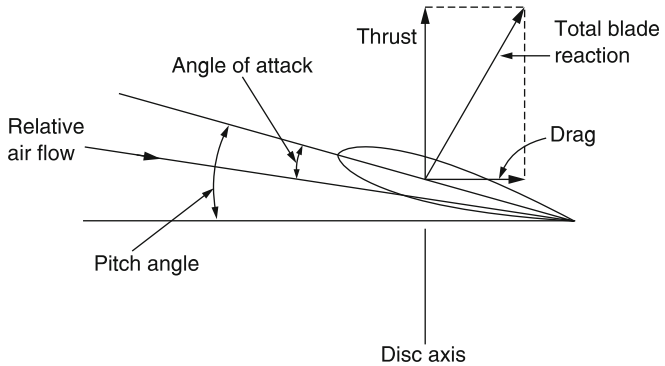


Figure 7.2-10 The blade reaction is usefully resolved into components both in and orthogonal to the tip path plane. These allow the rotor thrust and torque to be derived. In the presence of inflow, the angle of attack is the amount by which the pitch angle exceeds the angle of the relative airflow.

7.2.9 The Rotor as an Actuator

Without considering constructional details, it is possible to conceive of an ideal hovering rotor as an actuator disc that somehow accelerates air downwards over a circular region and develops thrust from the reaction. In this case the operation of an ideal rotor can be analysed theoretically from the change of momentum of the air passing through. The concept of an ideal rotor is useful because it can act as a benchmark with which to compare real designs. Clearly the performance of the ideal rotor can only be approached, but never exceeded.

Figure 7.2-11(a) shows the actuator concept which makes some simplifying assumptions. One of these is that the air is somehow constrained to flow only within the tapering column shown, known as a stream tube, and doesn't mix with the surrounding air. The mass of air passing any horizontal plane in the column per unit time must be constant. The actuator causes a pressure difference to exist across itself. This pressure must be given by the thrust divided by the disc area. A further assumption is that this pressure is uniform. The pressure difference across the actuator can only be sustained locally. Figure 7.2-11(b) shows that at a distance, the pressure must be the same as static pressure. To allow the pressure difference across the actuator, the pressure must have a falling gradient along the streamlines except at the actuator itself.

Bernouilli's theorem can be used to predict what happens along a stream tube. Bernouilli's theorem is simply another example of conservation of energy. Air has a static pressure and a dynamic pressure due to its motion. Bernouilli's theorem states that the sum of the two, known as the head, remains constant. As a result when static pressure falls, the dynamic pressure must increase to compensate. As the density doesn't change significantly, the velocity must increase. Clearly Bernouilli's theorem does not hold across the actuator, as energy is put into the air there.

Except for the pressure step at the actuator, the air in the stream tube experiences a falling pressure gradient which causes the stream to accelerate and contract. This is the phenomenon of wake contraction seen in propellers and rotors. Thus the velocity with which air arrives at the actuator is higher than it is some distance before the actuator. This difference is the induced velocity. The induced velocity is important because it determines the velocity the rotor sees which has a large bearing on the RAF and the power needed. It should not be confused with the final velocity which is higher still.

In the case of an actuator that is climbing vertically, the rotor thrust is given by the rate of change of momentum of the air passing through the disc. The power needed must be the product of the thrust and the velocity, where the velocity is the rate of climb plus the induced velocity. This power must be equal to the difference in the kinetic energy well above and well below the disc. Clearly this assumption implies that only the induced drag is being considered. Actuator theory cannot account for profile drag and assumes it to be zero.

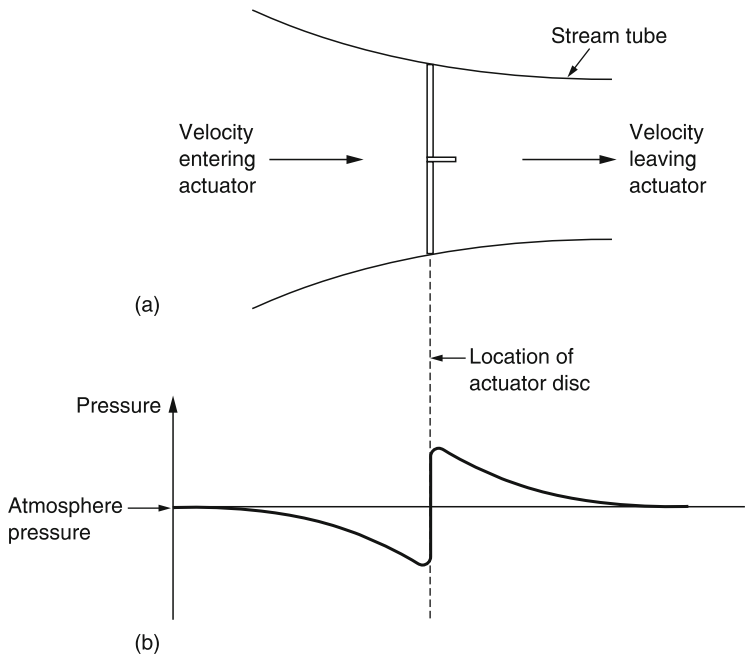


Figure 7.2-11 (a) In actuator theory, air passes along a stream tube which is intersected by the actuator. (b) In order to allow a step pressure difference across the rotor, pressure must fall both approaching and leaving. Note the wake contraction to about 0.7 of the rotor diameter.

The increased velocity applied to the downwash is twice the induced velocity at the disc. In a stationary hover, the total velocity is the induced velocity alone and the downstream velocity will be twice this. In practice the contraction has completed by about one diameter below the rotor. At this point, as the velocity has doubled, the cross-sectional area must have halved in order to have constant mass flow and so the downwash diameter will be about 0.7 of the rotor diameter.

In Figure 7.2-11, uniform inflow over the disc is assumed. In other words the induced velocity and the pressure are the same all over. This is the most efficient condition. It is easy to show that any other condition requires more power. For a given total thrust, if one area of the disc has greater induced velocity and thrust, another area must have less induced velocity and thrust. The thrust is proportional to the momentum change, which is in turn proportional to the induced velocity, whereas the kinetic energy imparted is proportional to the square of the induced velocity. Thus the power saved in an area of reduced thrust does not compensate for the extra power needed in an area of increased thrust because of the square law. It follows that the minimum power is required when the thrust per unit area is constant.

Helicopter flight is dependent on a pressure difference across the rotor disc that disturbs the air. Such pressure disturbances travel at the speed of sound and can thus easily travel long distances from the helicopter. It should be clear that a helicopter in flight is associated with the induced movement of a significant air mass. Such a moving mass cannot change its velocity in an instant. Consequently a rapid change of blade pitch does not result in an immediate change of inflow velocity. This will result in an angle of attack initially higher than in the steady state and this may result in a thrust overshoot. The same effect is observed in sailing vessels where there is a lag between sheeting the sails and a change in sail reaction. The phenomenon is known as *dynamic inflow* and is a macroscopic version of the dynamic overshoot phenomenon of Figure 7.2-5. The inertia of the associated air mass will also have effects in translational flight as will be seen in section 7.2.21.

7.2.10 Blade Element Theory

The actuator concept is useful, but it does not consider a number of real world factors, not least profile drag. Blade element theory takes this into account and allows a more accurate result to be obtained. In the hover the forward speed of some point on a blade is proportional to the radius of that point. It is possible to analyse the performance of a rotor by considering it to be made of small elements where the conditions over each element are substantially constant. The overall result is obtained by adding up the contribution from each element.

At each blade element, the inflow must be known so that the RAF and the effective airspeed can be calculated. The characteristics of the blade section employed at that element must be consulted to find the resultant aerodynamic force on the element that will be resolved into an element of rotor thrust and an element of drag. All of the axial thrusts from the elements can simply be added to find the total thrust. The drag of each element is multiplied by the radius of the element concerned to obtain an element of rotor torque. These elements are then added to obtain the total rotor torque.

Blade element analysis is only as accurate as the assumptions made about the inflow. Actuator theory gives some idea about the inflow, but in a real rotor the inflow at the tips and the blade roots will be non-ideal. More accurate results require the use of vortex theory to take into consideration the conditions at the blade root and tip.

7.2.11 Disc Loading

The thrust production mechanism in the helicopter creates a pressure difference, which causes the air to accelerate downwards. To maintain height, the rotor then has to climb up through the air at the same speed as the air is coming down. Work is being done on the air and the power is the product of the weight of the helicopter and the induced velocity. It follows that the power needed to hover can be minimized by reducing the induced velocity. This can be achieved by reducing the disc loading. This is the helicopter's equivalent of span loading in a fixed-wing aircraft. Where power is limited, as in piston engine helicopters, a low disc loading will be needed to extract more lift from the available power. Some of that lift will be lost because the larger rotor will be heavier.

In a turbine helicopter, the engine and fuel supply weigh less for the power the engine produces and so a better result may be obtained by using a smaller rotor and transmission that also weighs less to obtain a greater useful payload. A comparison of, for example, an Enstrom F-28 and a Hughes 500 illustrates this point.

This cannot be taken to extremes, as there is a practical limit to the downwash velocity. This is reached when objects on the ground are blown about and become a hazard. The CH-53E Stallion represents about the limit of acceptable downwash velocity as it is almost impossible to stand in the downwash. A further consideration is that in autorotation a machine with a high disc loading may have a rapid rate of descent, leaving the pilot little margin for

error in judging the flare-out in the case of an engine failure. On the other hand a machine with a low disc loading may have very good autorotation performance, the Enstrom being a particularly good glider.

Once a disc area and loading has been decided, some consideration has to be given to the tip speed. Profile power is proportional to the cube of the tip speed, whereas thrust is proportional to the square of the tip speed. The minimum profile power is where the minimum possible tip speed is used to operate the blades at L/D_{max}, which will be just below the stall. Reducing the tip speed will require the blade area to be increased to maintain thrust. Consequently the disc area is chosen to give the desired disc loading, whereas the blade area is chosen to produce sufficient rotor thrust (with an adequate load factor) at the chosen RPM. The ratio of the total blade area to the disc area is known as the solidity. Although low tip speeds and high solidity reduce profile drag and improve hovering performance, a machine built to these criteria would not have an adequate load factor for forward flight. It will be seen in section 7.2.24 that a low tip speed is also detrimental to forward flight as well as being a liability in the case of a power failure as the rotor stores little energy, giving the pilot little time to react and establish autorotation.

In a conventional helicopter there will always be a compromise between hover and forward flight performance. Possible technical solutions include variable rotor speed and variable diameter rotors.

In the hover a low rotor speed would reduce profile drag by allowing the blades to operate closer to L/D_{max}, whereas in cruise a high rotor speed would improve the load factor. To obtain maximum power from a piston engine at both speeds, a two-speed transmission would be needed, and there would be some weight and cost penalty. Changing the speed would be easy with a free turbine engine. There would also be a problem in detuning the rotor to minimize vibration at two different speeds. These problems are not insuperable.

A variable diameter rotor could have a low disc loading for efficient hover, but a raised disc loading in cruising flight. The technical problems here are enormous but may one day be economically viable.

7.2.12 Figure of Merit

Whatever the disc loading, in a constant height hover, the potential and kinetic energy of the helicopter remain constant, and so no work is being done *on the helicopter*. Thus the mechanical efficiency of all helicopters in the hover is zero. Clearly mechanical efficiency is not a useful metric, because it doesn't allow comparison. A better metric is to compare the power theoretically needed by an ideal actuator with the actual power needed in the real helicopter where the actuator has the same diameter as the rotor. This is the origin of the figure of merit.

7.2.13 Blade Twist and Taper

The forward velocity of the rotating blade increases linearly from zero at the shaft to a maximum at the tip. This is not a good starting point from which to obtain uniform inflow, but the rotor is the only practical approximation to an actuator disc that anyone has devised. The figure of merit may be improved by twisting and tapering the rotor blades. Figure 7.2-12(a) shows that in untwisted, untapered blades, the lift will be produced increasingly toward the tips as a parabolic function, until tip loss dominates.

Given the induced velocity is required to be uniform, it is easy to calculate the pitch required at any radius to make the angle of attack constant along the blade so that the whole blade operates at an efficient L/D ratio. However, in practice the induced velocity will be a function of radius and such a twist will not be optimal. Twist that takes account of induced velocity distribution is known as ideal twist. In fact ideal twist is academic because the amount of twist would have to change as a function of induced velocity and clearly it does not in any practical blade.

Ideal twist does not achieve uniform inflow, because of the reduced effectiveness of the inner slow-moving parts of the blade. Tapering the blade as well as twisting it removes area at the fast moving tips and places it in the slower moving areas where induced velocity is ordinarily lower. Ideally, the rotor chord needs to be inversely proportional to the radius. This results in an impractical rotor for a helicopter, but the technique will be found in axial blowers where the root chord may be several times the tip chord. Figure 7.2-12(b) shows a better lift distribution resulting from a combination of twist and taper. Blade taper has the further advantage that the various vibration modes of the blade will not be harmonically related. This will be considered in section 7.2-26.

In practical helicopters the blades cannot extend all the way to the mast because of the need to provide a rotor head mechanism. A further consideration is that, in the hover, the extreme roots of the blades simply provide a downwash onto the hull resulting in a download, negating some of the rotor lift. Section 7.2.20 will show that, in forward flight, the blade root of the retreating blade encounters reverse airflow. As a result blades often have a significant root cut. Ideal twist for hovering will not be ideal for forward flight or autorotation and a compromise is invariably necessary. In the proprotors of a tilt-rotor helicopter more twist will be possible because the rotors work with predominantly axial flow.

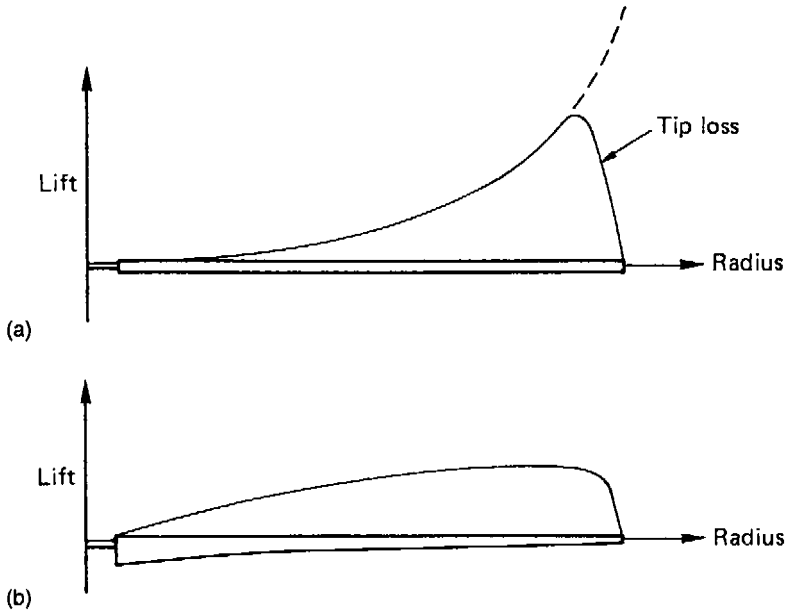


Figure 7.2-12 (a) An untwisted, untapered blade has a parabolic lift distribution and far from uniform inflow, (b) Taper and twist improve the uniformity of inflow.

7.2.14 Swirl

In a real rotor, the downwash does not just move downwards, but it also has a rotational component of motion known as swirl. Intuitively it is clear that the turning rotor must turn the air with it to some extent. Figure 7.2-10 showed that the reaction on a lifting blade is somewhat aft of the rotor shaft axis. It must follow that the direction in which the momentum of the inflow increases is in the opposite direction. The vertical component is the inflow velocity, which produces lift, whereas the horizontal component creates swirl which represents wasted power.

Some of the swirl is due to profile drag and some is due to induced drag. In the absence of profile drag, the blade reaction is still slightly behind the vertical and this must result in a small loss when compared with an ideal actuator that only causes vertical inflow. The main source of swirl is the result of profile drag adding to the rotational momentum of the air. In an ideal actuator the air leaves with increased vertical momentum only. In a real rotor the amount of swirl gives an indication of the figure of merit. The higher this is, the lower the swirl will be. In a twisted blade, swirl will be greatest near the root where the pitch angle is greater.

The torque delivered to the rotor shaft can be divided into two parts, first the torque needed to drive an ideal rotor which only produces an increase in vertical momentum sufficient to create the required thrust, and second an additional torque which is the reaction to creating swirl and tip vortices. With a single rotor, swirl energy is lost forever, but in contra-rotating rotors, the swirl of the second rotor can cancel the swirl of the first. Unfortunately this does not reduce profile drag, it only reduces induced drag.

In the single rotor helicopter, the main effect of swirl is that the downwash on the hull is not vertical, which adds to the general air of asymmetry surrounding the helicopter.

7.2.15 Vertical Autorotation

The direction of the reaction on an airfoil depends upon the angle of attack. When hovering, inflow causes the resultant to be tilted back and opposes the engine thrust. If, at a suitable height the collective pitch is lowered, the machine starts to fall, the inflow reverses and the relative airflow has an upward component. Figure 7.2-13(a) shows that the principle of the rotor blade in autorotation is no different to powered flight. It continues to accelerate air into a new direction. However, as the air approaches upwards, the acceleration has a rearward component. The reaction to this is forward, so the rotor can be driven against profile drag and transmission losses. At a suitable setting of the collective pitch, the overall reaction on the blades is brought to the vertical and so they will continue to turn at the same speed and provide lift. Note that in order to produce lift the blade must have a positive angle of attack with respect to RAF and so it is common for the blade to have positive pitch in autorotation.

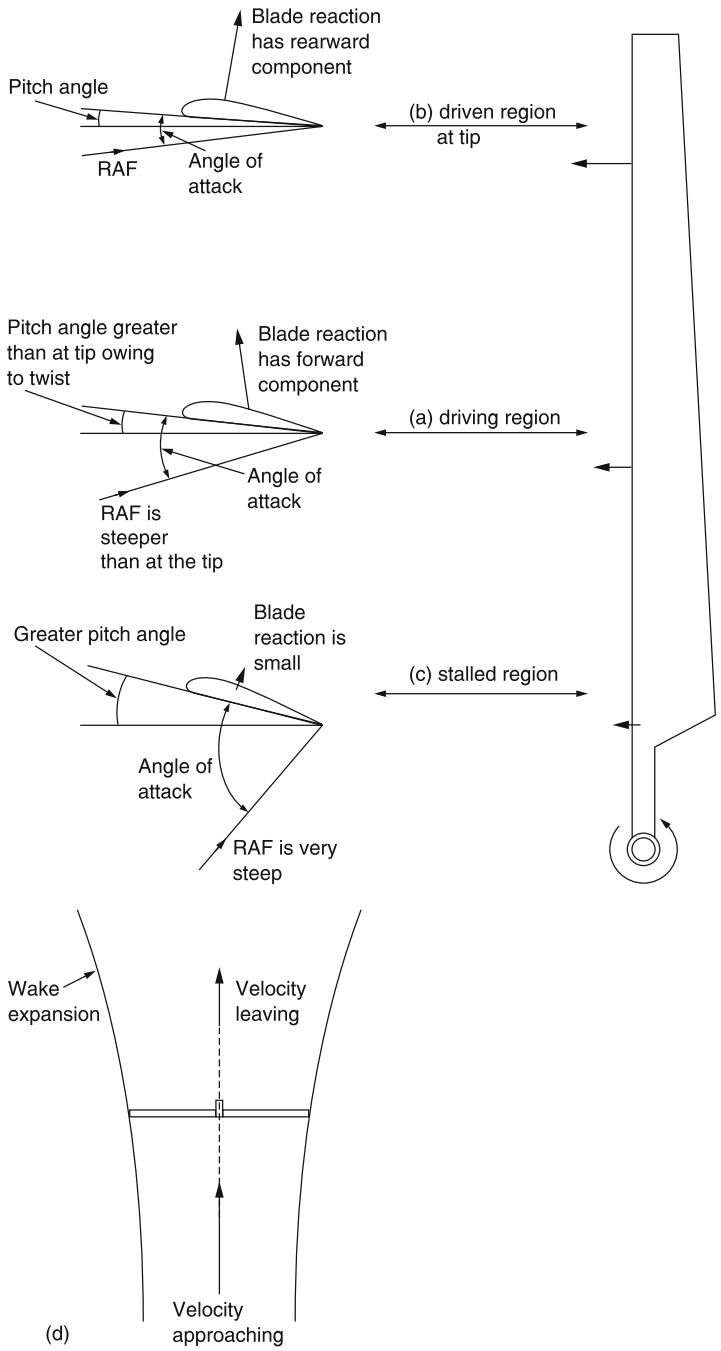


Figure 7.2-13 In autorotation, inflow is reversed. Air approaches from below so that the reaction has a forward component. (a) Middle part of blade is driving, taking momentum from the air. These conditions can be obtained with positive blade pitch. (b) Outer part of blade has a lower angle of attack and is being driven as in normal flight. (c) Inner part of blade has high angle of attack and is stalled. (d) The rotor takes momentum from the air and so the wake is slower than the approaching air and expands. This should be compared with Figure 7.2-11.

The machine is moving steadily downwards using potential energy to overcome profile drag. In fact the rotor is taking momentum from the air so that induced power becomes negative and provides the profile power needed. The strict definition of autorotation is the requirement for zero shaft torque. Where a rotor actually delivers shaft power it is said to be windmilling or in the windmill brake state. A real helicopter in autorotation is technically windmilling because the shaft torque is not zero. In the absence of the engine, some shaft torque is needed to turn the tail rotor and any other important items such as hydraulic pumps.

The speed of a blade element is proportional to radius. A further complication is that most rotor blades have twist to improve the hover figure of merit, but in autorotation the inflow is reversed and the twist is the opposite of what is

needed. It can be seen that on moving from tip to hub, the RAF decreases in speed but the angle of attack increases. As a result the central part of the blade, shown at Figure 7.2-13(a) is being driven by the inflow and is windmilling. Figure 7.2-13(b) shows that the tips have a small angle of attack and high airspeed. Profile drag dominates and L/D is poor, with the total blade reaction being behind the vertical. This area of the blade is producing lift and induced drag as if it was in the power-on condition. Figure 7.2-13(c) shows that the innermost parts of the blades are stalled as they have a high angle of attack and a low airspeed. The power needed to drive the tips and the inner parts of the blade comes from the windmilling of the centre parts. Only a small section of each blade at the boundary between the driven and driving regions is technically autorotating with a vertical total blade reaction. Thus what is meant in practice by autorotation is that the entire rotor has zero net torque. The entire rotor can also be considered to have an L/D ratio and in autorotation the slowest rate of descent will be where this is used.

In practice the loss of performance in autorotation due to twisted blades is surprisingly small. One reason for this is that the profile drag is determined only by the rotor speed and is largely unaffected by local angle of attack. Another is that real blades don't have anything like ideal taper and also have root cut-out so the stalled inner part of the rotor doesn't absorb as much power as might be expected.

In autorotation there is still a pressure step across the rotor, but the inflow is now from below. Figure 7.2-13(d) shows that as the rotor is taking momentum from the air, the slipstream from the rotor is slower than the inflow. The pressure gradient along a streamline is rising and so the wake expands.

In autorotation the best rate of descent will be where the entire rotor (not a blade section) operates at $L/D_{max.}$. The highest RRPM will also be obtained at $L/D_{max.}$, giving the greatest stored energy for landing. Figure 7.2-14 shows an autorotation diagram in which the vertical axis is drag over lift and the horizontal axis is angle of attack α. It will be seen from the figure that, at an idealized blade radius, the ratio of drag to lift determines the angle φ which must also be the angle of RAF if the blade reaction is to be vertical. If the pitch angle is θ, then $\varphi = \alpha - \theta$. The function $\alpha - \theta$ is shown on the graph to be a straight sloping line. The lowest profile drag will be at the bottom of the drag 'bucket' and the pitch angle to achieve that is where the line of $\alpha - \theta$ intersects the α axis. Most conventional helicopters will be rigged so that the minimum pitch stop on the collective lever is at about this value. Thus operation in autorotation is in the region to the right of the intersection, but not so far as the position of the dotted line where rotor speed may be irrevocably lost. In this region rotor speed may be changed by alteration of the balance of forces.

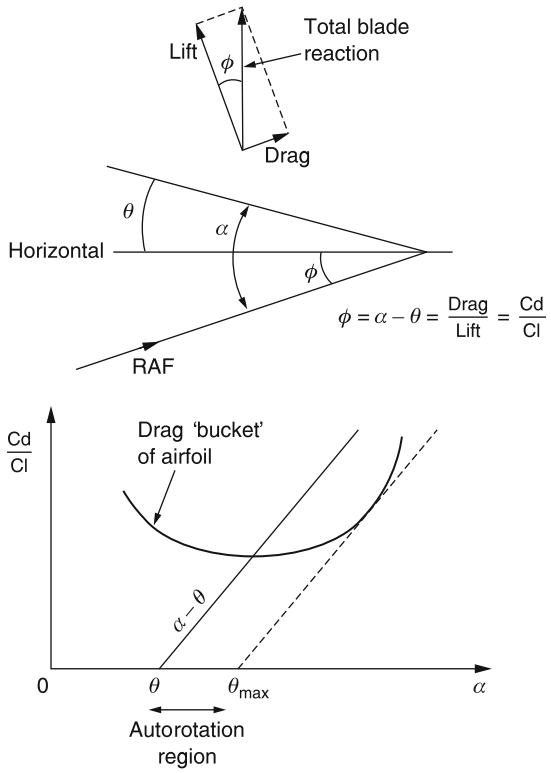


Figure 7.2-14 In ideal autorotation, the total blade reaction must be vertical as shown. For a given pitch angle θ the rotor operates at the intersection of the $\alpha - \theta$ line with the drag bucket. See text.

Reducing the angle of attack causes the blade reaction to move forward from the vertical. Negative induced power exceeds the profile power requirement until the rotor speed has increased. On the other hand, increasing the angle of attack causes the reaction to move behind the vertical, and the rotor speed falls.

7.2.16 Tip Loss and the Vortex Ring

The pressure differential between the upper and lower surface of the rotor disc causes air to try to flow transversely around the tip in order to equalize the pressure. This rotary flow combines with the blade velocity to leave a corkscrew-like trail shown in Figure 7.2-15(a) and known as a *vortex*. The loss of pressure differential due to this vortex generation cancels out the lift generated at the blade tip; an effect known as tip loss. Nearly 10% of the blade length is useless. The power needed to overcome induced drag at the tip is wasted generating the vortex. There must also be a vortex at the inboard end of the blade but it is less powerful because the lift gradient is much lower. In the steady hover and in forward flight, the tip vortices are swept away downwards by inflow.

Tip loss may be reduced by tapering and by twisting the blade near the tip so that the angle of attack is small. The pressure differential is then minimized across the tip. This cannot be taken too far as very small tip chords will operate at poor Reynolds numbers and suffer poor L/D. As a practical matter space will also be needed for tip weights.

It will be seen in Figure 7.2-15(b) that the result of non-uniform inflow is that the air disturbed by the passage of the blades changes from a disc to a cone and shrinks in diameter due to wake contraction.

In autorotation the vortices are swept away upwards. Thus in climb, hover or auto-rotation the vortices are always carried off by the slipstream. However, if a vertical powered descent is attempted, the machine descends into its own inflow and the vortices from successive blade sweeps are closer together and begin to reinforce one another. As the vertical rate of descent approaches the induced velocity, the vortices will not be swept away so much, and will begin to augment one another. The airflow will recirculate, in a toroidal motion known as a vortex ring. This is shown in Figure 7.2-16. The recirculation is increasing the inflow and so more collective pitch is needed to provide the same lift. The blade reaction will tilt back and more torque will be needed.

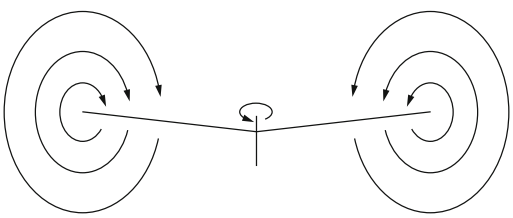


Figure 7.2-16 When the helicopter descends vertically, it can catch up with its own wake and the rotor vortices are not swept away, but result in recirculation. If allowed to develop fully, the result is the vortex ring condition.

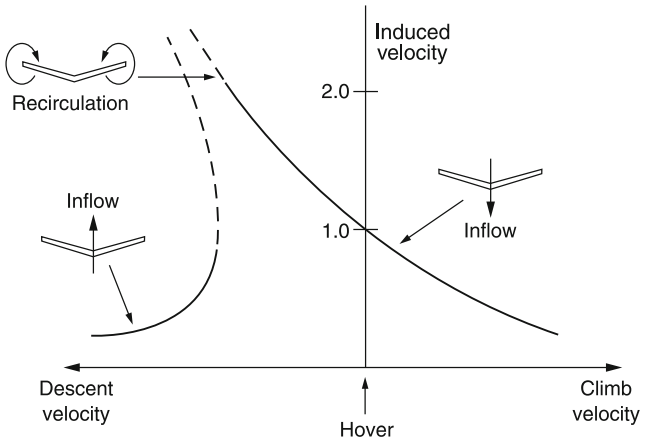


Figure 7.2-17 The relationship between induced velocity and vertical climb/descent velocity. Both have been normalized with respect to induced velocity in the hover. Note there is a discontinuity between powered descent and autorotation due to the vortex ring condition.

Paradoxically, a helicopter needs more power to descend in the hover than it does to maintain height. Early helicopters having marginal power used low disc loading and so had low induced velocity. Some suffered from power settling where above a certain rate of vertical descent the descent could not be arrested except by moving into forward flight.

Figure 7.2-17 shows the relationship between the induced velocity needed and the climb/descent velocity. As the effects are proportional to the hover induced velocity, Figure 7.2-17 has been normalized by dividing both parameters by that factor. As a result the curve passes through unity in the stationary hover. Note that as the rate of descent increases, following the curve to the left, there is a discontinuity where a fully developed vortex ring condition occurs. The recirculation has now enveloped the rotor whose ability to produce lift is considerably reduced. The effectiveness of the cyclic pitch will also be reduced. Increasing collective makes it worse. The machine drops like a stone, pitching and rolling randomly. If this condition is entered close to the ground, recovery may be impossible.

To create lift a rotor must have mass flow whose momentum it can increase. In hover or climb, the inflow is downward and the rotor puts momentum into the inflow by accelerating it. In a vertical autorotation the inflow is upward (effectively making the momentum negative) and the rotor puts momentum into the air by slowing it down. Both of these conditions are stable and the results are predictable from actuator disc theory. The problem is that in vertical flight to get from downward inflow to upward inflow the state of zero inflow must be traversed.

Zero inflow makes nonsense of actuator disc theory, and the real helicopter can't fly in that condition either, hence the discontinuity in the curve of Figure 7.2-17. The position of the discontinuity is affected by blade twist. A twisted blade produces more thrust at the inner parts of the disc and excites the vortex ring less, so it is possible to descend faster.

It must be accepted that there are certain combinations of vertical rate of descent and induced velocity in which a helicopter will be uncontrollable. Thus an important part of pilot training is to impart a practical knowledge of how to fly outside that region, how to recognize the onset of a vortex ring and how to recover from it.

Using forward speed to guarantee an inflow allows the vortex ring condition to be avoided. Figure 7.2-18 shows that to enter autorotation from forward flight the collective is simply lowered until the inflow reverses. Once in autorotation the forward speed can be brought to zero for a vertical descent. In order to terminate the vertical autorotation forward speed is regained before raising collective pitch again to obtain normal downward inflow.

7.2.17 Ground Effect

When hovering close to the ground, the air passing downward through the rotor cannot escape as freely as shown in Figure 7.2-19(a). Air pressure below the machine builds up, and reduces the induced velocity. The RAF is closer to the horizontal and the blade resultant is closer to the vertical. The same lift can be obtained from the rotor with less rotor drag. As Figure 7.2-19(b) shows, considerably less power is needed to hover close to the ground. As the figure shows, ground effect is noticeable up to a height approximately equal to the rotor radius. In still air, a helicopter will reach equilibrium in ground effect at a height where the lift just balances the weight without any input to the

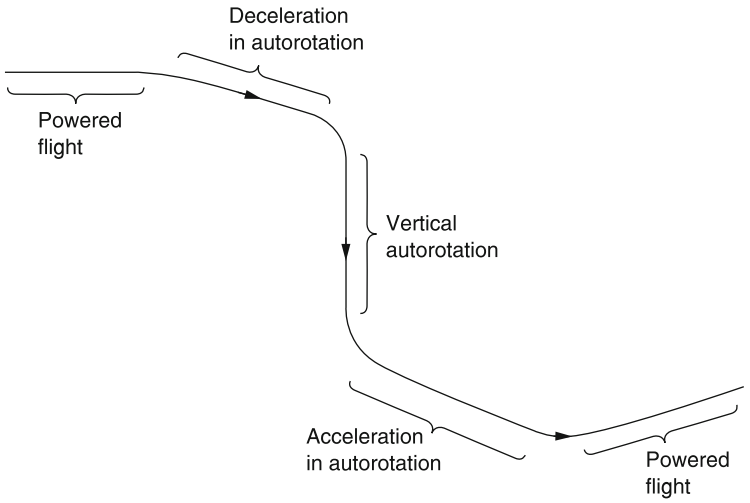


Figure 7.2-18 Avoiding the vortex ring. Autorotation is entered in forward flight, and then the forward speed is brought to zero. Prior to resuming powered flight, forward speed must be regained.

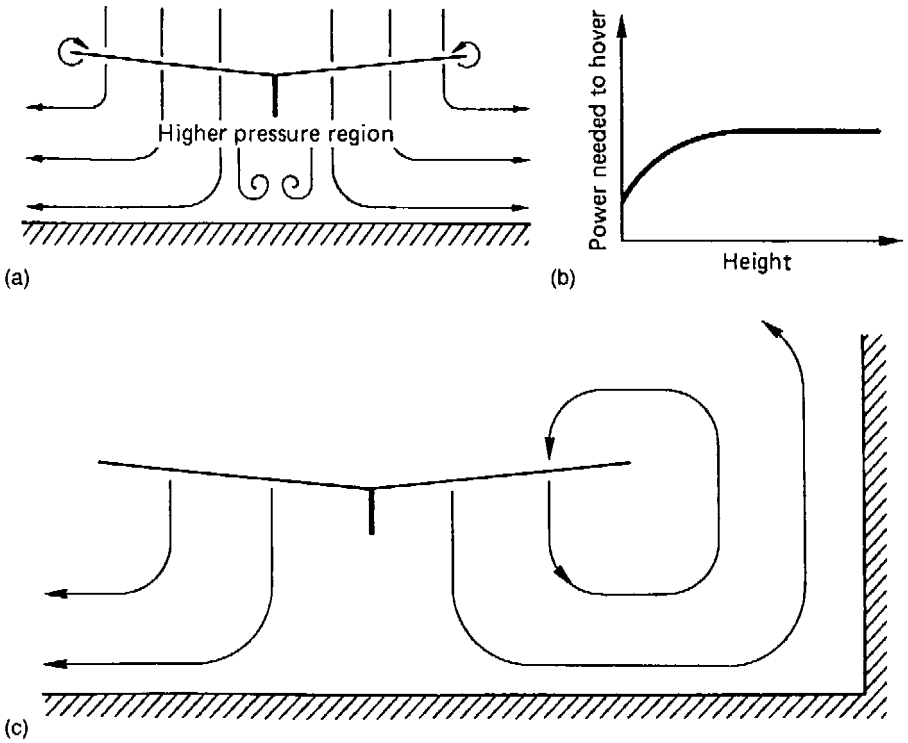


Figure 7.2-19 (a) When hovering close to the ground, a pressurized air cushion builds up under the rotor, reducing the induced power needed to hover (b). Hovering close to a building (c) may result in the downwash being returned to the rotor. Recirculation causes loss of lift.

collective pitch control. If the machine descends, it will obtain more lift from the air cushion; if it climbs, the lift will fall. In many respects the helicopter in ground effect is like a hovercraft.

A helicopter in ground effect over water will displace some of its weight in a saucer-shaped depression in the water surface. This is not often seen because hovering low over water is a recipe for spray in the engines.

The presence of the rotor head and blade root cut-out creates a hole in the middle of the disc through which air can escape upwards in ground effect. This is actually another form of tip loss, which is called the *fountain effect*. In practice

the fountain effect can be beneficial because it puts some of the hull into upflow and reduces the hull download. There will be vortices trailing from the inboard end of the blade. As the lift gradient is small here these are relatively weak and their main contribution is that in ground effect they make the hover conditions chaotic especially if they interact with the tail boom or tail rotor.

It is often stated by pilots that hovering over rough ground 'dissipates' the ground effect, when theory would suggest that restricting the ability of the downwash to escape would enhance ground effect. This aerodynamic effect ought to be very small, and the real effect may be psychological. When hovering over a crop, the pilot will estimate his height from the top of the crop, not from ground level. Thus the ground effect may not appear as powerful.

Ground effect normally increases lift so that it offsets the effect of vortex ring generation. The two effects are always opposing one another. However, hovering close to a wall or next to another helicopter can trigger recirculation (Figure 7.2-19(c)).

7.2.18 Cyclic Control

In order to move around in the hover, the pilot pushes the cyclic stick in the direction he wants to go. This superimposes a cyclic variation of blade pitch on the average or collective setting. The result is that the blades oscillate sinusoidally about the feathering axis as shown in Figure 7.2-20(a) at the same frequency as the rotor turns. As a result the lift will increase on one side of the rotor and decrease on the other, resulting in a rolling couple. The rotor is gyroscopic and so the cyclic pitch change is arranged to occur 90° ahead of the required result to allow for precession in the rotor.

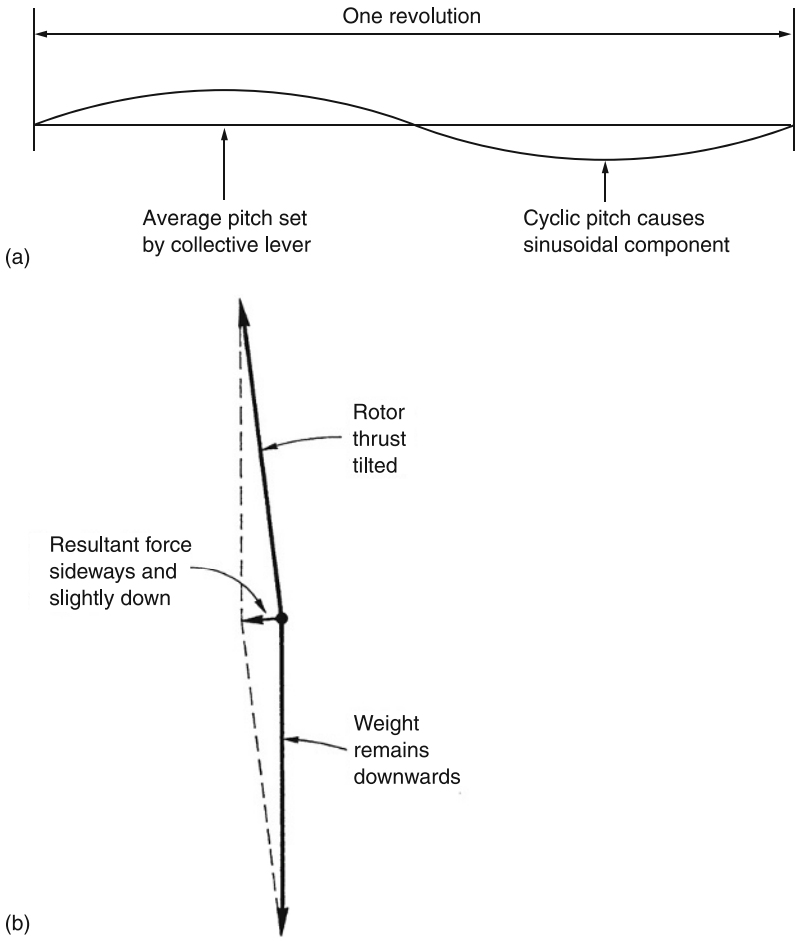


Figure 7.2-20 (a) The cyclic control causes the pitch of the blades to change sinusoidally at the rotor frequency. The rotor disc will tilt with a 90° phase lag. (b) Tilting the disc has the effect of reducing the vertical component of rotor thrust. In practice a slight increase in collective is required to maintain height as the cyclic control is used to move off.

Figure 7.2-20(b) shows that the application of cyclic control tilts the thrust vector, and that there will be a resultant of the thrust vector and gravity which is primarily horizontal but also has a slightly downward component. This is because the vertical component of the rotor thrust becomes smaller when the disc is tilted. As a result the helicopter will accelerate sideways and lose height. To compensate, the pilot slightly raises the collective lever so that the vertical component of the thrust remains the same.

7.2.19 Basic Manoeuvres

From the hover, the helicopter can be accelerated into translational flight by using cyclic pitch to tilt the disc. The greater the acceleration required, the more extreme will be the disc tilt and the greater the application of collective pitch. Figure 7.2-21(a) shows a helicopter at maximum horizontal acceleration. If done at low height, the blades can get very near the ground. Figure 7.2-21(b) shows that at constant speed and height the thrust necessary to balance the drag of the pure helicopter comes from the horizontal component of the inclined rotor thrust. The vertical component of the rotor thrust must balance the weight of the machine. Clearly the rotor thrust must be greater than in the hovering condition. The diagram simplifies the situation in that it assumes the drag of the fuselage acts at the rotor hub, which is not actually the case. In practice the drag acts below the hub to produce a couple which would depress the nose. There are two ways of dealing with that: using a rotor head which can apply an opposing couple from the blades and/or using a tail plane to produce a down thrust.

Note that the tip path plane may be inclined forwards with respect to the hull so that the balance of forces is obtained with the hull level. In this condition the drag will be minimized and the occupants will be most comfortable. A consequence is that when the machine comes to the hover it will sit with a tail down attitude. The length of the

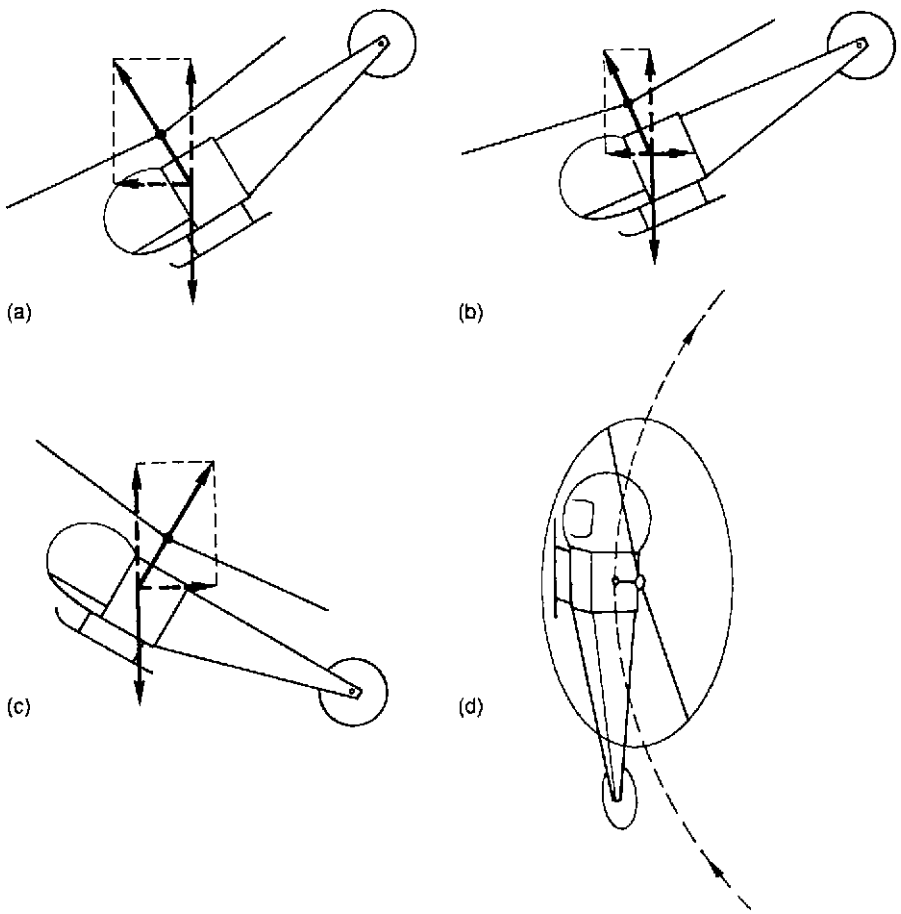


Figure 7.2-21 (a) The forces involved during maximum acceleration. Note the front of the disc can be very close to the ground. (b) The balance of forces in forward flight. The forward component of rotor thrust balances hull drag and the vertical component balances the weight. (c) The quickstop is a horizontal autorotation. Note the low position of the tail rotor. (d) A turn at speed is limited by the increased thrust necessary to maintain height whilst accelerating the machine towards the centre of the turn.

undercarriage legs or the angle of the skids is often arranged such that the machine can settle on the ground very nearly in this attitude.

Tilting the disc accelerates the machine along, but bringing the disc back to the horizontal only removes the acceleration, and the machine will continue along, slowing only because of drag. In order to stop, the disc must be tilted the opposite way to obtain acceleration in the opposite sense to reduce the velocity. Figure 7.2-21(c) shows a helicopter with maximum deceleration in a manoeuvre known as a quickstop. Note that the inflow has reversed so that a quickstop is actually a form of autorotation. Note that the tail boom is very low, and could strike the ground.

Figure 7.2-21(d) shows a turn at speed. The turn changes the direction but not the speed, so it is a change of velocity or acceleration. The acceleration must be towards the centre of the turn, and it is necessary to obtain a resultant sideways force to cause that acceleration. This is done by banking with lateral movement of the cyclic stick. Collective will need to be increased to maintain the vertical component of thrust equal to the weight. The machine will then fly around as if on the surface of a cone until the desired heading is reached, when the bank will be taken off and the collective reduced again. Note that in a 60° bank the thrust vector has to be twice that needed to hover in order to maintain height. The ratio of manoeuvre thrust to hover thrust is called the load factor. Not all machines can deliver that much power, and not all transmissions can accept it. However, some large transport helicopters can carry their own weight in cargo. When unladen, their load factor effectively doubles and they can safely accomplish some quite alarming manoeuvres.

Note that during acceleration, deceleration and a correctly banked turn gravity still appears to act vertically down through the cockpit floor, and it is possible to complete an entire flight without spilling a mug of coffee placed on the floor.

7.2.20 In Translational Flight

In translational flight, the forward velocity increases the velocity of the air passing the advancing blade and reduces that seen by the retreating blade. Figure 7.2-22 shows what happens. In (a) it can be seen that the velocity of a blade element due to rotation is given by the distance D from the shaft multiplied by the angular speed ω. When the blade is at 270°, the edge of the reverse flow region will be where ωD equals the airspeed V. In (b) the method for finding what happens at an arbitrary blade angle is shown. At any distance d from the shaft, the vector resultant R of ωd and V can be found. Figure 7.2-22(c) shows the derivation of the expression for d, the distance from the shaft to the boundary of the reverse flow region. Figure 7.2-22(d) shows the expression for d graphically and it will be seen that the reverse flow area is circular and that its diameter increases with speed. Instantaneously the rotor is turning with respect to the outer edge of the reverse flow region. The ratio of forward speed to tip speed is known as the *advance ratio* μ which is the same as the ratio of the diameter of the reverse flow region to the blade radius.

The retreating blade suffers reverse flow near the root; a small downward force may be experienced because the retreating blade effectively has a negative angle of attack due to flow reversal. Here is another reason for root cut-out, to reduce the amount of useless reverse-flowed blade. However, as the area is proportional to the square of the diameter, the amount of disc area lost to the reverse flow region is small. For example, at $\mu = 0.4$, the diameter of the reverse flow region is 0.2 of the rotor diameter but the area lost is only 4%.

Owing to the huge difference between the airspeeds, if nothing were done, the advancing blade would produce more lift and the retreating blade less. This would apply to the rotor a roll couple toward the retreating side. The gyroscopic action of the rotor would result in a response delayed by 90° of rotation corresponding to a rearward pitch effect. This rearward couple is colloquially known as 'flapback' and it tends to reduce the forward tilt of the machine and consequently reduces the forward component of the rotor thrust, slowing the machine down.

Thus sustained forward flight requires a continuous application of forward cyclic control that reduces the angle of attack of the advancing blade and increases that of the retreating blade in order to balance the lift moments on the two sides of the rotor. In cruise the amount of cyclic pitch applied may be of the order of 5°. Flapback results in speed stability because it automatically opposes the cyclic control.

The more forward cyclic is applied, the higher will be the forward speed. Thus in translational flight the fore-and-aft position of the cyclic stick primarily controls the airspeed. If the airspeed is changed, the collective lever will need to be adjusted to maintain height because tilting the rotor thrust vector alters its vertical component.

In forward flight the rotor has access to a greater mass of air. Thus the same momentum increase and the same thrust can be obtained with a smaller induced velocity and this reduces the power needed. Figure 7.2-23 shows that there is a minimum induced power speed where the inflow velocity is the least.

7.2.21 Inflow and Coning Roll

In flight, there is a region of low pressure above the rotor and a region of high pressure below. This pressure difference results in a tendency for air in the plane of the rotor to move upwards. Figure 7.2-24(a) shows that this happens for

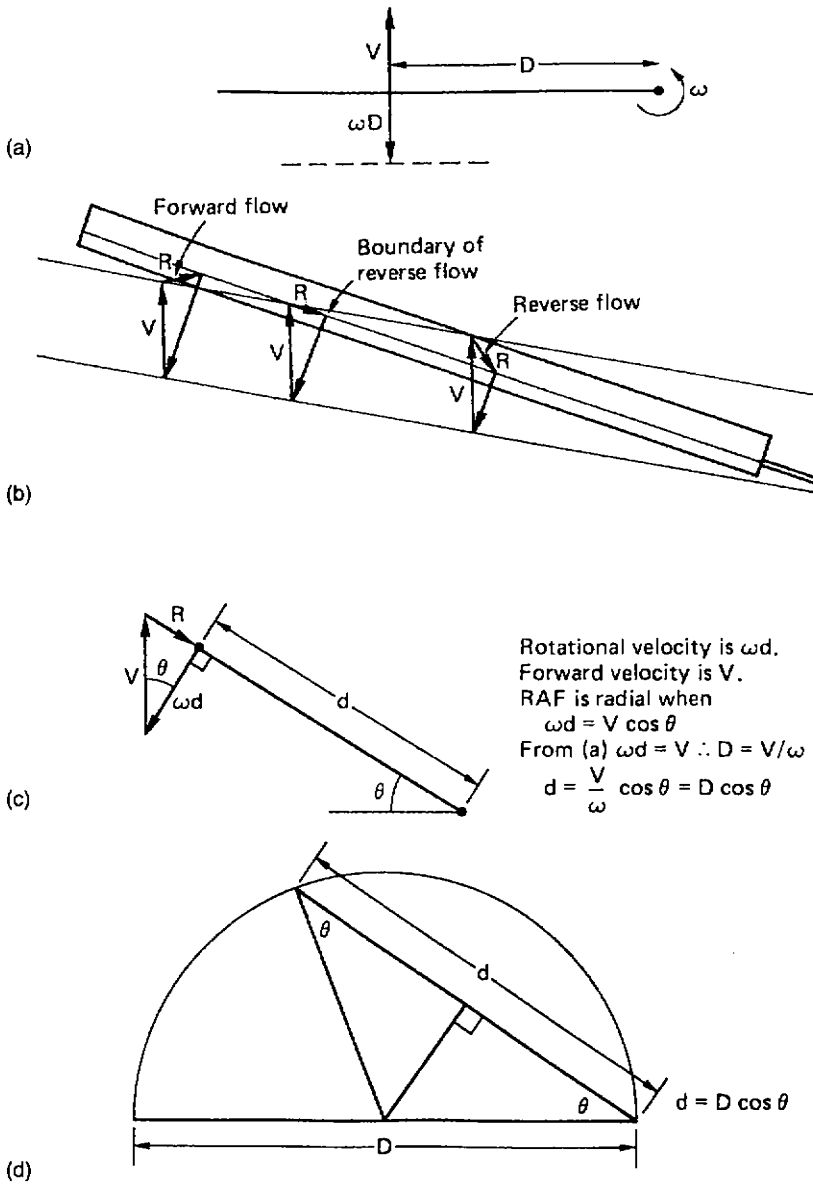


Figure 7.2-22 The reverse flow region can be shown to be circular. See text for details. Its diameter is determined by the ratio of the forward speed to the tip speed, which is the advance ratio μ.

some distance from the edge of the disc. Simple inertia keeps the air moving upwards so that in translational flight, the leading edge of the disc encounters upwash. The upwash ceases when the downward impulse of the rotor cancels the upward momentum of the air mass.

In slow forward flight, Figure 7.2-24(b), the low pressure above the rotor has time to bend the inflow increasingly downwards as it reaches the trailing edge of the disc. The disc can be thought of as a low aspect ratio wing producing tip vortices which roll inward at the top and down through the rear of the disc. The result is a reduction of the angle of attack of the blade at the rear of the disc compared to that at the front. The loss of lift at the rear of the disc is subject to the phase lag, or precession, of the rotor and manifests itself as a roll to the advancing side, which is known as inflow roll.

Whilst coning causes no undue aerodynamic problems to a hovering helicopter, it has an unwanted effect in forward flight. Figure 7.2-24(c) shows that the effect of coning is to reduce the angle of attack of the blade passing across the tail compared with that of the blade passing across the nose. As for the inflow roll, this results in a tail-down moment applied to the disc, which likewise manifests itself as a roll towards the advancing side because of the phase lag of the rotor. The two effects take place in parallel, but as different functions of airspeed. Inflow roll commences on leaving the hover but reduces significantly as speed increases, whereas coning roll slightly increases with airspeed.

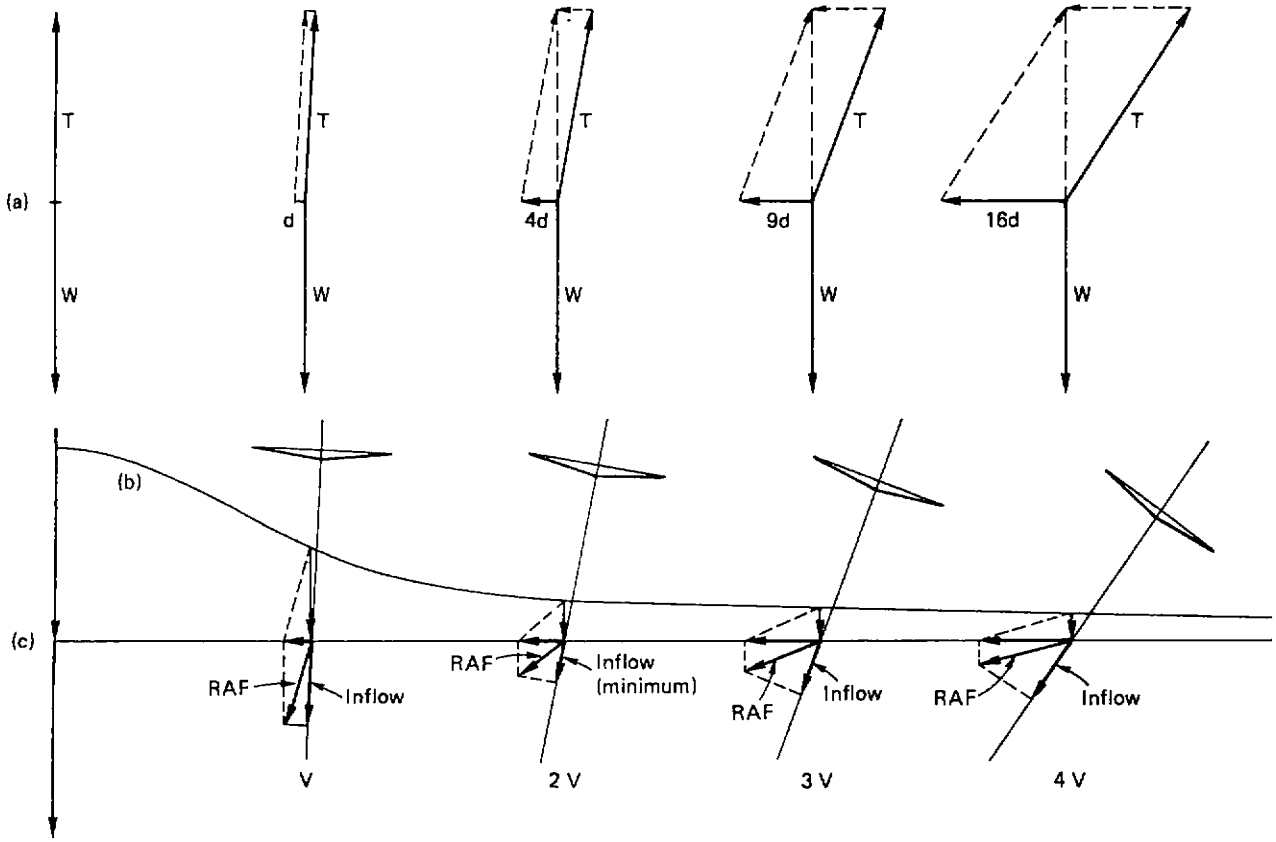


Figure 7.2-25 shows how the lateral cyclic control might have to be moved to the retreating side in order to fly straight at different airspeeds.

Thus in straight flight at constant speed, the cyclic stick must be held slightly towards the retreating side to counteract the inflow and coning induced rolls as well as forwards to counter the advancing/retreating induced pitch-up. In some machines the phasing of the controls is arranged so that forward cyclic automatically creates a degree of lateral cyclic. However, in the hover this results in the machine not responding precisely in the direction the stick is pushed.

Like the collective control, there is no one correct setting for the cyclic control. The cyclic trim control is used to shift the neutral position of the stick, relieving the pilot of the need to produce continuous control forces.

7.2.22 Rotor H-force

The retreating blade has the airspeed subtracted from the rotational speed and so has to operate with an increased angle of attack. The profile drag will be reduced but the induced drag will increase. Conversely the advancing blade will suffer significantly more profile drag because of the high relative airspeed, but the induced drag will fall because the angle of attack has been reduced. The overall effect of profile drag and induced drag does not balance between the advancing and retreating sides and the resultant is a rearward acting force called the H-force. Figure 7.2-26 shows that the H-force is the reason why the rotor thrust is not precisely at right angles to the tip path plane. A typical figure is 1°.

Because of inflow and coning, the blades crossing the nose produce more lift than the blades crossing the tail and so the induced drag is not the same. The resultant is the Y-force which acts at right angles to the direction of flight. The H-force is small and the Y-force is very small with the result that the effect on the magnitude of the thrust is insignificant.

7.2.23 Blade Stall and Compressibility

Stall occurs when air passing over the cambered upper surface of a blade can no longer accelerate rapidly enough to follow the surface. Airflow breaks away and lift is lost accompanied by an increase in drag. As acceleration is the

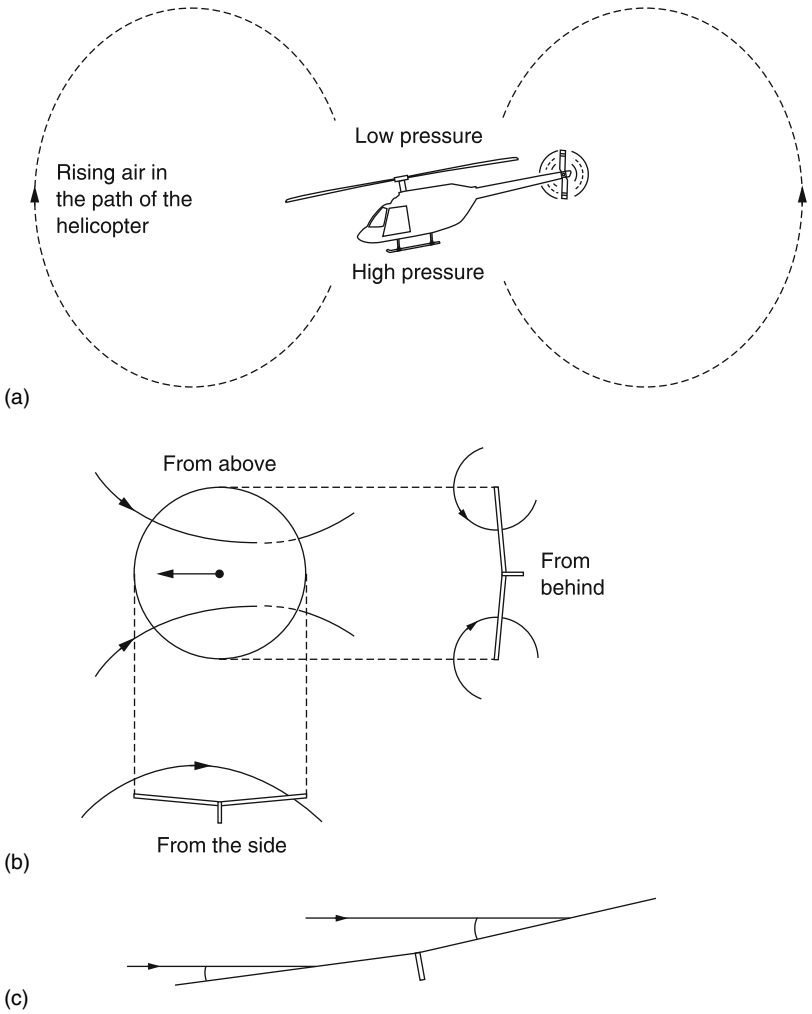


Figure 7.2-24 (a) Given the pressure difference across a rotor in the hover, surrounding air will tend to move upwards. In translational flight the edge of the rotor may actually encounter upwash. (b) In forward flight the rotor acts like a low aspect ratio wing. Vortices curl over the tips and down through the rear of the disc, reducing the angle of attack there. Loss of lift at the rear precesses into inflow roll. (c) As the rotor blades are coned, in forward flight the air approaches at a different angle as shown here. The component perpendicular to the blade is greater for blades at the rear, where the angle of attack is reduced.

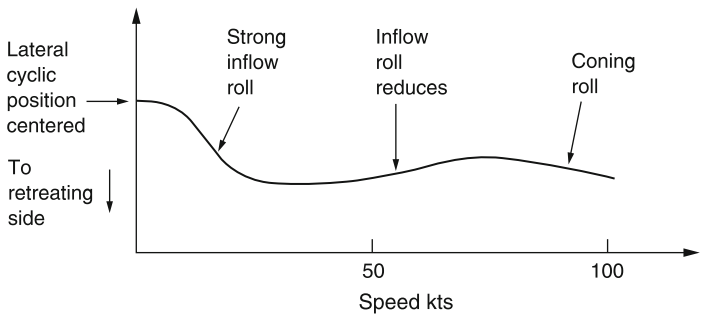


Figure 7.2-25 How the cyclic stick might be trimmed at different airspeeds. Inflow roll begins at low airspeed but has less effect as speed rises, whereas coning roll increases with speed as the coning angle may increase to produce more thrust.

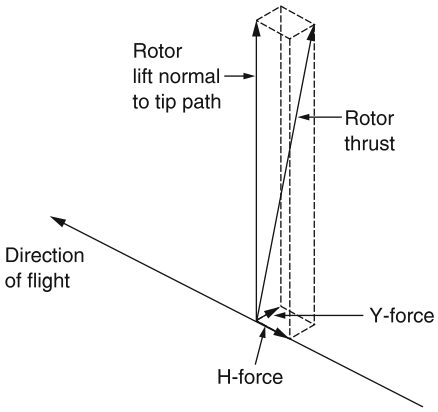


Figure 7.2-26 H-force and Y-force. Conditions experienced by the blades in forward flight are asymmetrical and this reflects as cyclic differences in blade drag that produce a force in the plane of the rotor. The rotor thrust is not precisely at right angles to the tip path plane because of this force. H- and Y-forces are the components of this force in the fore-and-aft and transverse directions. The effect is very small and not really noticeable to the pilot.

limiting factor, at low relative velocities air can negotiate the blade at a larger angle of attack than at high velocity. Figure 7.2-27(a) shows the stall limit diagram for a typical section. As the speed of sound is approached, the allowable angle of attack becomes very small. Note that the speed of sound falls with falling temperature and in very low temperature conditions helicopters with high tip speed will suffer a performance loss.

In fast translational flight, the tips of the rotor blades will be encountering relative airflow that alternately adds to and subtracts from the rotational speed. The pitch angle will change sinusoidally owing to the use of cyclic feathering but the angle of attack follows a more complex function owing to the effects of coning and inflow. Figure 7.2-27(b) shows the angle of attack of a part of the blade near the tip plotted against the relative airflow velocity. It will be seen that the plot is an elongated figure of eight. At positions A and C the blade movement is transverse, whereas at B and D the forward speed is subtracted from or added to the rotational velocity.

When the plot is superimposed on the stall limits of the airfoil section, it becomes possible to predict where blade stall will occur. This has been done in Figure 7.2-28. Figure 7.2-28(a) shows the conditions near the maximum airspeed. (b) shows that with a heavily loaded machine, stall can occur on the retreating blade where it has very low relative speed. The retreating blade needs to be feathered to a large angle of attack to resist lift asymmetry and inflow roll. The peak angle of attack will be reached at about 285° and the length of stalled blade will be greatest here. The loss of lift around 285° is converted by rotor phase lag to a pitch-up and roll. An attempt to correct the roll with cyclic will accentuate the stall and loss of control will result. Figure 7.2-28(c) shows that with a lightly loaded machine in straight and level flight or in a dive, advancing blade compressibility stall will occur first at high forward speed.

Partial blade stall can also be provoked at moderate speeds by tightly banked turns and if this is detected, the severity of the manoeuvre should be reduced immediately. Blade stall of either type causes serious vibration and control difficulty before the loss of lift becomes significant. Serious alternating stresses are set up and can cause blade delamination.

The never-exceed-speed for the machine, V_{ne}, will be set at a value which avoids advancing blade compressibility stall.

7.2.24 The Speed Limit

The pure helicopter is forever trapped in a forward flight region where it can retain control and contain blade forces. Figure 7.2-29 shows the constraints. The airspeed is added to the tip speed on the advancing blade, and this must be kept below about 0.92 of the speed of sound to avoid excessive noise and blade forces. On the other hand the airspeed is subtracted from the tip speed on the retreating blade and the advance ratio has to be kept below about 0.5 to avoid a large reverse flow area and retreating blade stall. As a result the pure helicopter with ideal tip speed is unable to exceed about 200 knots at the extreme right of the envelope of Figure 7.2-29.

If outright speed is not a priority, the possible range of tip speeds is greater. The upper limit is set by noise and the lower limit by the requirement to store enough kinetic energy to handle an engine failure. If a rotor has to provide thrust as well as lift it will be tilted well forward at high speed and the hull will be creating a lot of drag because of its nose down attitude. Higher speeds can be reached if an auxiliary form of forward thrust is available because the rotor thrust can then be vertical and minimized and the hull attitude will be better.

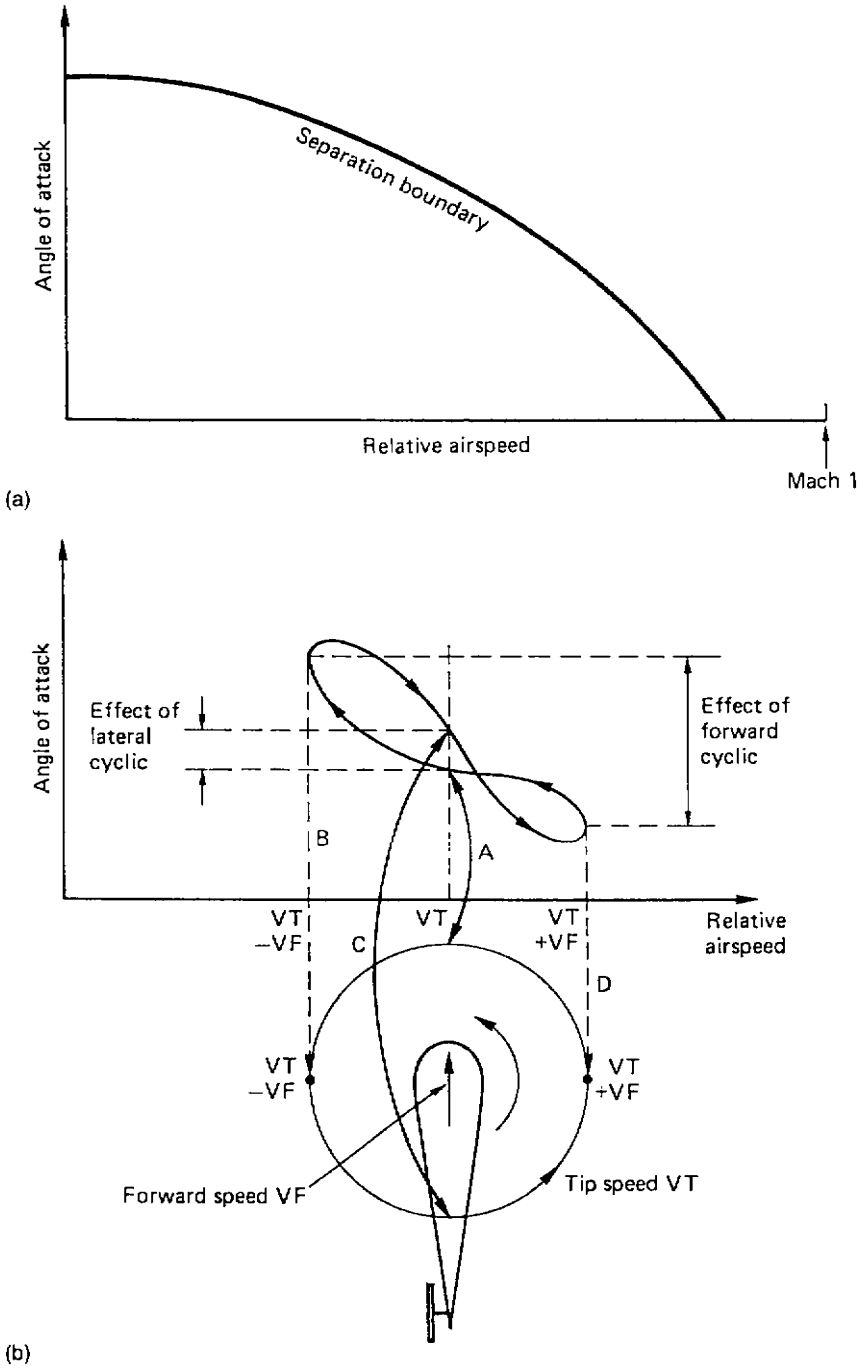


Figure 7.2-27 (a) As speed rises, air finds it increasingly difficult to follow the camber of an airfoil and the result is that the angle of attack must be limited as a function of speed to prevent stall. (b) The angle of attack of a blade may vary in the way shown here. The sinusoidal component due to cyclic pitch is modified due to variations in inflow over the disc.

Westland modified the turbine exhausts to produce thrust on the speed record-breaking Lynx, and the Lockheed Cheyenne had a second tail rotor facing rearwards to provide thrust for forward flight as well as a fixed wing to unload the rotor at high airspeeds.

7.2.25 Harmonic blade motion

So far the disc has been discussed as an entity since the disc attitude determines the path of the machine. Now the motion of the blades within the disc will be explored.

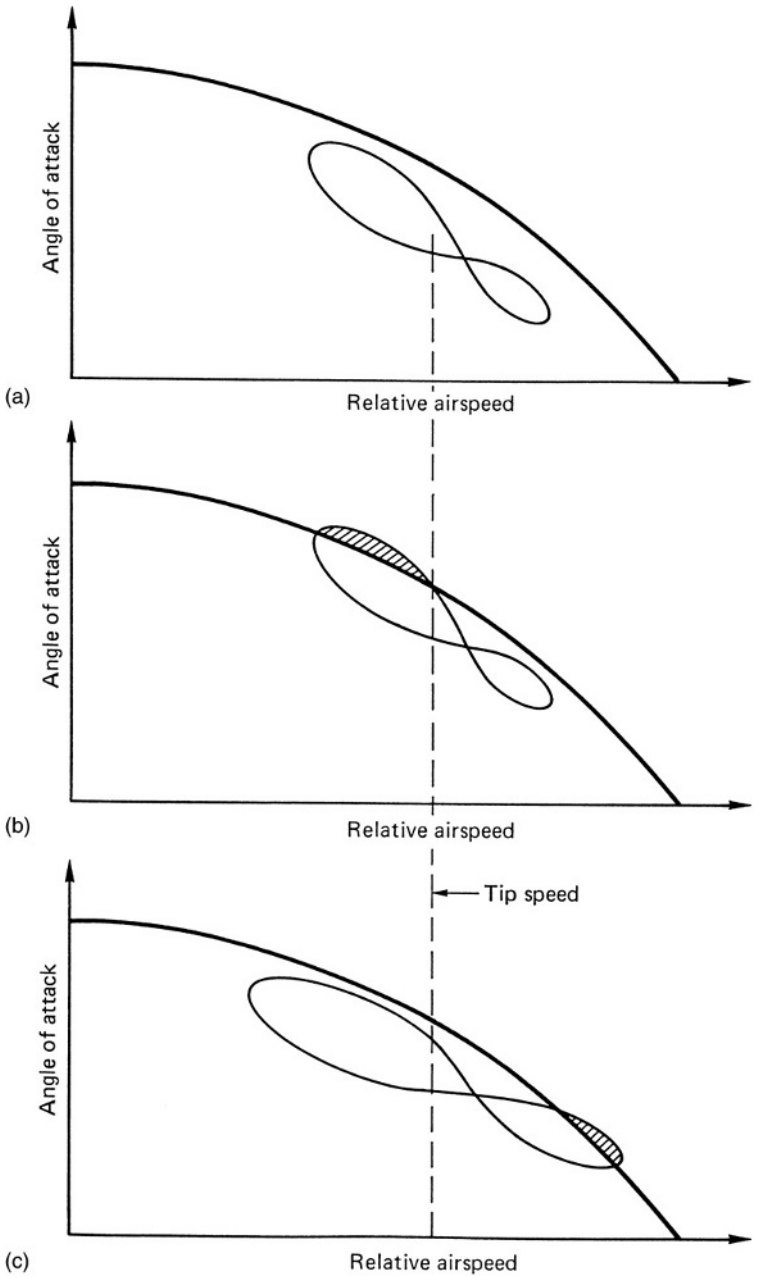


Figure 7.2-28 (a) No separation. (b) Retreating blade stall due to high rotor thrust. (c) Advancing blade compressibility at high airspeed.

In trimmed translational flight the application of cyclic feathering opposes the lift asymmetry. Cyclic feathering changes the blade pitch angle sinusoidally. If the lift asymmetry were also sinusoidal, the two effects would be in constant balance. Unfortunately this is not the case. Sinusoidal cyclic feathering is not strictly what is needed, but for practical reasons it is the most widely used solution. Cyclic feathering can only make the *average* lift moment the same on both sides of the disc. It cannot keep the lift of an individual blade constant at all angles of rotation.

Figure 7.2-30 shows what happens to a blade element as it rotates in forward flight with cyclic pitch applied. The diagram assumes uniform inflow across the disc and neglects coning for simplicity. Figure 7.2-30(a) shows the relative airspeed experienced by a blade element. The rotational speed produces a constant component, but the forward speed appears as a sinusoidal component to the revolving blade, adding speed to the advancing blade and subtracting it from the retreating blade. The lift generated by a blade is proportional to the square of the airspeed and this parameter is shown in Figure 7.2-30(b). Note that the squaring process makes the function at 90° very much greater than at 270°.

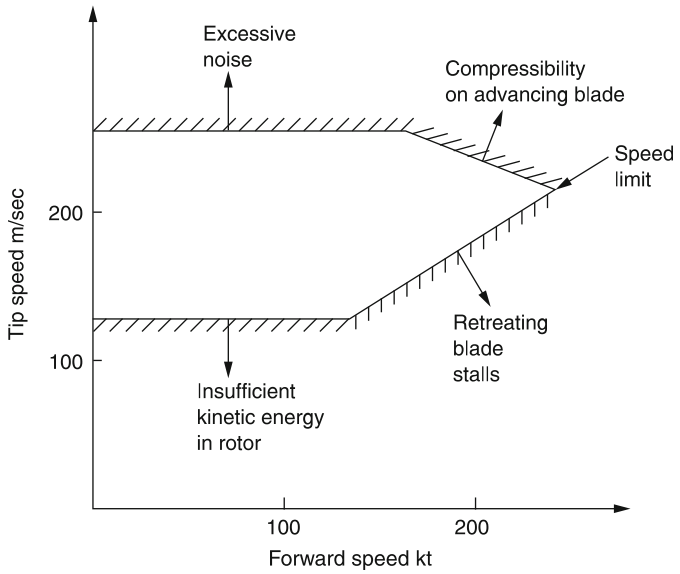


Figure 7.2-29 Pure helicopter speed is forever limited by the coincidence of advancing blade compressibility and retreating blade stall at the right of the diagram. Only compound helicopters can break this limit using auxiliary thrust and/or wings to offload the rotor. At lower speeds tip speed is constrained by noise and stored energy limits.

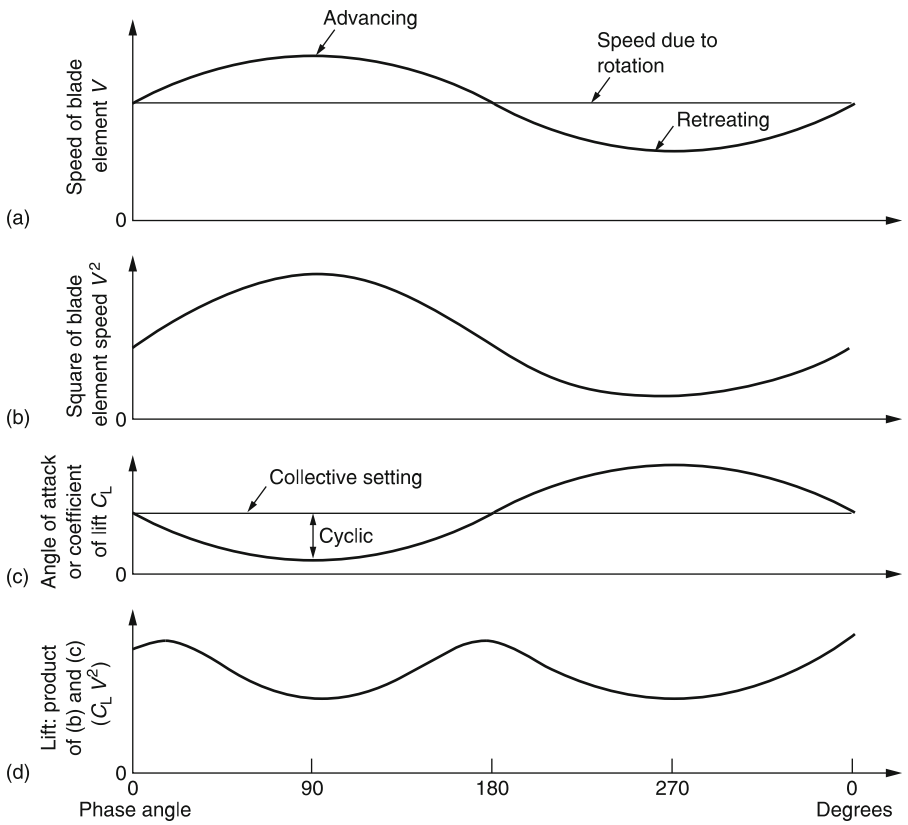


Figure 7.2-30 (a) Instantaneous airspeed experienced by a certain point on a blade is a fixed average added to a sinusoidal variation. (b) Lift is proportional to airspeed squared. (c) Sinusoidal control of angle of attack due to application of forward cyclic. (d) Non-constant lift around the disc is a source of vibration.

The coefficient of lift of the blade element is nearly proportional to the angle of attack. The collective pitch setting results in a constant component of the angle of attack and the cyclic pitch control causes this to become a shifted or offset sinusoid as shown in Figure 7.2-30(c). Figure 7.2-30(d) shows the lift function which is the angle of attack multiplied by the square of the airspeed. In other words waveform (d) is the product of (b) and (c).

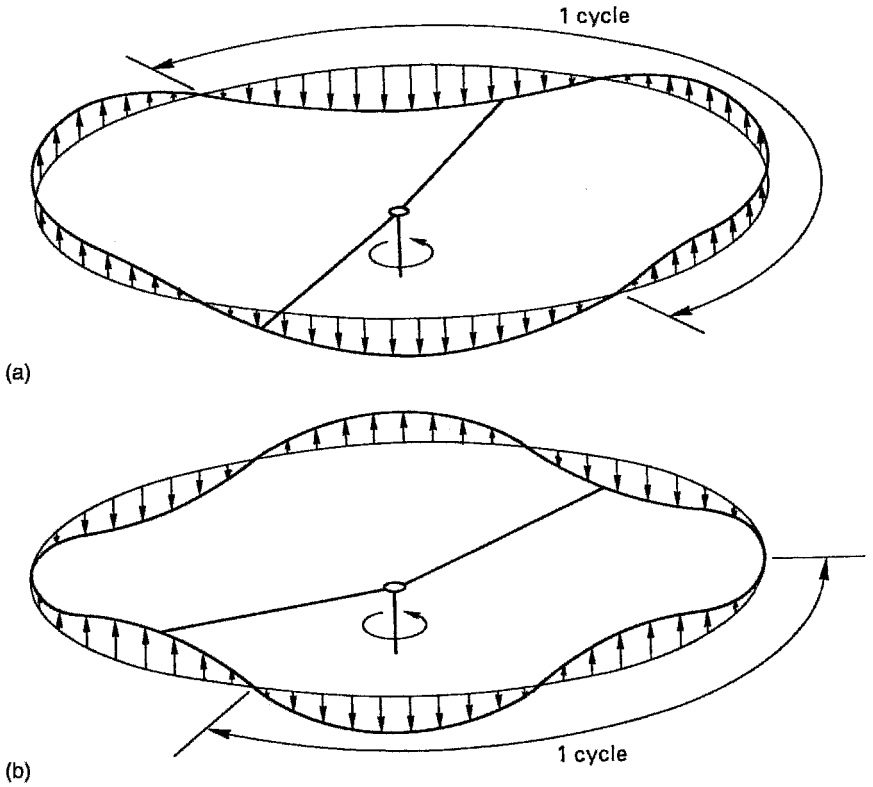


Figure 7.2-31 Lift on a given blade is not constant, but a function of azimuth angle. This results in harmonic blade flapping. (a) Shows second harmonic flapping. (b) Shows third harmonic flapping.

The amplitude of the cyclic control in (c) has been chosen to make the lift at 90° the same as the lift at 270°. Note that although the lift on the two sides of the rotor has been made the same by the application of cyclic feathering, the lift is not constant.

Unfortunately a sinusoidal function cannot cancel a sine-squared function. There is a lift trough at about 270° where the square of the relative airspeed has become very low and increasing the angle of attack fails completely to compensate. Lift symmetry is only obtained because the same cyclic input also strongly reduces the angle of attack at 90°. As a result there will be lift troughs at 90° and 270°.

The lift function contains a Fourier series of harmonics, or frequencies at integer multiples of the fundamental, which is the rotor speed. As the lift troughs on the two sides of the rotor are not the same shape, significant levels of odd harmonics will exist, especially the third and fifth harmonics. The lift function excites the blades in the flapping direction and they will respond according to the dynamic characteristics of the blade and its damping. The result is that the blades do not describe a perfect coning motion in forward flight, but as shown in Figure 7.2-31, they weave in and out of the cone due to the harmonics. Figure 7.2-31(a) shows second harmonic flapping, completing two cycles per revolution, whereas Figure 7.2-31(b) shows third harmonic flapping. The flapping action directly modulates the axial thrust delivered to the mast causing a hull motion variously described as 'hopping' or 'plunging'.

7.2.26 Sources of Vibration

Whilst vibration in the hover can be minimized by attention to dynamic balance and tracking, vibration in translational flight is inevitable because of the varying conditions caused by the alternate advancing and retreating of each blade. This causes harmonic flapping and dragging of the blades, resulting in forcing functions at the hub which may be axial, i.e. along the mast, lateral, i.e. tending to cause rocking, or torsional, i.e. tending to twist the mast.

On the retreating blade, the lift on the blade is concentrated near the tip with download near the root. This is because of the reverse flow area shown in Figure 7.2-22. As the same blade moves over the tail boom and into the advancing position, the lift distribution changes so that the lift is now generated more inboard. If the blade is twisted, the angle of attack at the tip may be negative at high forward speeds resulting in a download. The situation is now exactly the reverse of the retreating case. Clearly the blade is being excited by time varying forces and as it is not rigid it will flex.

Centrifugal stiffening keeps the blades taut giving them certain of the characteristics of the string of a musical instrument, except that the tension is not constant but is a function of radius. The response to excitation by a Fourier series is shown in Figure 7.2-32. A variety of harmonic structures will occur. At low frequencies the blade motion can be considered a form of flapping. At some higher frequency it will have to be considered a vibration. Note that blades carried on flapping bearings (a) will have a different harmonic response to that of hingeless blades. Teetering rotors and hingeless rotors have a different harmonic structure (b) from articulated rotors because the blades are fixed together in the centre of the rotor.

When a blade flexes, the CM of the blade elements must be closer to the rotor axis than when the blade is straight. Thus the blade mass is being hauled in and out against the rotational forces, modulating the blade tension. As the root tension is typically measured in tons, it should be clear that these lateral forces are considerable.

In the presence of harmonic flapping, conservation of blade momentum suggests that there will also be harmonic dragging. The azimuthal variation in lift caused by the application of the cyclic control will result in variations in induced drag that excite the blade in the dragging plane. This will be compounded by the azimuthal variation in profile drag. The blade will respond to these dragging excitations as a function of its dynamics and damping. When the blade drags, the effective moment-arm at which the blade outward pull is applied to the hub changes. Thus even if the blade pull were constant, the moment applied to the mast would be modulated. However, the blade pull is not constant because of blade flexing.

The application of cyclic pitch and the variations in airspeed seen by a blade segment will result in chord-wise movement of the centre of pressure which will excite the blade in torsion, placing alternating stresses on the pitch control mechanism.

A further complexity is that the finite number of blades in a real rotor cannot create a uniform downwash. The downwash velocity changes at the blade passing frequency. Figure 7.2-33 shows that in forward flight a given blade

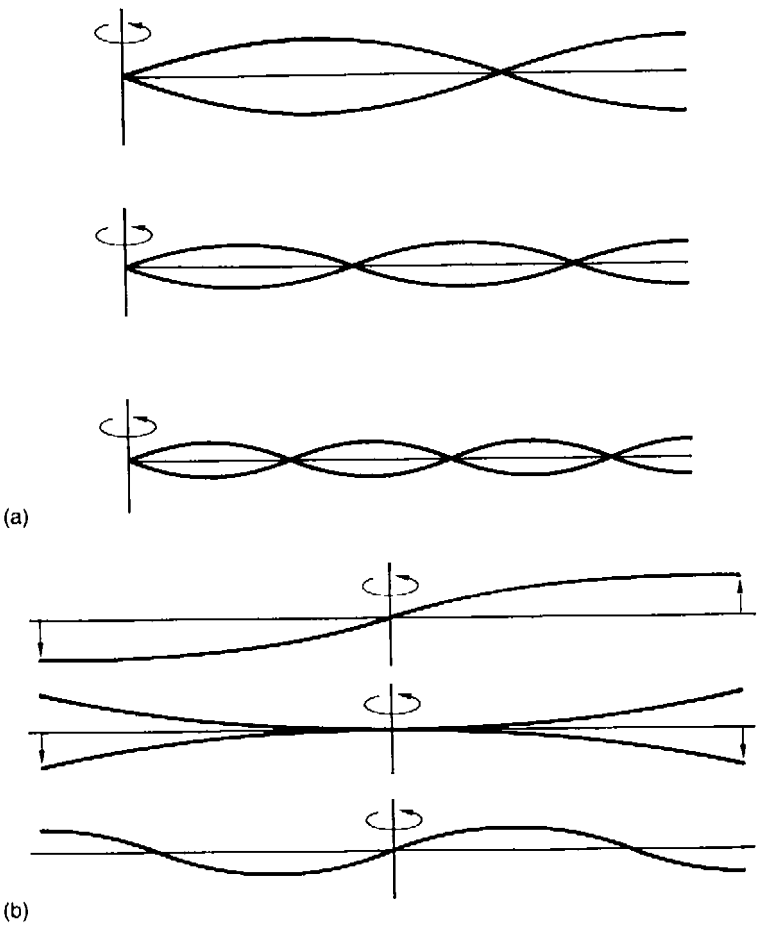


Figure 7.2-32 The blades are not rigid and the time variant lift distribution will excite them in various bending modes. (a) Modes for an articulated rotor in which the blades are hinged at the root. (b) Modes for a teetering or stiff hingeless rotor in which moments can pass from one blade to the other.

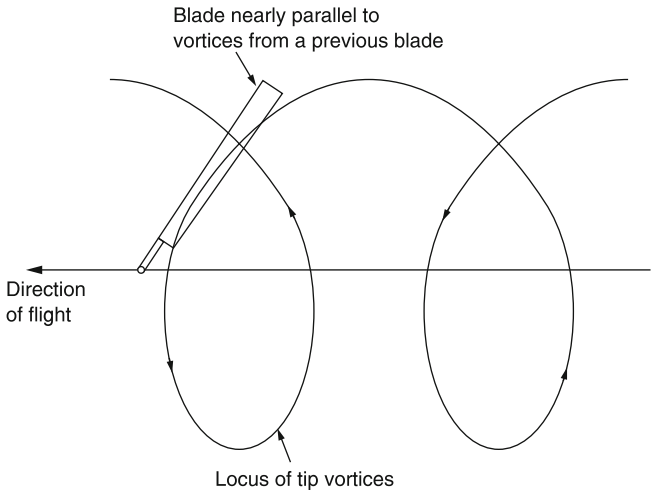


Figure 7.2-33 A rotor is not an ideal actuator disc. The downwash velocity is not constant but increases as each blade passes and falls in between. A fixed point on a blade will cross its own path and the paths of other blades in forward flight.

element traces out a path known as a cycloid. This path will take the blade element across its own path and those of a number of other blades depending on the forward speed and the number of blades. As the blade element crosses a previous path, the higher downwash velocity will reduce the angle of attack, whereas between the previous paths the angle of attack will increase. The result is that the lift of a particular blade element is modulated by the wakes of previous blades. Over most of the rotation the effect is generally not serious because the point on the blade at which the lift dips is constantly changing in radius. However, it can be seen from Figure 7.2-33 that at about 60° of azimuth the blade is nearly parallel to the cycloid, meaning that a significant length of the blade is experiencing the effect at the same time.

The non-uniformity of the downwash will also directly excite the hull and is commonly observed as flexing of the canopy at the blade passing frequency. In the absence of vibration structurally transmitted from the rotor, this mechanism will still vibrate the hull.

In addition to the vibration sources noted above, in some flight regimes the rotor may ingest its own tip vortices. These are mostly carried away by downwash in powered flight or upwash in autorotation, but in some conditions, such as the end of a landing approach or a roll manoeuvre, the vortex from one blade may hit the next blade head-on. Figure 7.2-34(a) shows that the effect of a vortex approaching a blade is a local modulation of the RAF direction causing a dip in the angle of attack. Figure 7.2-34(b) shows that as the vortex leaves the blade the angle of attack is

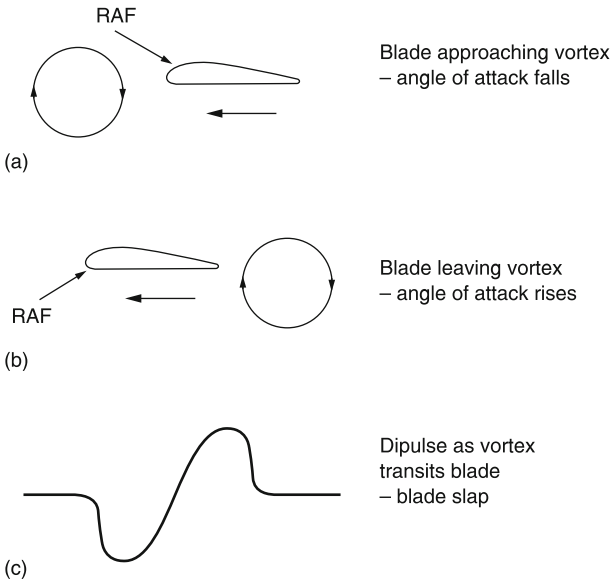


Figure 7.2-34 (a) A blade element encountering a vortex from an earlier blade will experience first an increase in angle of attack (a), then a reduction (b). The variations in angle of attack produce a pressure dipulse (c) rapid enough to be audible as blade slap.

increased. The time taken for the blade to transit the vortex is only a few milliseconds and so the resulting pressure dipulse shown at (c) produces audio frequencies. This is the origin of blade slap.

In order to simplify study and to allow comparison between rotors turning at different speeds, the frequencies of vibration are always divided by the rotor frequency. Thus the units of vibration are not cycles per second but cycles per revolution. An imbalanced rotor will cause a lateral vibration at the rotor frequency having a vibration frequency of one cycle per rev. This is usually abbreviated to 'one-per'.

In establishing the effect on the hull of vibrations created in the rotor, it is vital to appreciate that there are two completely different mechanisms at work. Vibrations in the direction of the mast axis and torsional vibrations about the mast axis are directly transmitted and will be felt in the hull at the same frequency. However, the rotor is turning with a frequency of its own and so vibrations in the plane of the rotor are taking place in a rotating frame of reference. When these vibrations reach the hull, which is in a stationary frame of reference, the same frequency will not be felt. Instead, in the case of lateral (in-plane) directions the vibration frequencies beat or heterodyne with the rotor frequency to produce sum and difference frequencies.

The concept of sidebands can be demonstrated graphically. It is possible conceptually to replace the blade generating in-plane vibrations due to modulation of its root pull with a device generating the same vibration. This consists of a pair of contra-rotating eccentrics called a Lanchester exciter shown in Figure 7.2-35(a) which produces linear sinusoidal vibration. Figure 7.2-35(b) shows that if the rotor blades vibrate radially at the second harmonic (two-per) this can be simulated by an exciter which makes two revolutions with respect to the blade in one rotor revolution. The result is that one of the eccentrics makes three revolutions with respect to the hull whilst the other

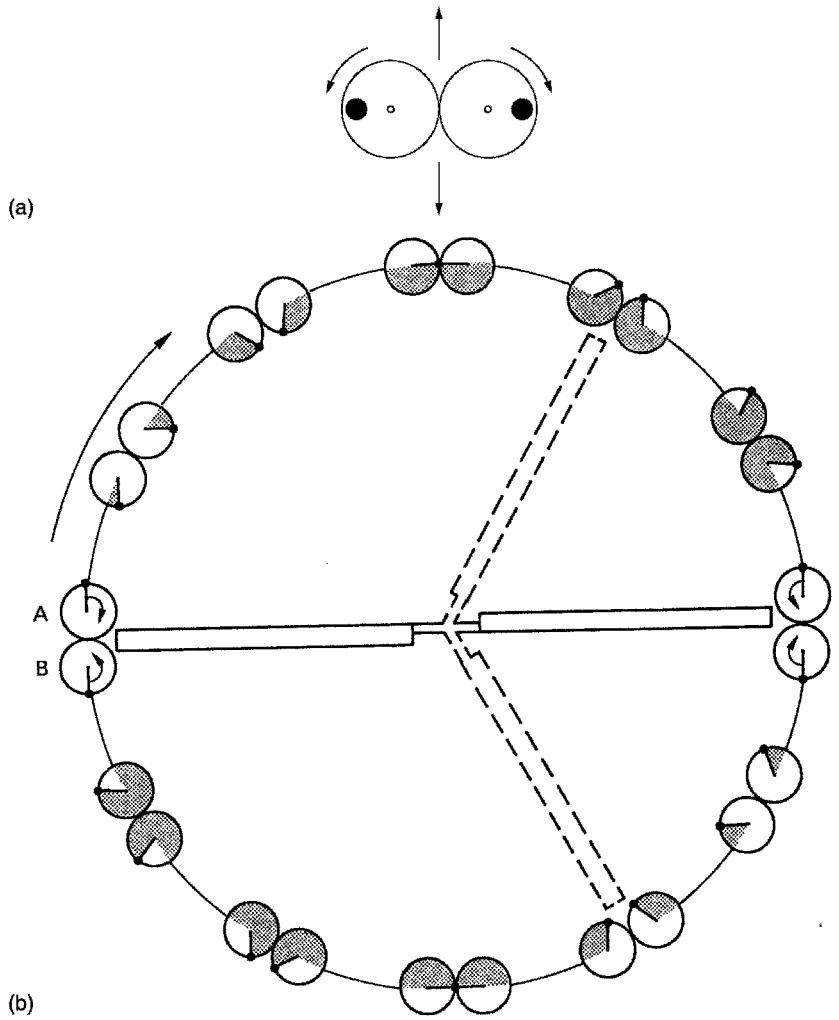


Figure 7.2-35 (a) A pair of contra-rotating eccentrics produces linear vibration. (b) If the mechanism of (a) is rotated, one of the eccentrics will turn faster and the other slower. Thus frequencies transmitted to the hull will be the vibration frequency plus or minus the rotor frequency.

makes only one. Thus a two-per in-plane blade vibration results in three-per and one-per at the hub in stationary co-ordinates. Third harmonic vibration produces four-per and two-per vibration in stationary co-ordinates and so on.

The discussion so far has considered the action of individual blades. However, in practice the hub will vectorially sum all of the forces and moments acting on it. Vectorial summation must consider the phase of the contributions from each blade. Clearly the phase difference between the blades is given by 360° divided by the number of blades. The overall result must then also depend on the number of blades in the rotor. If the blades are identical, and experience the same forces as they turn, they will develop the same vibration, except that the phase will be different for each blade. If, however, one blade has different characteristics to the others, a one-per function may result.

It is interesting to see how the number of blades affects the spectrum of vibration and how this effect is different for in-plane vibrations. Figure 7.2-36(a) depicts a two-bladed rotor and shows a graph of the fundamental flapping for each blade (one-per) and the second harmonic (two-per). Note that the two blades are 180° apart and so the one-per waveforms are in anti-phase and cancel. As one-per flapping is the result of tilting the disc this is no surprise and remains true for any number of blades. However, note that the second harmonic flapping of the two blades is in-phase and will therefore add.

It will be clear that a two-bladed rotor is at a disadvantage because both blades will be in the lift trough simultaneously at 90° and 270°, and in a lift peak simultaneously at 0° and 180°. The result at high speed is a vertical vibration known as 2P hop. Figure 7.2-36(b) shows the situation for the third flapping harmonic. As the third harmonic flapping has one and a half cycles in the time between blade passings, it will be clear that the 3P components will be out of phase between the two blades. Following similar logic it should be clear that in a two-bladed rotor, only even vertical harmonics exist; the odd ones cancel.

Figure 7.2-37 shows the case for a three-bladed rotor. At (a) it will be seen that the second flapping harmonic waveform of each blade is at 120° to that of the next. The sum of three sine waves at 120° is constant and so there can be no 2P hop in a three-bladed rotor, or indeed for a rotor with a higher number of blades. In Figure 7.2-37(b) is shown the effect for the third flapping harmonic. Here the third harmonic waveform is in the same phase for each blade and so there will be summation.

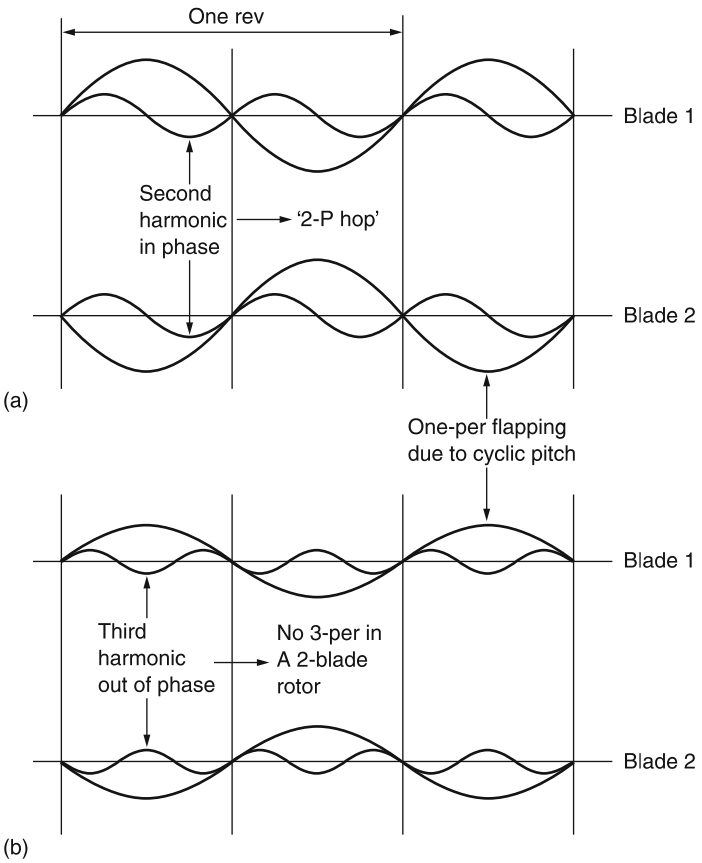


Figure 7.2-36 (a) In a two-bladed rotor the blades flap with a 180° relationship. The fundamental flapping cancels, whereas the second harmonic flapping adds. Two bladed rotors are prone to 2P hop. (b) The third harmonic cancels in a two-bladed rotor.

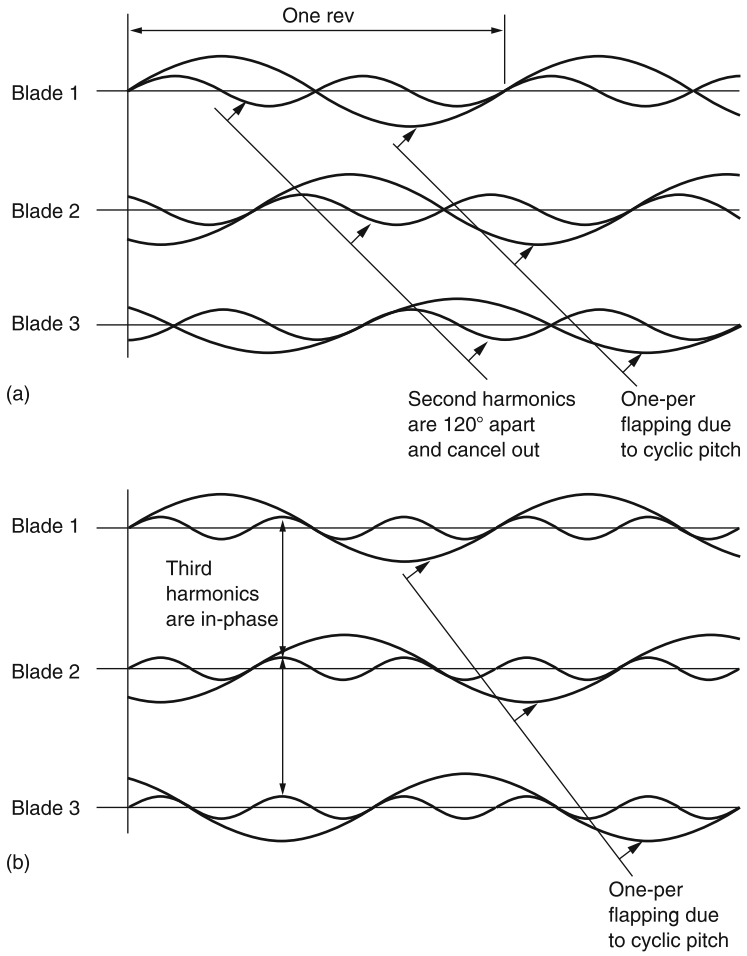


Figure 7.2-37 (a) In a three-bladed rotor, the second harmonic flapping is mutually at 120°, the sum of which is zero. (b) Third harmonic flapping is in the same phase in each blade and so will add.

Now that the general principle is clear, the result for any number of blades and any harmonic can readily be established. This is summarized in Figure 7.2-38(a) where it will be seen that the lowest hopping frequency is numerically equal to the number of blades and that these frequencies may be multiplied by integers to find the higher hopping modes.

When in-plane forces are considered, the results are quite different. Figure 7.2-38(b) shows the result of various harmonics with different numbers of blades. These forces will be heterodyned by rotor speed so that the result on the hull of a rotor frequency nP will be vibration at $(n–1)P$ and $(n+1)P$ whereas the frequency P is absent. Following the principles of vector summation explained above, certain harmonics cancel at the hub. Note that as with hop forces, the lowest frequencies at which rocking forces can exist are numerically the same as the number of blades, with the harmonics obtained by integer multiplication. The difference between Figure 7.2-38(a) and Figure 7.2-38(b) is that different frequencies at the blades are responsible for the same frequencies at the hull. For example, a five-bladed rotor can suffer 5P hopping and rocking, but the hopping is due to blade flapping at 5P whereas the rocking is due to blade flapping at 4P.

This section has concentrated on the main rotor, but clearly a tail rotor will show many of the same effects as it too is passing edge-on through the air. However, the tail rotor is at a further disadvantage because it is working in airflow that has been disturbed by the main rotor and the hull.

7.2.27 Vibration Control

Vibration and helicopters are inseparable. However well balanced the rotor may be, the lift asymmetry in forward flight will cause vibration. Vibration is unwelcome because it has a number of negative effects. Section 7.2.26 showed some of the many causes of vibration and it should be clear that many of these may result in severe alternating stresses in the rotor head, the blade roots, the pitch control mechanism and the transmission. This has a large impact on the service life of such parts where a failure cannot be tolerated. Many parts need to have more generous cross-sections, increasing mass and reducing payload. Periodic checking for cracks adds to maintenance costs.

A key problem due to vibration is that it impairs the pilot's vision. The human visual system (HVS) obtains its acuity by averaging visual information over a considerable length of time. In the presence of vibration, the image will be unstable on the retina and will thus appear blurred. There have been cases where damage to rotating components has caused vibration severe enough that the pilot was unable to read the instruments at all. Vibration contributes to pilot stress and workload. One of the earliest approaches therefore was to reduce vibration at the pilot's seat.

Given the low load factor of a helicopter, the hull is not highly stressed and could be very lightly built were it not for fatigue inducing vibration. Increasingly helicopters incorporate electronic systems and sensors used either to control or navigate the machine or as part of its mission equipment. Electronic equipment does not take kindly to vibration, or if specially adapted, must be more expensive.

Vibrating masses are a common origin of sound, so it follows that a vibrating helicopter will be noisy both inside and out. External noise causes disturbance in civil operations and may compromise military missions. A significant part of the design process of the modern helicopter is taken up in minimizing noise and vibration. There are a number of approaches that will be combined in various ways. Study of the aerodynamic origins of vibration may allow a reduction at the source. Once the forcing frequencies are known, the design of the structure should ensure that resonant responses to those frequencies are minimized. This is known as *detuning*. Vibration may be isolated or *decoupled* in some way, *damped* using lossy materials to convert flexing into heat, or opposed or *absorbed* through various cancelling techniques, both passive and active.

The reduction of the vibration at source has to be the technically superior solution. The disadvantage of decoupling the vibration from the passengers is that although their comfort is improved, the vibration experienced by the parts that are not decoupled may actually increase. It is easy to see why. For a given exciting force, the amplitude of a vibration is inversely proportional to the mass being vibrated. In a conventional helicopter the dominant masses are those of the transmission, hull and payload. If the hull and payload are decoupled, the mass seen by the vibration source is now only that of the transmission and the amplitude of the vibration there must be greater and not necessarily beneficial.

An obvious way of reducing vibration at source is to use non-sinusoidal cyclic pitch control that allows the lift moment to be more nearly constant as the blade turns. The means to do this are non-trivial and are described in section 7.2.28. Another way of reducing vibration is to use more blades. This has several advantages. First, the load on each blade is reduced and with it the magnitude of the forcing function of each blade. Second, the more blades, the higher the frequency of the vibration will be and the lower will be the amplitude for a given mass. Figure 7.2-37 shows that a three-bladed rotor eliminates the 2P hop of the two-bladed rotor, whereas a four-bladed rotor reintroduces some 2P lateral vibration. It will be seen from Figure 7.2-38(b) that an even number of blades gives lower rocking frequencies, making an odd number preferable. Consequently the change from four to five blades gives a significant improvement. Large transport helicopters inevitably have a large and generally odd number of blades – seven or nine – simply to contain the large vibrations that must result from generating high rotor thrust. Using many blades will be expensive and result in a large rotor head having high drag.

It is common to design the blades so that their natural frequencies of vibration are not in the spectrum of the excitation. If any of the excitation frequencies coincide with the resonant frequencies there could be a significant response at that frequency. Any coincidence can be seen from an interference diagram as shown in Figure 7.2-39. This interference diagram is for the Westland Lynx and shows the operating speed (318 RPM) as a vertical line. The excitation frequency may be at any integer multiple of the rotor frequency. Figure 7.2-39 shows that, as rotor speed increases, the excitation harmonics (thin lines starting at the origin) fan out. The resonant frequencies of the blades also change with rotor frequency because of the effect of centrifugal stiffening. Figure 7.2-39 shows how flapping, ωf, and lagging, ωl, resonant frequencies may change with rotor speed.

The only coincidence between excitation and response is the close correspondence between the fundamental flapping resonant frequency and the rotor fundamental frequency. This is the characteristic that results in rotor precession. Note that as the Lynx has a hingeless rotor, the flapping resonance is above the rotor frequency and so the rotor phase lag will be less than 90°.

In all other cases, coincidence is avoided so that at normal RRPM no resonances are excited. This result was not an accident. During the design process the blade was designed with a tapering spar resulting in the stiffness and the mass per unit length being non-uniform so that the flapping resonant frequencies are not integer multiples of the fundamental. Where constructional techniques such as extrusion do not allow the blade or spar to taper, the blades can also be detuned by placing weights at critical points along the span so that the effective mass per unit length is no longer uniform. Note that the technique of detuning only works at one rotor speed, at other speeds the excitation frequency is different and the response will also differ. Detuning will also be employed in the hull so that hull resonance does not magnify rotor-induced vibrations.

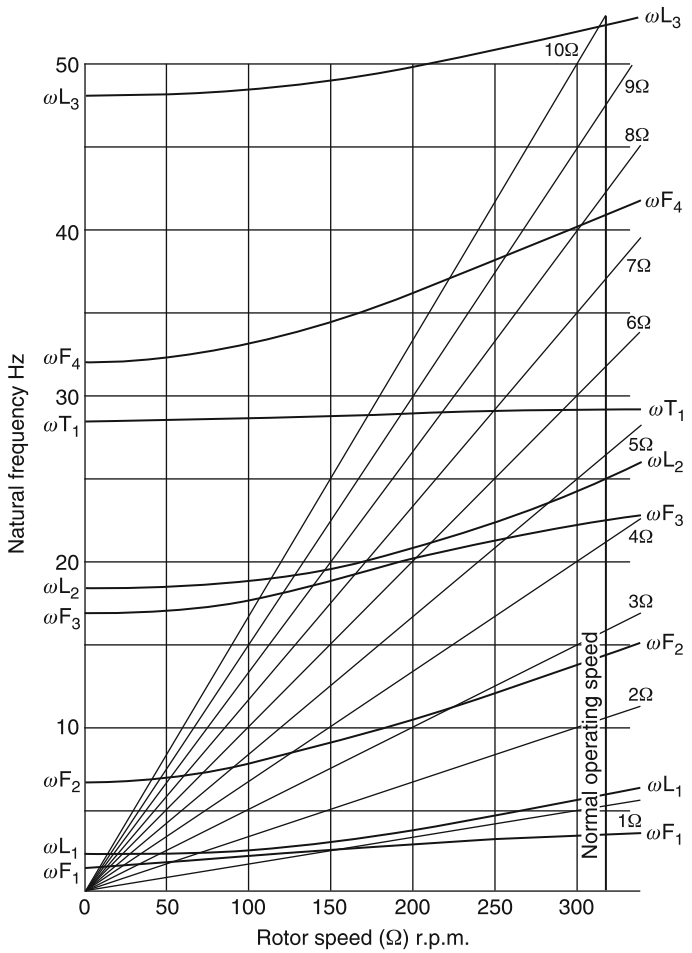


Figure 7.2-39 The excitation spectrum in a rotor fans out as RRPM increases. The natural frequencies of flapping and dragging of the blade will also change with RRPM due to centrifugal stiffening. If these two effects are superimposed, it is possible to see where the excitation frequencies coincide with response frequencies. This is known as an interference diagram. (AugustaWestland).

Although vibration may occur at a number of frequencies, often one of these will dominate and cause a problem. Vibration absorbers are devices intended to reduce vibration at a single specific frequency. The term is actually a misnomer because what they actually do is to create an opposing or near-anti-phase vibration. These devices work using mechanical resonance.

A mass supported on a spring displays resonance. If the damping is very low, the phase response just above the resonant frequency becomes nearly 180°. If the resonant frequency is set just below the frequency of vibration, the mass will automatically resonate in inverse phase and oppose the vibration. In helicopters the battery may be used as a suitable mass and this is supported on carefully calibrated springs. A mass absorber of this kind can only reduce vibration in the vicinity of the device itself. Flexibility of the hull will allow vibration elsewhere; however, such devices can oppose the hull vibration due to non-uniform downwash. Mass absorbers only work at a fixed frequency and require the rotor to turn at a precisely controlled RPM. The alternative is to use a system in which the spring stiffness can be modified by a servo so that the absorber can be tuned to the actual RRPM in use. Such a system is fitted in the Chinook. In some early helicopters vibrations fed into the control rods were opposed in this way. Figure 7.2-40(a) shows a resonator consisting of a bob weight on a leaf spring attached to a control bell crank.

A pendulum is a kind of resonator. The type used in a clock uses the earth's gravitational field to provide the restoring force. In helicopters, pendular vibration absorbers fitted on rotors use the acceleration caused by rotation. Figure 7.2-40(b) shows that pendular absorbers used to handle flapping vibration consist of weights that can swing on the end of cranked shafts pivoted in the blade roots. When the rotor is turning, the weights fly out horizontally and attempt to stay in the same plane. Vertical movement of the blade excites the pendulum which is tuned to resonate in anti-phase. These devices can be seen on the Kamov Ka-32 and on the Bo-105 and are known colloquially as bull's balls.

As these devices employ resonance, a single unit can only oppose vibration at a single frequency. However, it is possible to fit more than one pendulum to the same mounting. The Hughes 500 has a four-bladed rotor and so according to Figure 7.2-38(b) will be prone to 4P rocking forces due to flapping at 3P and 5P. The two pendula are tuned to the third and fifth harmonics of blade flapping.

Torsional vibration of the rotor shaft may be due to blade excitation in the dragging plane. This is an application for the bifilar pendulum. A pendulum can be made to swing in one axis only by attaching it with two strings, hence the name. Restoration forces are much smaller for in-plane motion, suggesting that a pendulum would need a very short arm in order to resonate at a harmonic of rotor frequency. Thus in practice the arms are replaced by a pair of pins which operate in oversized holes as in Figure 7.2-40(c). As the weight moves to one side, the geometry of the pins

moves the weight inwards. As the weight tends to fly out as the rotor turns, there will be a restoring force if the weight is displaced. This is all that is necessary to create a resonant system.

The restoring force that determines the resonant frequency in pendulum and bifilar absorbers comes from the acceleration due to rotation. This is a function of RRPM, as is the frequency to be absorbed. Consequently these devices are self-tuning.

Another popular and obvious way of isolating vibration is to use flexible mountings between the transmission and the hull. It is a fundamental concept of such mounts that there will be a resonance whose frequency is determined by the sprung mass and the stiffness of the mounts. The degree of isolation is proportional to the difference between that resonant frequency and the frequency to be filtered out. Consequently the lowest possible resonant frequency is required. This frequency may be lowered by softening the mounts, or by increasing the supported mass. Soft mounts may cause other problems, such as handling difficulties because the hull would tend not to follow the rotor in a manoeuvre. As a result, the best solution is to increase the mass to be isolated as much as possible.

This is the principle of the machinery raft, which is a rigid structure carrying the engines, transmission, control actuators, oil coolers, generators, etc. as a single unit. This raft is then mounted to the hull by widely spaced resilient supports. The large mass of the raft opposes the forcing function of the rotor and makes the resonant frequency as low as possible, as well as reducing the amplitude of vibration seen by the components on the raft.

The problem with resilient supports is that for good isolation they have to be soft, and this will always result in a static deflection. Some techniques have been developed for helicopters to circumvent the problem. Figure 7.2-41(a) shows a system developed by Kaman known as DAVI, or dynamic anti-resonant vibration isolator. This works using a mechanical impedance convertor to raise the moving mass artificially. Between the transmission and the hull is a lever carrying a bob weight. The mechanical advantage of the lever causes any relative motion to be amplified. The lever changes the mechanical impedance of the bob weight by a factor equal to the square of the mechanical advantage. Thus a 1 kg weight on a 10:1 lever would appear to weigh 100 kg. Mounting the transmission on, for example, four such mounts would raise the effective weight by 400 kg. Given the increase in apparent weight of the transmission, the supporting spring can be made much stiffer than would be possible with a conventional resilient mount, without increasing the resonant frequency.

Figure 7.2-41(b) shows the nodal beam system developed by Bell. At one time most Bell helicopters had two blades and suffered 2P hop. The nodal beam was developed to isolate the hop. Between the hull and the transmission is a flexible beam made from a sandwich of elastomer and steel. The transmission mounts are closer together on the beam than the hull mounts. As relative movement takes place between the hull and the transmission, the beam must bend. The proximity of the hull and transmission mounts makes a virtual lever system resulting in larger amplitude of

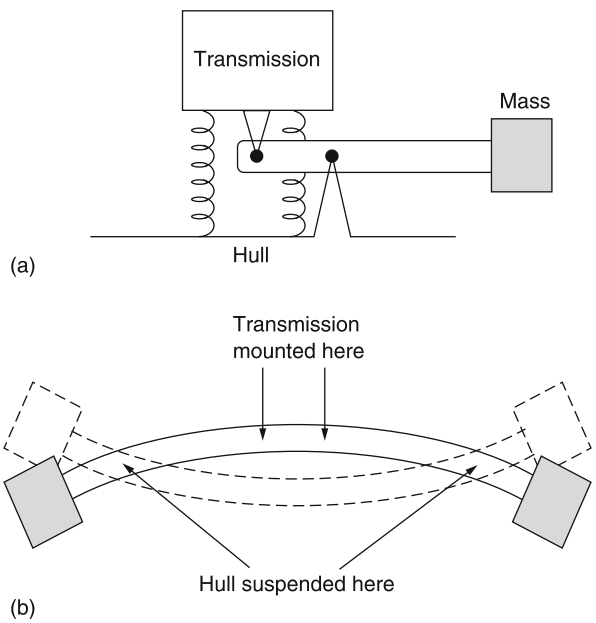


Figure 7.2-41 (a) In a dynamic anti-vibration isolator, a lever is used to amplify the motion of a weight thus giving it enormously increased effective mass. (b) The nodal beam system. A suitably tuned flexing beam will have nodes where the hull is attached.

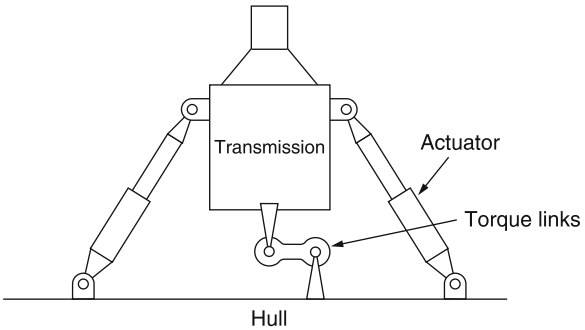


Figure 7.2-42 In an active anti-vibration system, the transmission is mounted on actuators that are driven by a suitable forcing function to cancel the vibration.

movement at the centre of the beam. The effective mass of the beam is amplified by the mechanical advantage of the virtual levers, making the system mathematically equivalent to a pair of DAVI isolators in parallel. The geometry of the system is tailored so that at the bending resonant frequency of the beam the hull attachment points are at nodes and so do not move.

In active vibration cancelling systems, Figure 7.2-42, a number of high speed actuators, typically hydraulically operated, are used to attach the hull to the transmission. The actuators are controlled by signals from processors and accelerometers in the hull. The signal processors are needed to adapt the waveforms fed to the actuators to changing conditions such as all up weight and airspeed. An increase in all up weight would require increased static pressure in the actuators to prevent them extending. In practice a feedback mechanism measures the average extension of the actuator and centralizes each one with a relatively long time constant.

No practical hull can be rigid, and so there will be various modes of vibration excited in it. One of the sources of vibration will be due to the non-uniform downwash impinging on the hull surface. Conventional isolators between hull and transmission can do nothing about this because they only oppose structure-borne vibration. However, an active system can use the transmission as a reaction mass to create forces opposing vibration from any source. Unfortunately the amplitude of vibration in the transmission will be increased. Better results may be obtained by applying active control to a raft carrying the engines, etc. as described above.

7.2.28 Harmonic Pitch Control

It was seen in Figure 7.2-30 that the use of sinusoidal feathering in a cyclic pitch control is incorrect. The airspeed seen by a blade element changes sinusoidally in forward flight, but the lift is proportional to the square of the airspeed. This non-linear function generates flapping harmonics that cannot be cancelled by sinusoidal feathering. It follows immediately that if there are to be no flapping harmonics, the feathering function must contain harmonics.

A harmonic cyclic pitch waveform is approximately a square root function of a raised sine wave although some modification will be necessary because of inflow effects. When multiplied by the sine-squared function due to the speed, the product, namely the lift, is constant. The pitch function above will need further refinement to account for upwash at the leading edge of the disc. Figure 7.2-43 shows how an actuator between the swashplate and the blade control arm could add the harmonic control.

In practice perfect constant lift can only be achieved at one radius, but it should be possible to find a pitch function that makes the average blade lift moment essentially constant. A more advanced system would have means to control the angle of attack of different parts of the blade independently. This may be possible using, for example, piezoelectric actuators. It is impossible to obtain constant lift over the whole blade because of the reverse flow area, and so some harmonic flapping will always remain.

Whilst theoretically obvious, harmonic blade control, also known as higher harmonic control, is rarely found. One of the difficulties is that the cyclic control becomes a function of airspeed, returning to a near sinusoidal function in the hover. This is probably beyond what might be achieved mechanically but some kind of electronic system would easily generate the required functions. High speed actuators are necessary to drive the blade feathering, and the blades must be torsionally stiff in order to respond to the rapid pitch variations. Another consideration is that as the number of blades increases the effects of lift troughs are averaged out more and there is less of a problem.

However, the elegant result of constant lift moment around the disc is one that is most desirable. The reduction in vibration makes the flight more comfortable for passengers and prolongs the life of the machine. A further

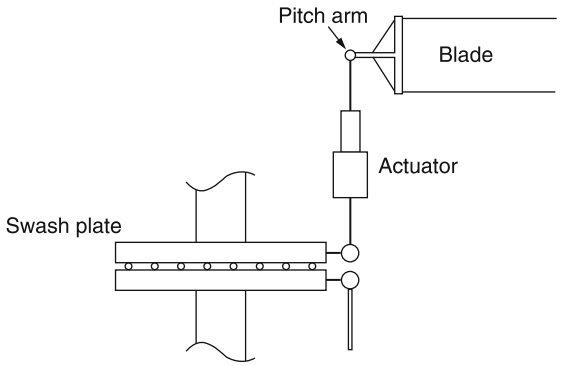


Figure 7.2-43 In harmonic pitch control, the blade pitch is not sinusoidal, but has a waveform calculated to make the lift more uniform around the disc. This may be controlled by an actuator between the swashplate and the blade pitch control arm.

consequence of constant lift is a reduction in power needed in forward flight. In a conventional rotor the increased power needed where the lift is greatest is not balanced by a reduction in power where the lift troughs are. Consequently constant lift needs the least power.

It would seem clear that the greatest cost benefit of harmonic blade control would be in the two-bladed helicopter where the 2P hop could be eliminated. The improvement in ride would be greatest but the number of actuators needed is least. In some light helicopters a two-bladed harmonically controlled rotor could replace a conventional three-bladed rotor and the overall cost might be the same. At some point designers will find ways of economically achieving the goal of constant lift. Some experiments have been done with systems which rock a conventional swashplate, but this cannot give the required result as each blade needs to be individually controlled.

7.2.29 Blade Design

The helicopter blade is a complex subject because of the number of interacting variables involved. These will be considered here. It might be thought that the aerodynamic characteristics of the blade would be determined only by shape. In general this is untrue. Figure 7.2-44 shows some of the issues. The long thin structure of a practical blade can never be rigid and flight loads will cause the blade to flex. However, it is not enough to know the stiffness of the blade to calculate the result. In forward flight the loads have a powerful alternating component and this generates harmonics. At the frequencies of the higher harmonics the motion may be mass controlled. A further complexity is that the flexing of the blade will affect the aerodynamics; a phenomenon called *aeroelasticity*. As it is clear that the aerodynamics affect the flexing, then there is a loop (Figure 7.2-44). This loop may be stable or unstable. Part of the design process must be to ensure that aeroelastic effects always result in blade stability.

The shape of the blade is completely defined by the root cut-out, the tip shape, the degree of twist and taper and by the blade section used, which may change with radius. Assuming first a rigid blade, the blade shape will allow the time variant aerodynamic forces to be predicted in a given flight regime. The way the blade responds to this

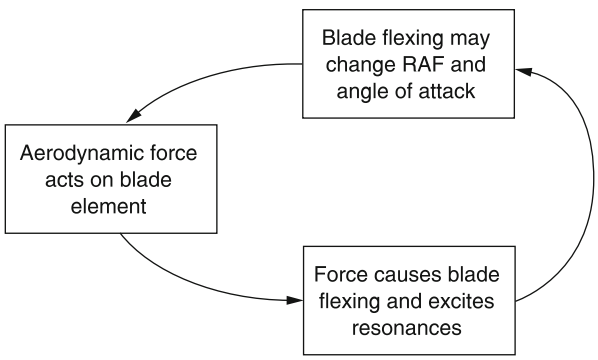


Figure 7.2-44 The aeroelasticity effect loop in a rotor blade. The blades are not rigid and can flex. This will affect their local angle of attack, changing the lift distribution that in turn causes flexing.

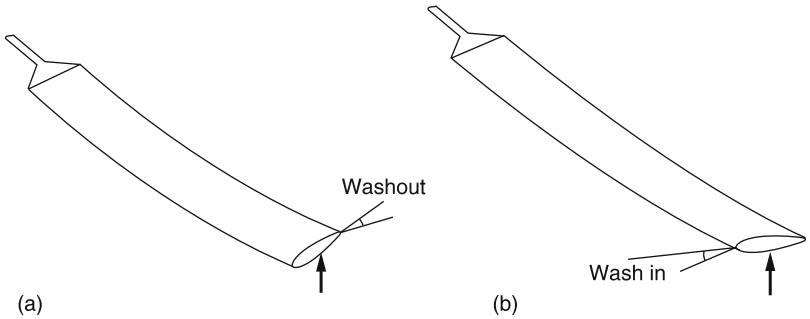


Figure 7.2-45 (a) If the blade has greater bending stiffness at the leading edge, it will wash out as lift increases. This is a stable condition as wash-out reduces the lift. (b) If the trailing edge is stiffer than the leading edge, lift causes wash-in that will further increase the lift. Such a blade would suffer serious flutter and probably disintegrate.

excitation is complicated. The torsional forces depend on the locus of the centre of pressure with respect to the mass centroid. The effect of flight loads on blade twisting is significant because twist will alter the local angle of attack as well as feeding loads back into the control system. The degree of twisting depends on the moment of inertia, the stiffness, the damping and the frequency.

The chord-wise distribution of stiffness needs to be known accurately. Figure 7.2-45(a) shows a blade in which the leading edge is stiffer than the trailing edge. The application of lift at the centre of pressure causes the blade to wash out; a stable condition. However, Figure 7.2-45(b) shows a blade in which the trailing edge is stiffer than the leading edge. Here lift causes the blade to wash in; leading to flutter. As a result blades are generally constructed in such a way that the mass centroid and the centre of bending stiffness is well forward on the chord. The result is a section where the structural mass is concentrated in a D-shaped structure at the leading edge, with the trailing part of the section being comprised of a low density structure.

Blade taper is not always employed because in practice the penalty of a constant chord blade is only a few per cent. Taper prevents constructional techniques such as extrusion that requires a constant section. Blade twist is advantageous in hover, but the ideal amount for hover will be detrimental in translational flight because the application of cyclic pitch may cause the finely pitched blade tips to go to a negative angle of attack. The exception to this restriction is in the tilt rotor aircraft that both hovers and cruises with axial inflow. This gives the designer more freedom to use blade twist and this is clearly visible in rotors of the Bell-Boeing Osprey.

This page intentionally left blank

SECTION 8

Space Vehicles and Rockets

This page intentionally left blank

Aerospace Engineering Desk Reference; ISBN 9781856175753

8.1 Satellite Attitude Dynamics

Howard D. Curtis

8.1.1 Introduction

In this chapter we apply the equations of rigid body motion to the study of the attitude dynamics of satellites. We begin with spin-stabilized spacecraft. Spinning a satellite around its axis is a very simple way to keep the vehicle pointed in a desired direction. We investigate the stability of a spinning satellite to show that only oblate spinners are stable over long times. Overcoming this restriction on the shape of spin-stabilized spacecraft led to the development of dual-spin vehicles, which consist of two interconnected segments rotating at different rates about a common axis. We consider the stability of that type of configuration as well. The nutation damper and its effect on the stability of spin-stabilized spacecraft is covered next.

The rest of the chapter is devoted to some of the common means of changing the attitude or motion of a spacecraft by applying external or internal forces or torques. The coning maneuver changes the attitude of a spinning spacecraft by using thrusters to apply impulsive torque, which alters the angular momentum and hence the orientation of the spacecraft. The much-used yo-yo despin maneuver reduces or eliminates the spin rate by releasing small masses attached to cords initially wrapped around the cylindrical vehicle.

An alternative to spin stabilization is three-axis stabilization by gyroscopic attitude control. In this case, the vehicle does not continuously rotate. Instead, the desired attitude is maintained by the spin of small wheels within the spacecraft. These are called reaction wheels or momentum wheels. If allowed to pivot relative to the vehicle, they are known as control moment gyros. The attitude of the vehicle can be changed by varying the speed or orientation of these internal gyros. Small thrusters may also be used to supplement the gyroscopic attitude control and to hold the spacecraft orientation fixed when it is necessary to despin or reorient gyros that have become saturated (reached their maximum spin rate or deflection) over time.

The chapter concludes with a discussion of how the earth's gravitational field by itself can stabilize the attitude of large satellites such as the space shuttle or space station in low earth orbits.

8.1.2 Torque-Free Motion

Gravity is the only force acting on a satellite coasting in orbit (if we neglect secondary drag forces and the gravitational influence of bodies other than the planet being orbited). Unless the satellite is unusually large, the gravitational force is concentrated at the center of mass G. Since the net moment about the center of mass is zero, the satellite is 'torque-free', and

$$\dot{\mathbf{H}}_G = \mathbf{0} \tag{8.1.1}$$

The angular momentum $\mathbf{H}_G$ about the center of mass does not depend on time. It is a vector fixed in inertial space. We will use $\mathbf{H}_G$ to define the Z axis of an inertial frame, as shown in Figure 8.1-1. The xyz axes in the figure comprise the principal body frame, centered at G. The angle between the z axis and $\mathbf{H}_G$ is (by definition of the Euler angles) the nutation angle θ. Let us determine the conditions for which θ is constant. From the dot product operation we know that

$$\cos\theta = \frac{\mathbf{H}_G}{\|\mathbf{H}_G\|} \cdot \hat{\mathbf{k}}$$

Differentiating this expression with respect to time, keeping in mind Equation 8.1.1, we get

$$\frac{d\cos\theta}{dt} = \frac{\mathbf{H}_G}{\|\mathbf{H}_G\|} \cdot \frac{d\hat{\mathbf{k}}}{dt}$$

But $d\hat{\mathbf{k}}/dt = \boldsymbol{\omega} \times \hat{\mathbf{k}}$, so

$$\frac{d\cos\theta}{dt} = \frac{\mathbf{H}_G \cdot (\boldsymbol{\omega} \times \hat{\mathbf{k}})}{\|\mathbf{H}_G\|} \tag{8.1.2}$$

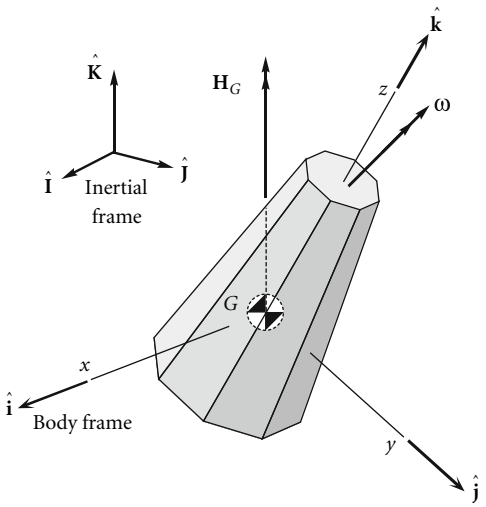


Figure 8.1-1 Rotationally symmetric satellite in torque-free motion.

Now

$$\boldsymbol{\omega} \times \hat{\mathbf{k}} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \omega_x & \omega_y & \omega_z \\ 0 & 0 & 1 \end{vmatrix} = \omega_y \hat{\mathbf{i}} - \omega_x \hat{\mathbf{j}}$$

Furthermore, we know that the angular momentum is related to the angular velocity in the principal body frame by the expression

$$\mathbf{H}_G = A\omega_x \hat{\mathbf{i}} + B\omega_y \hat{\mathbf{j}} + C\omega_z \hat{\mathbf{k}}$$

Thus

$$\mathbf{H}_G \cdot (\boldsymbol{\omega} \times \hat{\mathbf{k}}) = (A\omega_x \hat{\mathbf{i}} + B\omega_y \hat{\mathbf{j}} + C\omega_z \hat{\mathbf{k}}) \cdot (\omega_y \hat{\mathbf{i}} - \omega_x \hat{\mathbf{j}}) = (A - B)\omega_x \omega_y$$

so that Equation 8.1.2 can be written as

$$\dot{\theta} = \omega_n = -\frac{(A - B)\omega_x \omega_y}{\|\mathbf{H}_G\| \sin\theta} \tag{8.1.3}$$

From this we see that the nutation rate vanishes only if $A = B$. If $A \neq B$, the nutation angle θ will not in general be constant.

Relative to the body frame, Equation 8.1.1 is written as

$$\dot{\mathbf{H}}_{G\text{rel}} + \boldsymbol{\omega} \times \mathbf{H}_G = \mathbf{0}$$

This is Euler's equation with $\mathbf{M}_{G\text{net}} = \mathbf{0}$, the components of which are

$$\begin{aligned} A\dot{\omega}_x + (C - B)\omega_z \omega_y &= 0 \\ B\dot{\omega}_y + (A - C)\omega_x \omega_z &= 0 \\ C\dot{\omega}_z + (B - A)\omega_y \omega_x &= 0 \end{aligned} \tag{8.1.4}$$

In the interest of simplicity, let us consider the special case illustrated in Figure 8.1-1, namely that in which the z axis is an axis of rotational symmetry, so that $A = B$. Then Equations 8.1.4 become

$$\begin{aligned} A\dot{\omega}_x + (C - A)\omega_z \omega_y &= 0 \\ A\dot{\omega}_y + (A - C)\omega_x \omega_z &= 0 \\ C\dot{\omega}_z &= 0 \end{aligned} \tag{8.1.5}$$

From Equation $8.1.5_3$ we see that

$$\omega_z = \omega_0 \quad \text{(constant)} \tag{8.1.6}$$

The assumption of rotational symmetry therefore reduces the three differential equations 8.1.4 to just two. Substituting Equation 8.1.6 into Equations $8.1.5_1$ and $8.1.5_2$ and introducing the notation

$$\lambda = \frac{A - C}{A}\omega_0 \tag{8.1.7}$$

they can be written as

$$\begin{aligned} \dot{\omega}_x - \lambda\omega_y &= 0 \\ \dot{\omega}_y + \lambda\omega_x &= 0 \end{aligned} \tag{8.1.8}$$

To reduce these two equations in ω_x and ω_y down to just one equation in ω_x, we first differentiate Equation $8.1.8_1$ with respect to time to get

$$\ddot{\omega}_x - \lambda\dot{\omega}_y = 0 \tag{8.1.9}$$

We then solve Equation $8.1.8_2$ for $\dot{\omega}_y$ and substitute the result into Equation 8.1.9, which leads to

$$\ddot{\omega}_x + \lambda^2\omega_x = 0 \tag{8.1.10}$$

The solution of this well-known differential equation is

$$\omega_x = \Omega \sin \lambda t \tag{8.1.11}$$

where the constant amplitude Ω ($\Omega \neq 0$) has yet to be determined. (Without loss of generality, we have set the phase angle, the other constant of integration, equal to zero.) Substituting Equation 8.1.11 back into Equation $8.1.8_1$ yields the solution for ω_y,

$$\omega_y = \frac{1}{\lambda}\frac{d\omega_x}{dt} = \frac{1}{\lambda}\frac{d}{dt}(\Omega \sin \lambda t)$$

or

$$\omega_y = \Omega \cos \lambda t \tag{8.1.12}$$

Equations 8.1.6, 8.1.11 and 8.1.12 give the components of the absolute angular velocity $\boldsymbol{\omega}$ along the three principal body axes,

$$\boldsymbol{\omega} = \Omega \sin \lambda t\,\hat{\mathbf{i}} + \Omega \cos \lambda t\,\hat{\mathbf{j}} + \omega_0\hat{\mathbf{k}}$$

or

$$\boldsymbol{\omega} = \boldsymbol{\omega}_\perp + \omega_0\hat{\mathbf{k}} \tag{8.1.13}$$

where

$$\boldsymbol{\omega}_\perp = \Omega(\sin \lambda t\,\hat{\mathbf{i}} + \cos \lambda t\,\hat{\mathbf{j}}) \tag{8.1.14}$$

$\boldsymbol{\omega}_\perp$ ('omega-perp') is the component of $\boldsymbol{\omega}$ normal to the z axis. It sweeps out a circle of radius Ω in the xy plane at an angular velocity λ. Thus, $\boldsymbol{\omega}$ sweeps out a cone, as illustrated in Figure 8.1-2.

The three Euler orientation angles (and their rates) are related to the angular velocity components ω_x, ω_y and ω_z by

$$\begin{aligned} \omega_p &= \dot{\phi} = \frac{1}{\sin\theta}(\omega_x \sin\psi + \omega_y \cos\psi) \\ \omega_n &= \dot{\theta} = \omega_x \cos\psi - \omega_y \sin\psi \\ \omega_s &= \dot{\psi} = -\frac{1}{\tan\theta}(\omega_x \sin\psi + \omega_y \cos\psi) + \omega_z \end{aligned}$$

Substituting Equations 8.1.6, 8.1.11 and 8.1.12 into these three equations yields

$$\begin{aligned} \omega_p &= \frac{\Omega}{\sin\theta}\cos(\lambda t - \psi) \\ \omega_n &= \Omega \sin(\lambda t - \psi) \\ \omega_s &= \omega_0 - \frac{\Omega}{\tan\theta}\cos(\lambda t - \psi) \end{aligned} \tag{8.1.15}$$

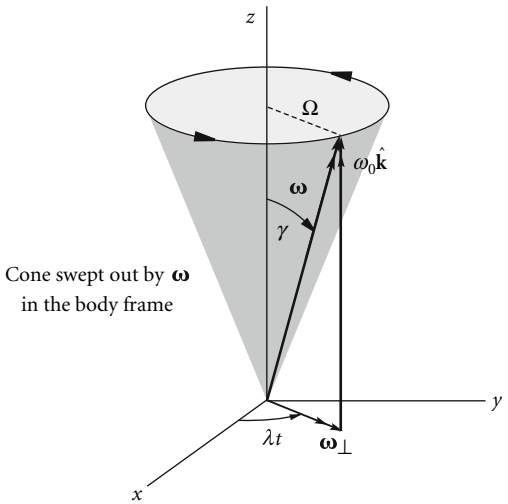


Figure 8.1-2 Components of the angular velocity * in the body frame.

Since $A = B$, we know from Equation 8.1.3 that $\omega_n = 0$. It follows from Equation $8.1.15_2$ that

$$\psi = \lambda t \tag{8.1.16}$$

(Actually, $\lambda t - \psi = n\pi$, $n = 0, 1, 2, \ldots$. We can set $n = 0$ without loss of generality.) Substituting Equation 8.1.16 into Equations $8.1.15_1$ and $8.1.15_3$ yields

$$\omega_p = \frac{\Omega}{\sin\theta} \tag{8.1.17}$$

and

$$\omega_s = \omega_0 - \frac{\Omega}{\tan\theta} \tag{8.1.18}$$

We have thus obtained the Euler angle rates ω_p and ω_s in terms of the components of the angular velocity $\boldsymbol{\omega}$.

Differentiating Equation 8.1.16 with respect to time shows that

$$\lambda = \dot{\psi} = \omega_s \tag{8.1.19}$$

That is, the rate λ at which $\boldsymbol{\omega}$ rotates around the body z axis equals the spin rate. Substituting the spin rate for λ in Equation 8.1.7 shows that ω_s is related to ω_0 alone,

$$\omega_s = \frac{A - C}{A}\omega_0 \tag{8.1.20}$$

Eliminating ω_s from Equations 8.1.18 and 8.1.20 yields the relationship between the magnitudes of the orthogonal components of the angular velocity in Equation 8.1.13,

$$\Omega = \frac{C}{A}\omega_0 \tan\theta \tag{8.1.21}$$

A similar relationship exists between ω_p and ω_s, which generally are *not* orthogonal. Substitute Equation 8.1.21 into Equation 8.1.17 to obtain

$$\omega_0 = \frac{A}{C}\omega_p \cos\theta \tag{8.1.22}$$

Placing this in Equation 8.1.20 leaves an expression involving only ω_p and ω_s, from which we get a useful formula relating the precession of a torque-free body to its spin,

$$\omega_p = \frac{C}{A - C}\frac{\omega_s}{\cos\theta} \tag{8.1.23}$$

Observe that if $A > C$ (i.e., the body is *prolate*, like a soup can or an American football), then ω_p has the same sign as ω_s, which means the precession is *prograde*. For an *oblate* body (like a tuna fish can or a frisbee), $A < C$ and the precession is *retrograde*.

The components of angular momentum along the body frame axes are obtained from the body frame components of $\boldsymbol{\omega}$,

$$\mathbf{H}_G = A\omega_x\hat{\mathbf{i}} + A\omega_y\hat{\mathbf{j}} + C\omega_z\hat{\mathbf{k}}$$

or

$$\mathbf{H}_G = \mathbf{H}_\perp + C\omega_0\hat{\mathbf{k}} \tag{8.1.24}$$

where

$$\mathbf{H}_\perp = A\Omega(\sin\omega_s t\,\hat{\mathbf{i}} + \cos\omega_s t\,\hat{\mathbf{j}}) = A\boldsymbol{\omega}_\perp \tag{8.1.25}$$

Since $\omega_0\hat{\mathbf{k}}$ and $C\omega_0\hat{\mathbf{k}}$ are collinear, as are $\boldsymbol{\omega}_\perp$ and $A\boldsymbol{\omega}_\perp$, it follows that $\hat{\mathbf{k}}$, $\boldsymbol{\omega}$ and $\mathbf{H}_G$ all lie in the same plane. $\mathbf{H}_G$ and $\boldsymbol{\omega}$ both rotate around the z axis at the same rate ω_s. These details are illustrated in Figure 8.1-3. See how the precession and spin angular velocities, $\boldsymbol{\omega}_p$ and $\boldsymbol{\omega}_s$, add up vectorially to give $\boldsymbol{\omega}$. Note also that from the point of view of inertial space, where $\mathbf{H}_G$ is fixed, $\boldsymbol{\omega}$ and $\hat{\mathbf{k}}$ rotate around $\mathbf{H}_G$ with angular velocity ω_p.

Let γ be the angle between $\boldsymbol{\omega}$ and the spin axis z, as shown in Figures 8.1-2 and 8.1-3. γ is sometimes referred to as the wobble angle. Then

$$\cos\gamma = \frac{\omega_z}{\|\boldsymbol{\omega}\|} = \frac{\omega_0}{\sqrt{\Omega^2+\omega_0^2}} = \frac{\omega_0}{\sqrt{\left(\omega_0\dfrac{C}{A}\tan\theta\right)^2+\omega_0^2}} = \frac{A}{\sqrt{A^2+C^2\tan^2\theta}}$$

γ is constant, since A, C and θ are fixed. Using trig identities, this expression can be recast as

$$\cos\gamma = \frac{\cos\theta}{\sqrt{\dfrac{C^2}{A^2}+\left(1-\dfrac{C^2}{A^2}\right)\cos^2\theta}} \tag{8.1.26}$$

From this we conclude that if $A > C$, then $\gamma < \theta$, whereas $C > A$ means $\gamma > \theta$. That is, the angular velocity vector $\boldsymbol{\omega}$ lies between the z axis and the angular momentum vector $\mathbf{H}_G$ when A > C (prolate body). On the other hand, when $C > A$ (oblate body), $\mathbf{H}_G$ lies between the z axis and $\boldsymbol{\omega}$. These two situations are illustrated in Figure 8.1-4, which also shows the *body cone* and *space cone*. The space cone is swept out in inertial space by the angular velocity vector as it rotates with angular velocity ω_p around $\mathbf{H}_G$, whereas the body cone is the trace of $\boldsymbol{\omega}$ in the body frame as it rotates with angular velocity ω_s about the z axis. From inertial space, the motion may be visualized as the body cone rolling on the space cone, with the line of contact being the angular velocity vector. From the body frame it appears as though the space cone rolls on the body cone. Figure 8.1-4 graphically confirms our deduction from Equation 8.1.23, namely, that precession and spin are in the same direction for prolate bodies and opposite in direction for oblate shapes.

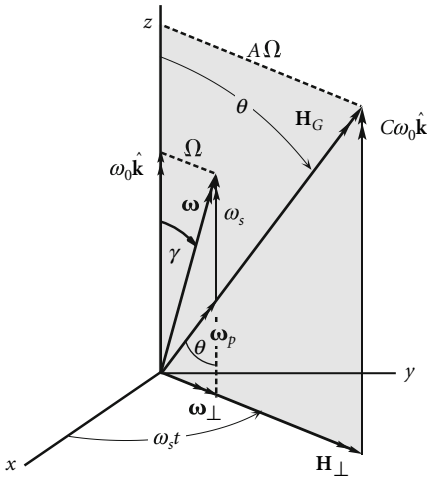
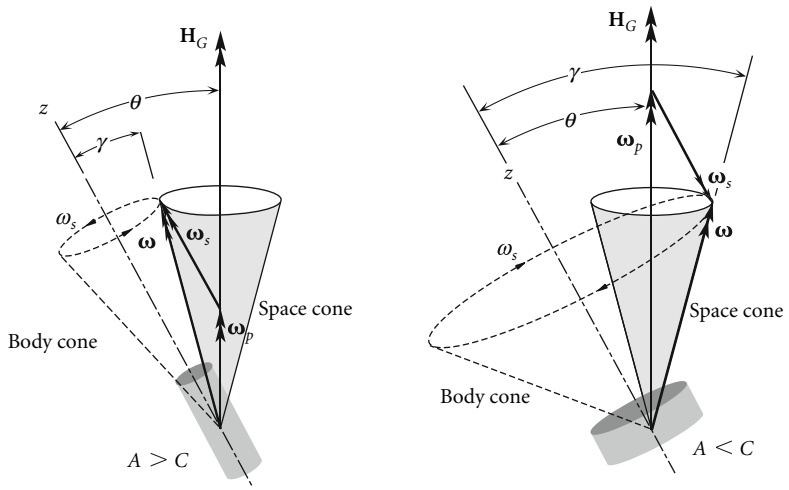


Figure 8.1-3 Angular velocity and angular momentum vectors in the body frame.



(a) Prograde precession (b) Retrograde precession

Figure 8.1-4 Space and body cones for a rotationally symmetric body in torque-free motion. (a) Prolate body. (b) Oblate body.

Finally, we know from Equations 8.1.24 and 8.1.25 that the magnitude $\|\mathbf{H}_G\|$ of the angular momentum is

$$\|\mathbf{H}_G\| = \sqrt{A^2\Omega^2 + C^2\omega_0^2}$$

Using Equation 8.1.21, we can write this as

$$\|\mathbf{H}_G\| = \sqrt{A^2\left(\omega_0\frac{C}{A}\tan\theta\right)^2 + C^2\omega_0^2} = C\omega_0\sqrt{1+\tan^2\theta} = \frac{C\omega_0}{\cos\theta}$$

Substituting Equation 8.1.22 into this expression yields a surprisingly simple formula for the magnitude of the angular momentum,

$$\|\mathbf{H}_G\| = A\omega_p \tag{8.1.27}$$

Example 8.1.1

A cylindrical shell is rotating in torque-free motion about its longitudinal axis (Figure 8.1-5). If the axis is wobbling slightly, determine the ratios of l/r for which the precession will be prograde or retrograde.

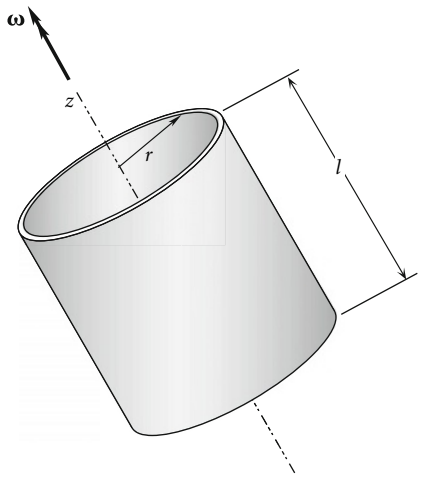


Figure 8.1-5 Cylindrical shell in torque-free motion.

For a thin-walled circular cylinder,

$$C = mr^2 \quad A = \frac{1}{2}mr^2 + \frac{1}{12}ml^2$$

According to Equation 8.1.23 and Figure 8.1-4, direct or prograde precession exists if $A > C$, that is, if

$$\frac{1}{2}mr^2 + \frac{1}{12}ml^2 > mr^2$$

or

$$\frac{1}{12}ml^2 > \frac{1}{2}mr^2$$

Thus

$l > 2.45r \quad \Rightarrow$ Direct precession.

$l < 2.45r \quad \Rightarrow$ Retrograde precession.

Example 8.1.2

In the previous example, let $r = 1$ m, $l = 3$ m, $m = 100$ kg and the nutation angle θ is 20°. How long does it take the cylinder to precess through 180° if the spin rate is 2π radians per minute?

Since $l > 2.45r$, the precession is direct. Furthermore,

$$C = mr^2 = 100 \cdot 1^2 = 100 \text{ kg}\cdot\text{m}^2$$

$$A = \frac{1}{2}mr^2 + \frac{1}{12}ml^2 = \frac{1}{2}\cdot 100 \cdot 1^2 + \frac{1}{12}100 \cdot 3^2 = 125 \text{ kg}\cdot\text{m}^2$$

Thus, Equation 8.1.23 yields

$$\omega_p = \frac{C}{A - C}\frac{\omega_s}{\cos\theta} = \frac{100}{125 - 100}\frac{2\pi}{\cos 20^\circ} = 26.75 \text{ rad/min}$$

At this rate, the time for a precession angle of 180° is

$$t = \frac{\pi}{\omega_p} = \underline{0.1175 \text{ min}}$$

Example 8.1.3

What is the torque-free motion of a satellite for which $A = B = C$?

If $A = B = C$, the satellite is spherically symmetric. Any orthogonal triad at G is a principal body frame, so $\mathbf{H}_G$ and $\boldsymbol{\omega}$ are collinear,

$$\mathbf{H}_G = C\boldsymbol{\omega}$$

Substituting this and $\mathbf{M}_{\text{Gnet}} = \mathbf{0}$, into Euler's equations, Equation 8.1.72a, yields

$$C\frac{d\boldsymbol{\omega}}{dt} + \boldsymbol{\omega} \times (C\boldsymbol{\omega}) = 0$$

That is,

$$\boldsymbol{\omega} = \text{constant}$$

The angular velocity vector of a spherically symmetric satellite is fixed in magnitude and direction.

Example 8.1.4

The inertial components of the angular momentum of a torque-free rigid body are

$$\mathbf{H}_G = 320\hat{\mathbf{I}} - 375\hat{\mathbf{J}} + 450\hat{\mathbf{K}}(\text{kg}\cdot\text{m}^2/\text{s}) \qquad \text{(a)}$$

The Euler angles are

$$\phi = 20^\circ \quad \theta = 50^\circ \quad \psi = 75^\circ \tag{b}$$

If the inertia tensor in the body-fixed principal frame is

$$[\mathbf{I}_G] = \begin{bmatrix} 1000 & 0 & 0 \\ 0 & 2000 & 0 \\ 0 & 0 & 3000 \end{bmatrix} (\text{kg}\cdot\text{m}^2) \tag{c}$$

calculate the inertial components of the (absolute) angular acceleration.
Using the Euler angles from (b), we obtain the matrix of the transformation from the inertial frame to the body frame,

$$[\mathbf{Q}]_{Xx} = \begin{bmatrix} 0.03086 & 0.6720 & 0.7399 \\ -0.9646 & -0.1740 & 0.1983 \\ 0.2620 & -0.7198 & 0.6428 \end{bmatrix} \tag{d}$$

We use this to obtain the components of $\mathbf{H}_G$ in the body frame,

$$\begin{aligned} \{\mathbf{H}_G\}_x = [\mathbf{Q}]_{Xx}\{\mathbf{H}_G\}_X &= \begin{bmatrix} 0.03086 & 0.6720 & 0.7399 \\ -0.9646 & -0.1740 & 0.1983 \\ 0.2620 & -0.7198 & 0.6428 \end{bmatrix} \begin{Bmatrix} 320 \\ -375 \\ 450 \end{Bmatrix} \\ &= \begin{Bmatrix} 90.86 \\ -154.2 \\ 643.0 \end{Bmatrix} (\text{kg}\cdot\text{m}^2/\text{s}) \end{aligned} \tag{e}$$

In the body frame $\{\mathbf{H}_G\}_x = [\mathbf{I}_G]\{\boldsymbol{\omega}\}_x$, where $\{\boldsymbol{\omega}\}_x$ are the components of angular velocity in the body frame. Thus

$$\begin{Bmatrix} 90.86 \\ -154.2 \\ 643.0 \end{Bmatrix} = \begin{bmatrix} 1000 & 0 & 0 \\ 0 & 2000 & 0 \\ 0 & 0 & 3000 \end{bmatrix} \{\boldsymbol{\omega}\}_x$$

or, solving for $\{\boldsymbol{\omega}\}_x$,

$$\{\boldsymbol{\omega}\}_x = \begin{bmatrix} 1000 & 0 & 0 \\ 0 & 2000 & 0 \\ 0 & 0 & 3000 \end{bmatrix}^{-1} \begin{Bmatrix} 90.86 \\ -154.2 \\ 643.0 \end{Bmatrix} = \begin{Bmatrix} 0.09086 \\ -0.07709 \\ 0.2144 \end{Bmatrix} (\text{rad/s}) \tag{f}$$

Euler's equations of motion may be written for the case at hand as

$$[\mathbf{I}_G]\{\boldsymbol{\alpha}\}_x + \{\boldsymbol{\omega}\}_x \times \left([\mathbf{I}_G]\{\boldsymbol{\omega}\}_x\right) = \{\mathbf{0}\} \tag{g}$$

where $\{\boldsymbol{\alpha}\}_x$ is the absolute acceleration in body frame components. Substituting (c) and (f) into this expression, we get

$$\begin{bmatrix} 1000 & 0 & 0 \\ 0 & 2000 & 0 \\ 0 & 0 & 3000 \end{bmatrix} \{\boldsymbol{\alpha}\}_x + \begin{Bmatrix} 0.09086 \\ -0.07709 \\ 0.2144 \end{Bmatrix}$$

$$\times \left(\begin{bmatrix} 1000 & 0 & 0 \\ 0 & 2000 & 0 \\ 0 & 0 & 3000 \end{bmatrix} \begin{Bmatrix} 0.09086 \\ -0.07709 \\ 0.2144 \end{Bmatrix} \right) = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

$$\begin{bmatrix} 1000 & 0 & 0 \\ 0 & 2000 & 0 \\ 0 & 0 & 3000 \end{bmatrix} \{\boldsymbol{\alpha}\}_x + \begin{Bmatrix} -16.52 \\ -38.95 \\ -7.005 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix}$$

so that, finally,

$$\{\boldsymbol{\alpha}\}_x = -\begin{bmatrix} 1000 & 0 & 0 \\ 0 & 2000 & 0 \\ 0 & 0 & 3000 \end{bmatrix}^{-1} \begin{Bmatrix} -16.52 \\ -38.95 \\ -7.005 \end{Bmatrix} = \begin{Bmatrix} 0.01652 \\ 0.01948 \\ 0.002335 \end{Bmatrix} (\text{rad/s}^2) \tag{h}$$

These are the components of the angular acceleration in the body frame. To transform them into the inertial frame we use

$$\{\boldsymbol{\alpha}\}_X = [\mathbf{Q}]_{xX}\{\boldsymbol{\alpha}\}_x = ([\mathbf{Q}]_{Xx})^T\{\boldsymbol{\alpha}\}_x = \begin{bmatrix} 0.03086 & -0.9646 & 0.2620 \\ 0.6720 & -0.1740 & -0.7198 \\ 0.7399 & 0.1983 & 0.6428 \end{bmatrix} \begin{Bmatrix} 0.01652 \\ 0.01948 \\ 0.002335 \end{Bmatrix}$$

$$= \begin{Bmatrix} -0.01766 \\ 0.006033 \\ 0.01759 \end{Bmatrix} (\text{rad/s}^2)$$

That is,

$$\underline{\boldsymbol{\alpha} = -0.01766\hat{\mathbf{I}} + 0.006033\hat{\mathbf{J}} + 0.01759\hat{\mathbf{K}}\ (\text{rad/s}^2)}$$

8.1.3 Stability of Torque-Free Motion

Let a rigid body be in torque-free motion with its angular velocity vector directed along the principal body z axis, so that $\boldsymbol{\omega} = \omega_0\hat{\mathbf{k}}$, where ω_0 is constant. The nutation angle is zero and there is no precession. Let us perturb the motion slightly, as illustrated in Figure 8.1-6, so that

$$\omega_x = \delta\omega_x \quad \omega_y = \delta\omega_y \quad \omega_z = \omega_0 + \delta\omega_z \tag{8.1.28}$$

'δ' means a very small quantity. In this case, $\delta\omega_x \ll \omega_0$ and $\delta\omega_y \ll \omega_0$. Thus, the angular velocity vector has become slightly inclined to the z axis. For torque-free motion, $M_{Gx} = M_{Gy} = M_{Gz} = 0$, so that Euler's equations become

$$\begin{aligned} A\dot{\omega}_x + (C-B)\omega_y\omega_z &= 0 \\ B\dot{\omega}_y + (A-C)\omega_x\omega_z &= 0 \\ C\dot{\omega}_z + (B-A)\omega_x\omega_y &= 0 \end{aligned} \tag{8.1.29}$$

Observe that we have not assumed $A = B$, as we did in the previous section. Substituting Equations 8.1.28 into Equations 8.1.29 and keeping in mind our assumption that $\dot{\omega}_0 = 0$, we get

$$\begin{aligned} A\delta\dot{\omega}_x + (C-B)\omega_0\delta\omega_y + (C-B)\delta\omega_y\delta\omega_z &= 0 \\ B\delta\dot{\omega}_y + (A-C)\omega_0\delta\omega_x + (C-B)\delta\omega_x\delta\omega_z &= 0 \\ C\delta\dot{\omega}_z + (B-A)\delta\omega_x\delta\omega_y &= 0 \end{aligned} \tag{8.1.30}$$

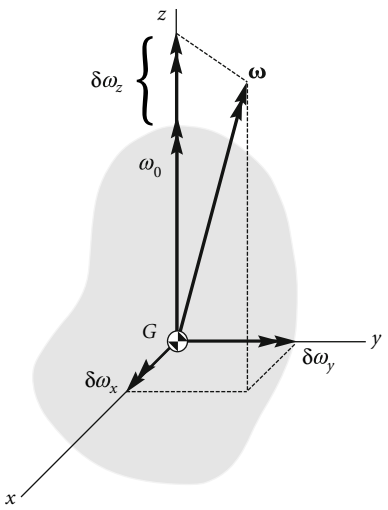


Figure 8.1-6 Principal body axes of a rigid body rotating primarily about the body z axis.

Neglecting all products of the $\delta\omega$s (because they are arbitrarily small), Equations 8.1.30 become

$$\begin{aligned} A\delta\dot{\omega}_x + (C-B)\omega_0\delta\omega_y &= 0 \\ B\delta\dot{\omega}_y + (A-C)\omega_0\delta\omega_x &= 0 \\ C\delta\dot{\omega}_z &= 0 \end{aligned} \tag{8.1.31}$$

Equation $8.1.31_3$ implies that $\delta\omega_z$ is constant. Differentiating Equation $8.1.31_1$ with respect to time, we get

$$A\delta\ddot{\omega}_x + (C-B)\omega_0\delta\dot{\omega}_y = 0 \tag{8.1.32}$$

Solving Equation $8.1.31_2$ for $\delta\dot{\omega}_y$ yields $\delta\dot{\omega}_y = -[(A-C)/B]\omega_0\delta\omega_x$, and substituting this into Equation 8.1.32 gives

$$\delta\ddot{\omega}_x - \frac{(A-C)(C-B)}{AB}\omega_0^2\delta\omega_x = 0 \tag{8.1.33}$$

Likewise, by differentiating Equation $8.1.31_2$ and then substituting $\delta\dot{\omega}_x$ from Equation $8.1.31_1$ yields

$$\delta\ddot{\omega}_y - \frac{(A-C)(C-B)}{AB}\omega_0^2\delta\omega_y = 0 \tag{8.1.34}$$

If we define

$$k = \frac{(A-C)(B-C)}{AB}\omega_0^2 \tag{8.1.35}$$

then both Equations 8.1.33 and 8.1.34 may be written in the form

$$\delta\ddot{\omega} + k\delta\omega = 0 \tag{8.1.36}$$

If $k > 0$, then $\delta\omega \propto e^{\pm i\sqrt{k}t}$, which means $\delta\omega_x$ and $\delta\omega_y$ vary sinusoidally with small amplitude. The motion is therefore bounded and neutrally stable. That means the amplitude does not die out with time, but it does not exceed the small amplitude of the perturbation. Observe from Equation 8.1.35 that $k > 0$ if either $C > A$ and $C > B$ or $C < A$ and $C < B$. This means that the spin axis (z axis) is either the major axis of inertia or the minor axis of inertia. That is, if the spin axis is either the major or minor axis of inertia, the motion is stable. The stability is neutral for a rigid body, because there is no damping.

On the other hand, if $k < 0$, then $\delta\omega \propto e^{\pm\sqrt{k}t}$, which means that the initially small perturbations $\delta\omega_x$ and $\delta\omega_y$ increase without bound. The motion is unstable. From Equation 8.1.35 we see that $k < 0$ if either $A > C$ and C $> B$ or $A < C$ and $C < B$. This means that the spin axis is the intermediate axis of inertia ($A > C > B$ or $B > C > A$). If the spin axis is the intermediate axis of inertia, the motion is unstable.

If the angular velocity of a satellite lies in the direction of its major axis of inertia, the satellite is called a major axis spinner or oblate spinner. A minor axis spinner or prolate spinner has its minor axis of inertia aligned with the angular velocity. 'Intermediate axis spinners' are unstable and will presumably end up being major or minor axis spinners, if the satellite is a rigid body. However, the flexibility inherent in any real satellite leads to an additional instability, as we shall now see.

Consider again the rotationally symmetric satellite in torque-free motion discussed in Section 8.1.2. From Equations 8.1.24 and 8.1.25, we know that the angular momentum $\mathbf{H}_G$ is given by

$$\mathbf{H}_G = A\boldsymbol{\omega}_\perp + C\omega_z\hat{\mathbf{k}} \tag{8.1.37}$$

Hence,

$$H_G^2 = A^2\omega_\perp^2 + C^2\omega_z^2 \tag{8.1.38}$$

Differentiating this equation with respect to time yields

$$\frac{dH_G^2}{dt} = A^2\frac{d\omega_\perp^2}{dt} + 2C^2\omega_z\dot{\omega}_z \tag{8.1.39}$$

But, according to Equation 8.1.1, $\mathbf{H}_G$ is constant, so that $dH_G^2/dt = 0$ and Equation 8.1.39 can be written as

$$\frac{d\omega_\perp^2}{dt} = -2\frac{C^2}{A^2}\omega_z\dot{\omega}_z \tag{8.1.40}$$

The rotary kinetic energy of a rotationally symmetric body ($A = B$) is

$$T_R = \frac{1}{2}A\omega_x^2 + \frac{1}{2}A\omega_y^2 + \frac{1}{2}C\omega_z^2 = \frac{1}{2}A(\omega_x^2 + \omega_y^2) + \frac{1}{2}C\omega_z^2$$

From Equation 8.1.13 we know that $\omega_x^2 + \omega_y^2 = \omega_\perp^2$, which means

$$T_R = \frac{1}{2}A\omega_\perp^2 + \frac{1}{2}C\omega_z^2 \tag{8.1.41}$$

The time derivative of T_R is, therefore,

$$\dot{T}_R = \frac{1}{2}A\frac{d\omega_\perp^2}{dt} + C\omega_z\dot{\omega}_z$$

Solving this for $\dot{\omega}_z$, we get

$$\dot{\omega}_z = \frac{1}{C\omega_z}\left(\dot{T}_R - \frac{1}{2}A\frac{d\omega_\perp^2}{dt}\right)$$

Substituting this expression for $\dot{\omega}_z$ into Equation 8.1.40 and solving for $d\omega_\perp^2/dt$ yields

$$\frac{d\omega_\perp^2}{dt} = 2\frac{C}{A}\frac{\dot{T}_R}{C-A} \tag{8.1.42}$$

Real bodies are not completely rigid, and their flexibility, however slight, gives rise to small dissipative effects which cause the kinetic energy to decrease over time. That is,

$$\dot{T}_R < 0 \quad \text{for satellites with dissipation.} \tag{8.1.43}$$

Substituting this inequality into Equation 8.1.42 leads us to conclude that

$$\begin{aligned} &\frac{d\omega_\perp^2}{dt} < 0 \quad \text{if } C > A \text{ (oblate spinner)} \\ &\frac{d\omega_\perp^2}{dt} > 0 \quad \text{if } C < A \text{ (prolate spinner)} \end{aligned} \tag{8.1.44}$$

If $d\omega_\perp^2/dt$ is negative, the spin is asymptotically stable. Should a non-zero value of $\omega_\perp$ develop for some reason, it will drift back to zero over time so that once again the angular velocity lies completely in the spin direction. On the other hand, if $d\omega_\perp^2/dt$ is positive, the spin is unstable. $\omega_\perp$ does not damp out, and the angular velocity vector drifts away from the spin axis as $\omega_\perp$ increases without bound. We pointed out above that spin about a minor axis of inertia is stable with respect to small disturbances. Now we see that only major axis spin is stable in the long run if dissipative mechanisms exist.

For some additional insight into this phenomenon, solve Equation 8.1.38 for $\omega_\perp^2$,

$$\omega_\perp^2 = \frac{H_G^2 - C^2\omega_z^2}{A^2}$$

and substitute this result into the expression for kinetic energy, Equation 8.1.41, to obtain

$$T_R = \frac{1}{2}\frac{H_G^2}{A} + \frac{1}{2}\frac{(A-C)C}{A}\omega_z^2 \tag{8.1.45}$$

According to Equation 8.1.24

$$\omega_z = \frac{H_{G_z}}{C} = \frac{H_G\cos\theta}{C}$$

Substituting this into Equation 8.1.45 yields the kinetic energy as a function of just the inclination angle θ,

$$T_R = \frac{1}{2}\frac{H_G^2}{A}\left(1 + \frac{A-C}{C}\cos^2\theta\right) \tag{8.1.46}$$

The extreme values of T_R occur at $\theta = 0$ or $\theta = \pi$,

$$T_R = \frac{1}{2}\frac{H_G^2}{C} \quad \text{(major axis spinner)}$$

and $\theta = \pi/2$,

$$T_R = \frac{1}{2}\frac{H_G^2}{A} \quad \text{(minor axis spinner)}$$

Clearly, the kinetic energy of a torque-free satellite is smallest when the spin is around the major axis of inertia. We may think of a satellite with dissipation ($dT_R/dt < 0$) as seeking the state of minimum kinetic energy that occurs when it spins about its major axis.

Example 8.1.5

A rigid spacecraft is modeled by the solid cylinder B which has a mass of 300 kg and the slender rod R which passes through the cylinder and has a mass of 30 kg. Which of the principal axes x, y, z can be an axis about which stable torque-free rotation can occur?

For the cylindrical shell A, we have (Figure 8.1-7)

$$r_B = 0.5 \text{ m} \quad l_B = 1.0 \text{ m} \quad m_B = 300 \text{ kg}$$

The principal moments of inertia about the center of mass are found in Figure 8.1-9(b),

$$I_{B_x} = \frac{1}{4}m_B r_B^2 + \frac{1}{12}m_B l_B^2 = 43.75 \text{ kg.m}^2$$
$$I_{B_y} = I_{B_{xx}} = 43.75 \text{ kg.m}^2$$
$$I_{B_z} = \tfrac{1}{2}m_B r_B^2 = 37.5 \text{ kg.m}^2$$

The properties of the transverse rod are

$$l_R = 1.0 \text{ m} \quad m_R = 30 \text{ kg}$$

Figure 8.1-9(a), with $r = 0$, yields the moments of inertia,

$$I_{R_y} = 0$$
$$I_{R_z} = I_{R_x} = \frac{1}{12}m_A r_A^2 = 10.0 \text{ kg}\cdot\text{m}^2$$

The moments of inertia of the assembly is the sum of the moments of inertia of the cylinder and the rod,

$$I_x = I_{B_x} + I_{R_x} = 53.75 \text{ kg} \cdot \text{m}^2$$
$$I_y = I_{B_y} + I_{R_y} = 43.75 \text{ kg} \cdot \text{m}^2$$
$$I_z = I_{B_z} + I_{R_z} = 47.50 \text{ kg} \cdot \text{m}^2$$

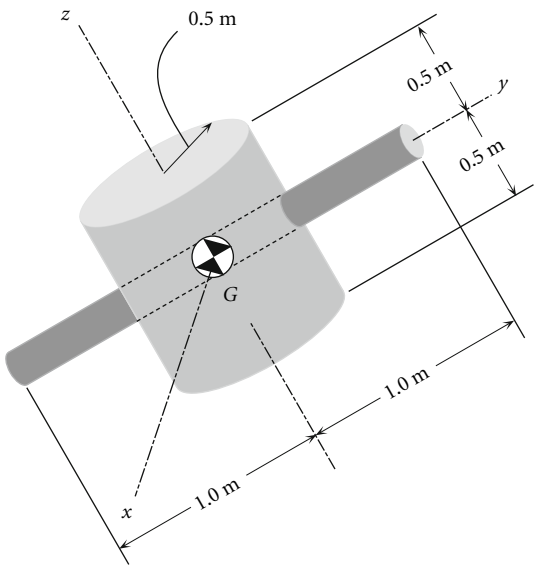


Figure 8.1-7 Built-up satellite structure.

Since I_z is the intermediate mass moment of inertia, rotation about the z axis is unstable. With energy dissipation, rotation is stable in the long term only about the major axis, which in this case is the x axis.

8.1.4 Dual-Spin Spacecraft

If a satellite is to be spin stabilized, it must be an oblate spinner. The diameter of the spacecraft is restricted by the cross-section of the launch vehicle's upper stage, and its length is limited by stability requirements. Therefore, oblate spinners cannot take full advantage of the payload volume available in a given launch vehicle, which after all are slender, prolate shapes for aerodynamic reasons. The dual-spin design permits spin stabilization of a prolate shape.

The axisymmetric, dual-spin configuration, or gyrostat, consists of an axisymmetric rotor and a smaller axisymmetric platform joined together along a common longitudinal spin axis at a bearing, as shown in Figure 8.1-8. The platform and rotor have their own components of angular velocity, $\boldsymbol{\omega}_p$ and $\boldsymbol{\omega}_r$, respectively, along the spin axis direction $\hat{\mathbf{k}}$. The platform spins at a much slower rate than the rotor. The assembly acts like a rigid body as far as transverse rotations are concerned; i.e., the rotor and the platform have $\boldsymbol{\omega}_\perp$ in common. An electric motor integrated into the axle bearing connecting the two components acts to overcome frictional torque which would otherwise eventually cause the relative angular velocity between the rotor and platform to go to zero. If that should happen, the satellite would become a single spin unit, probably an unstable prolate spinner, since the rotor of a dual-spin spacecraft is likely to be prolate.

The first dual-spin satellite was OSO-I (Orbiting Solar Observatory), which NASA launched in 1962. It was a major-axis spinner. The first prolate dual-spin spacecraft was the two-storey tall TACSAT I (Tactical Communications Satellite). It was launched into geosynchronous orbit by the US Air Force in 1969. Typical of many of today's communications satellites, TACSAT's platform rotated at one revolution per day to keep its antennas pointing towards the earth. The rotor spun at about one revolution per second. Of course, the axis of the spacecraft was normal to the plane of its orbit. The first dual-spin interplanetary spacecraft was Galileo. Galileo's platform was completely despun to provide a fixed orientation for cameras and other instruments. The rotor spun at three revolutions per minute.

The equations of motion of a dual-spin spacecraft will be developed later in Section 8.1.8. Let us determine the stability of the motion by following the same 'energy sink' procedure employed in the previous section for

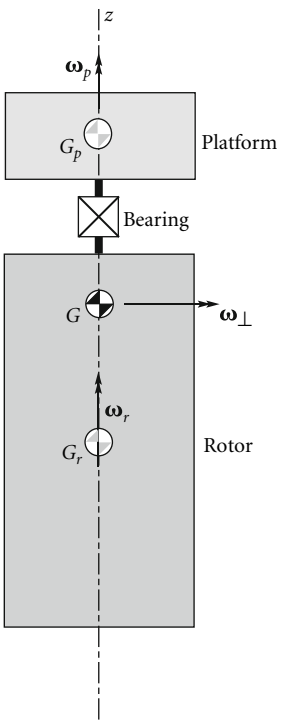


Figure 8.1-8 Axisymmetric, dual-spin satellite.

a single-spin stabilized spacecraft. The angular momentum of the dual-spin configuration about the space-craft's center of mass G is the sum of the angular momenta of the rotor (r) and the platform (p) about G,

$$\mathbf{H}_G = \mathbf{H}_G^{(p)} + \mathbf{H}_G^{(r)} \tag{8.1.47}$$

The angular momentum of the platform about the spacecraft center of mass is

$$\mathbf{H}_G^{(p)} = C_p \omega_p \hat{\mathbf{k}} + \mathrm{A}_p \boldsymbol{\omega}_\perp \tag{8.1.48}$$

where C_p is the moment of inertia of the platform about the spacecraft spin axis, and A_p is its transverse moment of inertia about G (not G_p). Likewise, for the rotor,

$$\mathbf{H}_G^{(r)} = C_r \omega_r \hat{\mathbf{k}} + A_r \boldsymbol{\omega}_\perp \tag{8.1.49}$$

where C_r and A_r are its longitudinal and transverse moments of inertia about axes through G. Substituting Equations 8.1.48 and 8.1.49 into 8.1.47 yields

$$\mathbf{H}_G = (C_r \omega_r + C_p \omega_p)\hat{\mathbf{k}} + A_\perp \boldsymbol{\omega}_\perp \tag{8.1.50}$$

where $A_\perp$ is the total transverse moment of inertia,

$$A_\perp = A_p + A_r$$

From this it follows that

$$H_G^2 = (C_r \omega_r + C_p \omega_p)^2 + A_\perp^2 \omega_\perp^2$$

For torque-free motion, $\dot{\mathbf{H}}_G = \mathbf{0}$, so that $dH_G^2/dt = 0$, or

$$2(C_r \omega_r + C_p \omega_p)(C_r \dot{\omega}_r + C_p \dot{\omega}_p) + A_\perp^2 \frac{d\omega_\perp^2}{dt} = 0 \tag{8.1.51}$$

Solving this for $d\omega_\perp^2/dt$ yields

$$\frac{d\omega_\perp^2}{dt} = -\frac{2}{A_\perp^2}(C_r \omega_r + C_p \omega_p)(C_r \dot{\omega}_r + C_p \dot{\omega}_p) \tag{8.1.52}$$

The total rotational kinetic energy of rotation of the dual spin spacecraft is the sum of that of the rotor and the platform,

$$T = \frac{1}{2} C_r \omega_r^2 + \frac{1}{2} C_p \omega_p^2 + \frac{1}{2} A_\perp \omega_\perp^2$$

Differentiating this expression with respect to time and solving for $d\omega_\perp^2/dt$ yields

$$\frac{d\omega_\perp^2}{dt} = \frac{2}{A_\perp}(\dot{T} - C_r \omega_r \dot{\omega}_r - C_p \omega_p \dot{\omega}_p) \tag{8.1.53}$$

$\dot{T}$ is the sum of the power $P^{(r)}$ dissipated in the rotor and the power $P^{(p)}$ dissipated in the platform,

$$\dot{T} = P^{(r)} + P^{(p)} \tag{8.1.54}$$

Substituting Equation 8.1.54 into 8.1.53 we find

$$\frac{d\omega_\perp^2}{dt} = \frac{2}{A_\perp}(P^{(r)} - C_r \omega_r \dot{\omega}_r + P^{(p)} - C_p \omega_p \dot{\omega}_p) \tag{8.1.55}$$

Equating the two expressions for $d\omega_\perp^2/dt$ in Equations 8.1.52 and 8.1.55 yields

$$\frac{2}{A_\perp}(\dot{T} - C_r \omega_r \dot{\omega}_r - C_p \omega_p \dot{\omega}_p) = -\frac{2}{A_\perp^2}(C_r \omega_r + C_p \omega_p)(C_r \dot{\omega}_r + C_p \dot{\omega}_p)$$

Solve this for $\dot{T}$ to obtain

$$\dot{T} = \frac{C_r}{A_\perp}[(A_\perp - C_r)\omega_r - C_p \omega_p]\dot{\omega}_r + \frac{C_p}{A_\perp}[(A_\perp - C_p)\omega_p - C_r \omega_r]\dot{\omega}_p \tag{8.1.56}$$

Following Likins (1967), we identify the terms containing $\dot{\omega}_r$ and $\dot{\omega}_p$ as the power dissipation in the rotor and platform, respectively. That is, comparing Equations 8.1.54 and 8.1.56,

$$P^{(r)} = \frac{C_r}{A_\perp}[(A_\perp - C_r)\omega_r - C_p\omega_p]\dot{\omega}_r \tag{8.1.57a}$$

$$P^{(p)} = \frac{C_p}{A_\perp}[(A_\perp - C_p)\omega_p - C_r\omega_r]\dot{\omega}_p \tag{8.1.57b}$$

Solving these two expressions for $\dot{\omega}_r$ and $\dot{\omega}_p$, respectively, yields

$$\dot{\omega}_r = \frac{A_\perp}{C_r}\frac{p^{(r)}}{(A_\perp - C_r)\omega_r - C_p\omega_p} \tag{8.1.58a}$$

$$\dot{\omega}_p = \frac{A_\perp}{C_p}\frac{P^{(p)}}{(A_\perp - C_p)\omega_p - C_r\omega_r} \tag{8.1.58b}$$

Substituting these results into Equation 8.1.55 leads to

$$\frac{d\omega_\perp^2}{dt} = \frac{2}{A_\perp}\left[\frac{P^{(r)}}{C_p\dfrac{\omega_p}{\omega_r} - (A_\perp - C_r)} + \frac{P^{(p)}}{C_r - (A_\perp - C_p)\dfrac{\omega_p}{\omega_r}}\right]\left(C_r + C_p\frac{\omega_p}{\omega_r}\right) \tag{8.1.59}$$

As pointed out above, for geosynchronous dual-spin communication satellites,

$$\frac{\omega_p}{\omega_r} \approx \frac{2\pi \text{ rad/d}}{2\pi \text{ rad/s}} \approx 10^{-5}$$

whereas for interplanetary dual-spin spacecraft, $\omega_p = 0$. Therefore, there is an important class of spin stabilized spacecraft for which $\omega_p/\omega_r \approx 0$. For a despun platform wherein ω_p is zero (or nearly so), Equation 8.1.59 yields

$$\frac{d\omega_\perp^2}{dt} = \frac{2}{A_\perp}\left[P^{(p)} + \frac{C_r}{C_r - A_\perp}P^{(r)}\right] \tag{8.1.60}$$

If the rotor is oblate ($C_r > A_\perp$), then, since $P^{(r)}$ and $P^{(p)}$ are both negative, it follows from Equation 8.1.60 that $d\omega_\perp^2/dt < 0$. That is, the oblate dual spin configuration with a despun platform is unconditionally stable. In practice, however, the rotor is likely to be prolate ($C_r < A_\perp$), so that

$$\frac{C_r}{C_r - A_\perp}P^{(r)} > 0$$

In that case, $d\omega_\perp^2/dt < 0$ only if the dissipation in the platform is significantly greater than that of the rotor. Specifically, for a prolate design it must be true that

$$|P^{(p)}| > \left|\frac{C_r}{C_r - A_\perp}P^{(r)}\right|$$

The platform dissipation rate $P^{(p)}$ can be augmented by adding nutation dampers, which are discussed in the next section.

For the despun prolate dual-spin configuration, Equations 8.1.58 imply

$$\dot{\omega}_r = \frac{P^{(r)}}{(A_\perp - C_r)}\frac{A_\perp}{C_r\omega_r}$$

$$\dot{\omega}_p = -\frac{P^{(p)}}{C_p}\frac{A_\perp}{C_r\omega_r}$$

Clearly, the signs of $\dot{\omega}_r$ and $\dot{\omega}_p$ are opposite. If $\omega_r > 0$, then dissipation causes the spin rate of the rotor to decrease and that of the platform to increase. Were it not for the action of the motor on the shaft connecting the two components of the spacecraft, eventually $\omega_p = \omega_r$. That is, the relative motion between the platform and rotor would cease and the dual-spinner would become an unstable single spin spacecraft. Setting $\omega_p = \omega_r$ in Equation 8.1.59 yields

$$\frac{d\omega_\perp^2}{dt} = 2\frac{C_r + C_p}{A_\perp}\frac{P^{(r)} + P^{(p)}}{(C_r + C_p) - A_\perp}$$

which is the same as Equation 8.1.42, the energy sink conclusion for a single spinner.

8.1.5 Nutation Damper

Nutation dampers are passive means of dissipating energy. A common type consists essentially of a tube filled with viscous fluid and containing a mass attached to springs, as illustrated in Figure 8.1-9. Dampers may contain just fluid, only partially filling the tube so it can slosh around. In either case, the purpose is to dissipate energy through fluid friction. The wobbling of the spacecraft due to non-alignment of the angular velocity with the principal spin axis induces accelerations throughout the satellite, giving rise to the sloshing of fluids, stretching and flexing of non-rigid components, etc., all of which dissipate energy to one degree or another. Nutation dampers are added to deliberately increase energy dissipation, which is desirable for stabilizing oblate single spinners and dual-spin spacecraft.

Let us focus on the motion of the mass within the nutation damper of Figure 8.1-9 in order to gain some insight into how relative motion and deformation are induced by the satellite's precession. Note that point P is the center of mass of the rigid satellite body itself. The center of mass G of the satellite-damper mass combination lies between P and m, as shown in Figure 8.1-9. We suppose that the tube is lined up with the z axis of the body-fixed xyz frame, as shown. The mass m in the tube is therefore constrained by the tube walls to move only in the z direction. When the springs are undeformed, the mass lies in the xy plane. In general, the position vector of m in the body frame is

$$\mathbf{r} = R\hat{\mathbf{i}} + z_m\hat{\mathbf{k}} \tag{8.1.61}$$

where z_m is the z coordinate of m and R is the distance of the damper from the centerline of the spacecraft. The velocity and acceleration of m relative to the satellite are, therefore,

$$\mathbf{v}_{\text{rel}} = \dot{z}_m\hat{\mathbf{k}} \tag{8.1.62}$$

$$\mathbf{a}_{\text{rel}} = \ddot{z}_m\hat{\mathbf{k}} \tag{8.1.63}$$

The absolute angular velocity $\boldsymbol{\omega}$ of the satellite (and, therefore, the body frame) is

$$\boldsymbol{\omega} = \omega_x\hat{\mathbf{i}} + \omega_y\hat{\mathbf{j}} + \omega_z\hat{\mathbf{k}} \tag{8.1.64}$$

When $\boldsymbol{\omega}$ is given in a body frame, we find the absolute angular acceleration by taking the time derivative of $\boldsymbol{\omega}$, holding the unit vectors fixed. Thus,

$$\dot{\boldsymbol{\omega}} = \dot{\omega}_x\hat{\mathbf{i}} + \dot{\omega}_y\hat{\mathbf{j}} + \dot{\omega}_z\hat{\mathbf{k}} \tag{8.1.65}$$

The absolute acceleration of m for the case at hand becomes

$$\mathbf{a} = \mathbf{a}_P + \dot{\boldsymbol{\omega}} \times \mathbf{r} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) + 2\boldsymbol{\omega} \times \mathbf{v}_{\text{rel}} + \mathbf{a}_{\text{rel}} \tag{8.1.66}$$

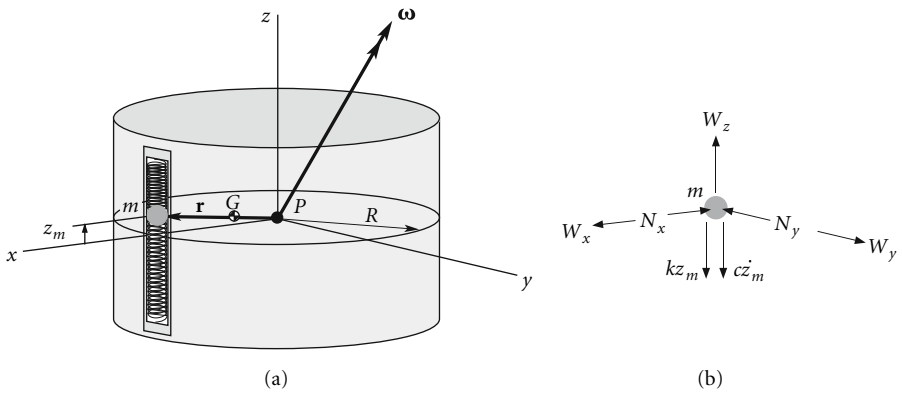


Figure 8.1-9 (a) Precessing oblate spacecraft with a nutation damper aligned with the z axis. (b) Free-body diagram of the moving mass in the nutation damper.

in which $\mathbf{a}_P$ is the absolute acceleration of the reference point P. Substituting Equations 8.1.61 through 8.1.65 into Equation 8.1.66, carrying out the vector operations, combining terms, and simplifying leads to the following expressions for the three components of the inertial acceleration of m,

$$\begin{aligned} a_x &= a_{P_x} - R(\omega_y^2 + \omega_z^2) + z_m\dot{\omega}_y + z_m\omega_x\omega_z + 2\dot{z}_m\omega_y \\ a_y &= a_{P_y} - R\dot{\omega}_z + R\omega_x\omega_y - z_m\dot{\omega}_x + z_m\omega_y\omega_z - 2\dot{z}_m\omega_x \\ a_z &= a_{P_z} - z_m(\omega_x^2 + \omega_y^2) - R\dot{\omega}_y + R\omega_x\omega_z + \ddot{z}_m \end{aligned} \tag{8.1.67}$$

Figure 8.1-9(b) shows the free-body diagram of the damper mass m. In the x and y directions the forces on m are the components of the force of gravity (W_x and W_y) and the components N_x and N_y of the force of contact with the smooth walls of the damper tube. The directions assumed for these components are, of course, arbitrary. In the z direction, we have the z component W_z of the weight, plus the force of the springs and the viscous drag of the fluid. The spring force ($-kz_m$) is directly proportional and opposite in direction to the displacement z_m. k is the net spring constant. The viscous drag ($-c\dot{z}_m$) is directly proportional and opposite in direction to the velocity $\dot{z}_m$ of m relative to the tube. c is the damping constant. Thus, the three components of the net force on the damper mass m are

$$\begin{aligned} F_{\text{net}_x} &= W_x - N_x \\ F_{\text{net}_y} &= W_y - N_y \\ F_{\text{net}_z} &= W_z - kz_m - c\dot{z}_m \end{aligned} \tag{8.1.68}$$

Substituting Equations 8.1.67 and 8.1.68 into Newton's second law, $\mathbf{F}_{\text{net}} = m\mathbf{a}$, yields

$$\begin{aligned} N_x &= mR(\omega_y^2 + \omega_z^2) - mz_m\dot{\omega}_y - mz_m\omega_x\omega_y - 2m\dot{z}_m\omega_y + \overbrace{(W_x - ma_{P_x})}^{=0} \\ N_y &= -mR\dot{\omega}_z - mR\omega_x\omega_y + mz_m\dot{\omega}_x - mz_m\omega_y\omega_z \\ &\quad + 2m\dot{z}_m\omega_x + \overbrace{(W_y - ma_{P_y})}^{=0} \\ m\ddot{z}_m + c\dot{z}_m + [k - m(\omega_x^2 + \omega_y^2)]z_m &= mR(\dot{\omega}_y - \omega_x\omega_z) + \overbrace{(W_z - ma_{P_z})}^{=0} \end{aligned} \tag{8.1.69}$$

The last terms in parentheses in each of these expressions vanish if the acceleration of gravity is the same at m as at the reference point P of the spacecraft. This will be true unless the satellite is of enormous size.

If the damper mass m is vanishingly small compared to the mass M of the rigid spacecraft body, then it will have little effect on the rotary motion. If the rotational state is that of an axisymmetric satellite in torque-free motion, then we know from Equations 8.1.13, 8.1.14 and 8.1.19 that

$$\begin{aligned} \omega_x &= \Omega\sin\omega_s t \qquad & \omega_y &= \Omega\cos\omega_s t \qquad & \omega_z &= \omega_0 \\ \dot{\omega}_x &= \Omega\omega_s\cos\omega_s t \qquad & \dot{\omega}_y &= -\Omega\omega_s\sin\omega_s t \qquad & \dot{\omega}_z &= 0 \end{aligned}$$

in which case Equations 8.1.69 become

$$\begin{aligned} N_x &= mR(\omega_0^2 + \Omega^2\cos^2\omega_s t) + m(\omega_s - \omega_0)\Omega z_m\sin\omega_s t - 2m\Omega\dot{z}_m\cos\omega_s t \\ N_y &= -mR\Omega^2\cos\omega_s t\sin\omega_s t + m(\omega_s - \omega_0)\Omega z_m\cos\omega_s t + 2m\Omega\dot{z}_m\sin\omega_s t \\ m\ddot{z}_m + c\dot{z}_m + (k - m\Omega^2)z_m &= -mR(\omega_s + \omega_0)\Omega\sin\omega_s t \end{aligned} \tag{8.1.70}$$

Equation $8.1.70_3$ is that of a single degree of freedom, damped oscillator with a sinusoidal forcing function. The precession produces a force of amplitude $m(\omega_0 + \omega_s)\Omega R$ and frequency ω_s which causes the damper mass m to oscillate back and forth in the tube, such that

$$z_m = \frac{mR\Omega(\omega_s + \omega_0)}{[k - m(\omega_s^2 + \Omega^2)]^2 + (c\omega_s)2}\{c\omega_s\cos\omega_s t - [k - m(\omega_s^2 + \Omega^2)\sin\omega_s t]\}$$

Observe that the contact forces N_x and N_y depend exclusively on the amplitude and frequency of the precession. If the angular velocity lines up with the spin axis, so that $\Omega = 0$ (precession vanishes), then

$$\begin{aligned} N_x &= m\omega_0^2 R \\ N_y &= 0 \qquad \text{No precession.} \\ z_m &= 0 \end{aligned}$$

If precession is eliminated, so there is pure spin around the principal axis, the time-varying motions and forces vanish throughout the spacecraft, which thereafter rotates as a rigid body with no energy dissipation.

Now, the whole purpose of a nutation damper is to interact with the rotational motion of the satellite so as to damp out any tendencies to precess. Therefore, its mass should not be ignored in the equations of motion of the satellite. We will derive the equations of motion of the rigid satellite with nutation damper to show how rigid body mechanics is brought to bear upon the problem and, simply, to discover precisely what we are up against in even this extremely simplified system. We will continue to use P as the origin of our body frame. Since a moving mass has been added to the rigid satellite and since we are not using the center of mass of the system as our reference point, we cannot use Euler's equations. The equation of rotational motion of the system of satellite plus damper is

$$\dot{\mathbf{H}}_{P_{\text{rel}}} + \mathbf{r}_{G/P} \times (M+m)\mathbf{a}_{P/G} = \mathbf{M}_{G_{\text{net}}} \tag{8.1.71}$$

The angular momentum of the satellite body plus that of the damper mass, relative to point P on the spacecraft, is

$$\mathbf{H}_{P_{\text{rel}}} = \overbrace{A\omega_x\hat{\mathbf{i}} + B\omega_y\hat{\mathbf{j}} + C\omega_z\hat{\mathbf{k}}}^{\text{body of the spacecraft}} + \overbrace{\mathbf{r} \times m\dot{\mathbf{r}}}^{\text{damper mass}} \tag{8.1.72}$$

where the position vector $\mathbf{r}$ is given by Equation 8.1.61. We have

$$\dot{\mathbf{r}} = \left(\frac{d\mathbf{r}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega} \times \mathbf{r} = \dot{z}_m\hat{\mathbf{k}} + \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \omega_x & \omega_y & \omega_z \\ R & 0 & z \end{vmatrix}$$

$$= \omega_y z_m\hat{\mathbf{i}} + (\omega_z R - \omega_x z_m)\hat{\mathbf{j}} + (\dot{z}_m - \omega_y R)\hat{\mathbf{k}}$$

After substituting this into Equation 8.1.72 and collecting terms we obtain

$$\begin{aligned} \mathbf{H}_{P_{\text{rel}}} &= [(A + mz_m^2)\omega_x - mRz_m\omega_z]\hat{\mathbf{i}} \\ &\quad + [(B + mR^2 + mz_m^2)\omega_y - mR\dot{z}_m]\hat{\mathbf{j}} \\ &\quad + [(C + mR^2)\omega_z - mRz_m\omega_x]\hat{\mathbf{k}} \end{aligned} \tag{8.1.73}$$

Thus,

$$\dot{\mathbf{H}}_{P_{\text{rel}}} = \left(\frac{d\mathbf{H}_{P_{\text{rel}}}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega} \times \mathbf{H}_{P_{\text{rel}}}$$

Carrying out the operations on the right leads eventually to

$$\begin{aligned} \dot{\mathbf{H}}_{P_{\text{rel}}} &= [(A + mz_m^2)\dot{\omega}_x - mRz_m\dot{\omega}_z + (C - B - mz_m^2)\omega_y\omega_z \\ &\quad - mRz_m\omega_x\omega_y + 2mz_m\dot{z}_m\omega_x]\hat{\mathbf{i}} \\ &\quad + \{(B + mR^2 + mz_m^2)\dot{\omega}_y + mRz_m(\omega_x^2 - \omega_z^2) \\ &\quad + [A + mz_m^2 - (C + mR^2)]\omega_x\omega_z + 2mz_m\dot{z}_m\omega_y - mR\ddot{z}_m\}\hat{\mathbf{j}} \\ &\quad + [-mRz_m\dot{\omega}_x + (C + mR^2)\dot{\omega}_z + (B + mR^2 - A)\omega_x\omega_y \\ &\quad + mRz_m\omega_y\omega_z - 2mR\dot{z}_m\omega_x]\hat{\mathbf{k}} \end{aligned} \tag{8.1.74}$$

To calculate the second term on the left of Equation 8.1.71, we keep in mind that P is the center of mass of the body of the satellite and first determine the position vector of the center of mass G of the vehicle plus damper relative to P,

$$(M+m)\mathbf{r}_{G/P} = M(\mathbf{0}) + m\mathbf{r} \tag{8.1.75}$$

where $\mathbf{r}$, the position of the damper mass m relative to P, is given by Equation 8.1.61.

Thus

$$\mathbf{r}_{G/P} = \frac{m}{m+M}\mathbf{r} = \mu\mathbf{r} = \mu(R\hat{\mathbf{i}} + z_m\hat{\mathbf{k}}) \tag{8.1.76}$$

in which

$$\mu = \frac{m}{m+M} \tag{8.1.77}$$

Thus,

$$\mathbf{r}_{G/P} \times (M+m)\mathbf{a}_{P/G} = \left(\frac{m}{M+m}\right)\mathbf{r} \times (M+m)\mathbf{a}_{P/G} = \mathbf{r} \times m\mathbf{a}_{P/G} \tag{8.1.78}$$

The acceleration of P relative to G is

$$\mathbf{a}_{P/G} = -\ddot{\mathbf{r}}_{G/P} = -\mu\frac{d^2\mathbf{r}}{dt^2} = -\mu\left[\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{\text{rel}} + \dot{\boldsymbol{\omega}} \times \mathbf{r} + \boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) + 2\boldsymbol{\omega} \times \left(\frac{d\mathbf{r}}{dt}\right)_{\text{rel}}\right] \tag{8.1.79}$$

where

$$\left(\frac{d\mathbf{r}}{dt}\right)_{\text{rel}} = \frac{dR}{dt}\hat{\mathbf{i}} + \frac{dz_m}{dt}\hat{\mathbf{k}} = \dot{z}_m\hat{\mathbf{k}} \tag{8.1.80}$$

and

$$\left(\frac{d^2\mathbf{r}}{dt^2}\right)_{\text{rel}} = \frac{d^2R}{dt^2}\hat{\mathbf{i}} + \frac{d^2z_m}{dt^2}\hat{\mathbf{k}} = \ddot{z}_m\hat{\mathbf{k}} \tag{8.1.81}$$

Substituting Equations 8.1.61, 8.1.64, 8.1.65, 8.1.80 and 8.1.81 into Equation 8.1.79 yields

$$\begin{aligned}\mathbf{a}_{P/G} = {} & [-\mu z_m\dot{\omega}_y + \mu R(\omega_y^2 + \omega_z^2) - \mu z_m\omega_x\omega_z - 2\mu\dot{z}_m\omega_y]\hat{\mathbf{i}} \\ & + (\mu z_m\dot{\omega}_x - \mu R\dot{\omega}_z - \mu R\omega_x\omega_y - \mu z_m\omega_y\omega_z + 2\mu\dot{z}_m\omega_x)\hat{\mathbf{j}} \\ & + [\mu R\dot{\omega}_y + \mu z_m(\omega_x^2 + \omega_y^2) - \mu R\omega_x\omega_z - \mu\ddot{z}_m]\hat{\mathbf{k}}\end{aligned} \tag{8.1.82}$$

We move this expression into Equation 8.1.78 to get

$$\begin{aligned}&\mathbf{r}_{G/P} \times (M+m)\mathbf{a}_{P/G} \\ &\quad = [-\mu m z_m^2\dot{\omega}_x - 2\mu m\ddot{z}_m\omega_x + \mu mR(\omega_x\omega_y + \dot{\omega}_z) + \mu m z_m^2\omega_y\omega_z]\hat{\mathbf{i}} \\ &\qquad + [-\mu m(R^2 + z_m^2)\dot{\omega}_y - 2\mu m z_m\dot{z}_m\omega_y - \mu mRz_m(\omega_z^2 - \omega_x^2) + \mu m(R^2 - z_m^2)\omega_x\omega_z + \mu mR\ddot{z}_m]\hat{\mathbf{j}} \\ &\qquad + \left(\mu mRz_m\dot{\omega}_x - \mu mR^2\dot{\omega}_z + 2\mu mR\dot{z}_m\omega_x - \mu mR^2\omega_x\omega_y - \mu mRz_m\omega_y\omega_z\right)\hat{\mathbf{k}}\end{aligned}$$

Placing this result and Equation 8.1.74 in Equation 8.1.71, and using the fact that $\mathbf{M}_{G_{\text{net}}} = \mathbf{0}$, yields a vector equation whose three components are

$$\begin{aligned}&A\dot{\omega}_x + (C-B)\omega_y\omega_z + (1-\mu)mz_m^2\dot{\omega}_x - (1-\mu)mz_m^2\omega_y\omega_z \\ &\quad + 2(1-\mu)mz_m\dot{z}_m\omega_x - (1-\mu)mRz_m\omega_x\omega_y = 0 \\ &[B + (1-\mu)mR^2]\dot{\omega}_y + [A - C - (1-\mu)mR^2]\omega_x\omega_z \\ &\quad + (1-\mu)mz_m^2(\omega_x\omega_z + \dot{\omega}_y) + 2(1-\mu)mz_m\dot{z}_m\omega_y \\ &\quad - (1-\mu)mR\ddot{z}_m + (1-\mu)mRz_m(\omega_x^2 - \omega_z^2) = 0 \\ &[C + (1-\mu)mR^2]\dot{\omega}_z + [B - A + (1-\mu)mR^2]\omega_x\omega_y \\ &\quad + (1-\mu)mRz_m\omega_y\omega_z - 2(1-\mu)mR\dot{z}_m\omega_x - (1-\mu)mRz_m\dot{\omega}_x = 0\end{aligned} \tag{8.1.83}$$

These are three equations in the four unknowns ω_x, ω_y, ω_z and z_m. The fourth equation is that of the motion of the damper mass m in the z direction,

$$W_z - kz_m - c\dot{z}_m = ma_z \tag{8.1.84}$$

where a_z is given by Equation $8.1.67_3$, in which $a_{P_z} = a_{P_z} - a_{G_z} + a_{G_z} = a_{P/G_z} + a_{G_z}$, so that

$$a_z = a_{P/G_z} + a_{G_z} - z_m(\omega_x^2 + \omega_y^2) - R\dot{\omega}_y + R\omega_x\omega_z + \ddot{z}_m \tag{8.1.85}$$

Substituting the z component of Equation 8.1.82 into this expression and that result into Equation 8.1.84 leads (with $W_z = ma_{G_z}$) to

$$(1-\mu)m\ddot{z}_m + c\dot{z}_m + [k - (1-\mu)m(\omega_x^2 + \omega_y^2)]z_m = (1-\mu)mR[\dot{\omega}_y - \omega_x\omega_z] \tag{8.1.86}$$

Compare Equation $8.1.69_3$ with this expression, which is the fourth equation of motion we need.

Equations 8.1.83 and 8.1.86 are a rather complicated set of non-linear, second order differential equations, which must be solved (numerically) to obtain a precise description of the motion of the semirigid spacecraft. That is beyond our scope. However, to study their stability we can linearize the equations in much the same way as we did in Section 8.1.3. (Note that Equations 8.1.83 reduce to 8.1.29 when $m = 0$.) We assume the satellite is in pure spin with angular velocity ω_0 about the z axis and that the damper mass is at rest ($z_m = 0$). This motion is slightly perturbed, in such a way that

$$\omega_x = \delta\omega_x \qquad \omega_y = \delta\omega_y \qquad \omega_z = \omega_0 + \delta\omega_z \qquad z_m = \delta z_m \tag{8.1.87}$$

It will be convenient for this analysis to introduce operator notation for the time derivative, $D = d/dt$. Thus, given a function of time $f(t)$, for any integer n, $D^n f = d^n f/dt^n$, and $D^0 f(t) = f(t)$. Then the various time derivatives throughout the equations will, in accordance with Equation 8.1.87, be replaced as follows,

$$\dot{\omega}_x = D\delta\omega_x \qquad \dot{\omega}_y = D\delta\omega_y \qquad \dot{\omega}_z = D\delta\omega_z \qquad \dot{z}_m = D\delta z_m \qquad \ddot{z}_m = D^2\delta z_m \tag{8.1.88}$$

Substituting Equations 8.1.87 and 8.1.88 into Equations 8.1.83 and 8.1.86 and retaining only those terms which are at most linear in the small perturbations leads to

$$\begin{aligned}
&AD\delta\omega_x + (C - B)\omega_0\delta\omega_y = 0\\
&[A - C - (1 - \mu)mR^2]\omega_0\delta\omega_x + [B + (1 - \mu)mR^2]D\delta\omega_y\\
&\quad -(1 - \mu)mR(D^2 + \omega_0^2)\delta z_m = 0\\
&[C + (1 - \mu)mR^2]D\delta\omega_z = 0\\
&(1 - \mu)mR\omega_0\delta\omega_x - (1 - \mu)mRD\delta\omega_y + [(1 - \mu)mD^2 + cD + k]\delta z_m = 0
\end{aligned} \tag{8.1.89}$$

$\delta\omega_z$ appears only in the third equation, which states that $\delta\omega_z =$ constant. The first, second and fourth equations may be combined in matrix notation,

$$\begin{bmatrix} AD & (C - B)\omega_0 & 0\\ [A - C - (1 - \mu)mR^2]\omega_0 & [B + (1 - \mu)mR^2]D & -(1 - \mu)mR(D^2 + \omega_0^2)\\ (1 - \mu)mR\omega_0 & -(1 - \mu)mRD & (1 - \mu)mD^2 + cD + k \end{bmatrix} \times \begin{Bmatrix} \delta\omega_x\\ \delta\omega_y\\ \delta z_m \end{Bmatrix} = \begin{Bmatrix} 0\\ 0\\ 0 \end{Bmatrix} \tag{8.1.90}$$

This is a set of three linear differential equations in the perturbations $\delta\omega_x$, $\delta\omega_y$ and δz_m. We won't try to solve them, since all we are really interested in is the stability of the satellite-damper system. It can be shown that the determinant Δ of the 3 by 3 matrix in Equation 8.1.90 is

$$\Delta = a_4D^4 + a_3D^3 + a_2D^2 + a_1D + a_0 \tag{8.1.91}$$

in which the coefficients of the characteristic equation $\Delta = 0$ are

$$\begin{aligned}
a_4 &= (1 - \mu)mAB\\
a_3 &= cA[B + (1 - \mu)mR^2]\\
a_2 &= k[B + (1 - \mu)mR^2]A + (1 - \mu)m[(A - C)(B - C)\\
&\quad -(1 - \mu)AmR^2]\omega_0^2\\
a_1 &= c\{[A - C - (1 - \mu)mR^2](B - C)\}\omega_0^2\\
a_0 &= k\{[A - C - (1 - \mu)mR^2](B - C)\}\omega_0^2 + [(B - C)(1 - \mu)^2]m^2R^2\omega_0^4
\end{aligned} \tag{8.1.92}$$

According to the Routh–Hurwitz stability criteria (see any text on control systems, e.g., Palm, 1983), the motion represented by Equations 8.1.90 is asymptotically stable if and only if the signs of all of the following quantities, defined in terms of the coefficients of the characteristic equation, are the same

$$r_1 = a_4 \qquad r_2 = a_3 \qquad r_3 = a_2 - \frac{a_4a_1}{a_3} \qquad r_4 = a_1 - \frac{a_3a_0}{a_3a_2 - a_4a_1} \qquad r_5 = a_0 \tag{8.1.93}$$

Example 8.1.6

A satellite is spinning about the z axis of its principal body frame at 2π radians per second. The principal moments of inertia about its center of mass are

$$A = 300 \text{ kg} \cdot \text{m}^2 \quad B = 400 \text{ kg} \cdot \text{m}^2 \quad C = 500 \text{ kg} \cdot \text{m}^2 \tag{a}$$

For the nutation damper, the following properties are given

$$R = 1 \text{ m} \qquad \mu = 0.01 \qquad m = 10 \text{ kg} \qquad k = 10\,000 \text{ N/m} \qquad c = 150 \text{ N} \cdot \text{s/m} \tag{b}$$

Use the Routh–Hurwitz stability criteria to assess the stability of the satellite as a major-axis spinner, a minor-axis spinner, and an intermediate-axis spinner.

The data in (a) are for a major-axis spinner. Substituting into Equations 8.1.92 and 8.1.93, we find

$$\begin{aligned} r_1 &= +1.188 \times 10^6 \text{ kg}^3\text{m}^4 \\ r_2 &= +18.44 \times 10^6 \text{ kg}^3\text{m}^4/\text{s} \\ r_3 &= +1.228 \times 10^9 \text{ kg}^3\text{m}^4/\text{s}^2 \\ r_4 &= +92\,820 \text{ kg}^3\text{m}^4/\text{s}^3 \\ r_5 &= +8.271 \times 10^9 \text{ kg}^3\text{m}^4/\text{s}^4 \end{aligned} \tag{c}$$

Since the rs are all positive, spin about the major axis is asymptotically stable. As we know from Section 8.1.3, without the damper the motion is neutrally stable.

For spin about the minor axis,

$$A = 500 \text{ kg} \cdot \text{m}^2 \quad B = 400 \text{ kg} \cdot \text{m}^2 \quad C = 300 \text{ kg} \cdot \text{m}^2 \tag{d}$$

For these moment of inertia values, we obtain

$$\begin{aligned} r_1 &= +1.980 \times 10^6 \text{ kg}^3\text{m}^4 \\ r_2 &= +30.74 \times 10^6 \text{ kg}^3\text{m}^4/\text{s} \\ \text{r}_3 &= +2.048 \times 10^9 \text{ kg}^3\text{m}^4/\text{s}^2 \\ r_4 &= -304\,490 \text{ kg}^3\text{m}^4/\text{s}^3 \\ r_5 &= +7.520 \times 10^9 \text{ kg}^3\text{m}^4/\text{s}^4 \end{aligned} \tag{e}$$

Since the rs are not all of the same sign, spin about the minor axis is not asymptotically stable. Recall that for the rigid satellite, such a motion was neutrally stable. Finally, for spin about the intermediate axis,

$$A = 300 \text{ kg} \cdot \text{m}^2 \quad B = 500 \text{ kg} \cdot \text{m}^2 \quad C = 400 \text{ kg} \cdot \text{m}^2 \tag{f}$$

We know this motion is unstable, even without the nutation damper, but doing the Routh–Hurwitz stability check anyway, we get

$$\begin{aligned} r_1 &= +1.485 \times 10^6 \text{ kg}^3\text{m}^4 \\ r_2 &= +22.92 \times 10^6 \text{ kg}^3\text{m}^4/\text{s} \\ \text{r}_3 &= +1.529 \times 10^9 \text{ kg}^3\text{m}^4/\text{s}^2 \\ r_4 &= -192\,800 \text{ kg}^3\text{m}^4/\text{s}^3 \\ r_5 &= -4.323 \times 10^9 \text{ kg}^3\text{m}^4/\text{s}^4 \end{aligned}$$

The motion, as we expected, is not stable.

8.1.6 Coning Maneuver

Like the use of nutation dampers, the coning maneuver is an example of the attitude control of spinning spacecraft. In this case, the angular momentum is changed by the use of on-board thrusters (small rockets) to apply pure torques.

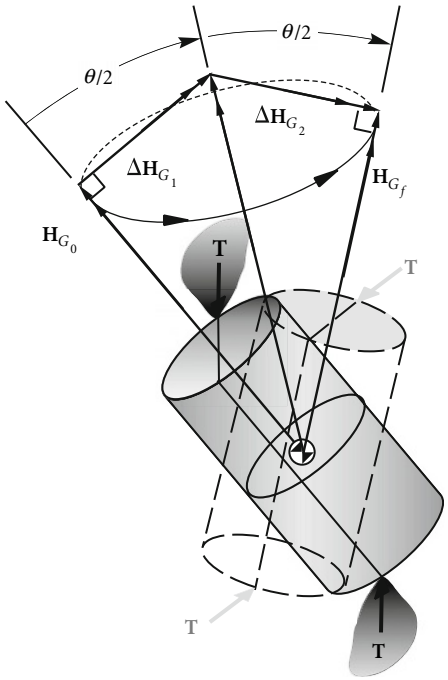


Figure 8.1-10 Impulsive coning maneuver.

Consider a satellite in pure spin with angular momentum $\mathbf{H}_{G_0}$. Suppose we wish to maintain the magnitude of the angular momentum but change its direction by rotating the spin axis through an angle θ, as illustrated in Figure 8.1-10. To change the angular momentum of the spacecraft requires applying an external moment,

$$\Delta\,\mathbf{H}_G = \int_0^{\Delta t} \mathbf{M}_G\; dt$$

Thrusters may be used to provide the external impulsive torque required to produce an angular momentum increment $\Delta\mathbf{H}_{G_1}$, normal to the spin axis. Since the spacecraft is spinning, this induces coning (precession) of the satellite about an axis at an angle $\theta/2$ to $\mathbf{H}_{G_0}$. The precession rate is given by Equation 8.1.23,

$$\omega_p = \frac{C}{A-C}\frac{\omega_s}{\cos\left(\frac{\theta}{2}\right)} \tag{8.1.94}$$

After precessing 180°, an angular momentum increment $\Delta\mathbf{H}_{G_2}$ normal to the spin axis and in the same direction relative to the spacecraft as the initial torque impulse, with $\|\,\Delta\mathbf{H}_{G_2}\,\| = \|\,\Delta\mathbf{H}_{G_1}\,\|$, stabilizes the spin vector in the desired direction. The time required for an angular reorientation θ using a single coning maneuver is found by simply dividing the precession angle, π radians, by the precession rate ω_p,

$$t_1 = \frac{\pi}{\omega_p} = \pi\frac{A-C}{C\omega_s}\cos\frac{\theta}{2} \tag{8.1.95}$$

Propellant expenditure is reflected in the magnitude of the individual angular momentum increments, in obvious analogy to delta-v calculations for orbital maneuvers. The total delta-H required for the single coning maneuver is therefore given by

$$\Delta H_{\text{total}} = \|\,\Delta\mathbf{H}_{G_1}\,\| + \|\,\Delta\mathbf{H}_{G_2}\,\| = 2\left(\|\,\mathbf{H}_{G_0}\,\|\tan\frac{\theta}{2}\right) \tag{8.1.96}$$

Figure 8.1-11 illustrates the fact that ΔH_{total} can be reduced by using a sequence of small coning maneuvers (small θ's) rather than one big θ. The large number of small ΔH's approximates a circular arc of radius $\|\mathbf{H}_{G_0}\|$, subtended by the angle θ. Therefore, approximately,

$$\Delta H_{\text{total}} = 2\left(\|\,\mathbf{H}_{G_0}\,\|\,\frac{\theta}{2}\right) = \|\,\mathbf{H}_{G_0}\|\theta \tag{8.1.97}$$

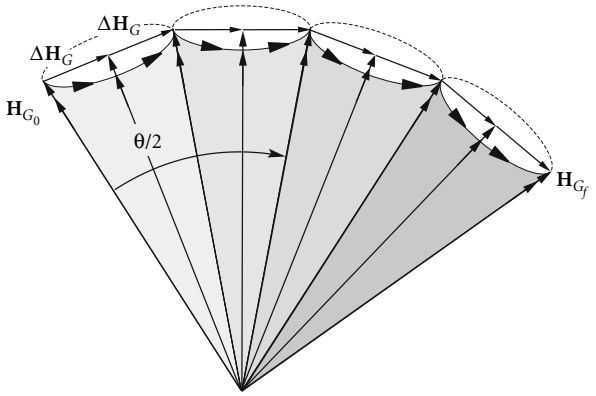


Figure 8.1-11 A sequence of small coning maneuvers.

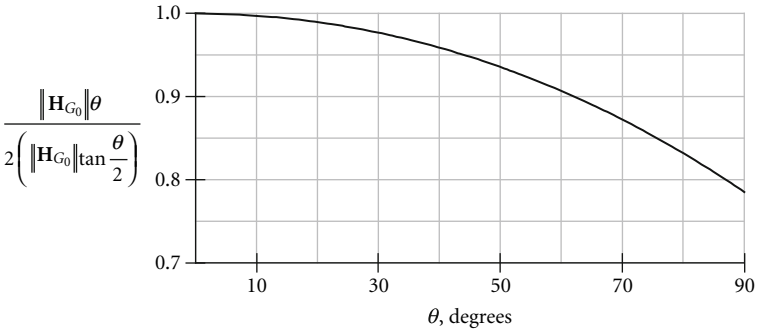


Figure 8.1-12 Ratio of delta-H for a sequence of small coning maneuvers to that for a single coning maneuver, as a function of the angle of swing of the spin axis.

This expression becomes more precise as the number of intermediate maneuvers increases. Figure 8.1-12 reveals the extent to which the multiple coning maneuver strategy reduces energy requirements. The difference is quite significant for large reorientation angles.

One of the prices to be paid for the reduced energy of the multiple coning maneuver is time. (The other is the risk involved in repeating the maneuver over and over again.) From Equation 8.1.95, the time required for n small-angle coning maneuvers through a total angle of θ is

$$t_n = n\pi\frac{A-C}{C\omega_s}\cos\frac{\theta}{2n} \tag{8.1.98}$$

The ratio of this to the time t_1 required for a single coning maneuver is

$$\frac{t_n}{t_1} = n\frac{\cos\dfrac{\theta}{2n}}{\cos\dfrac{\theta}{2}} \tag{8.1.99}$$

The time is directiy proportional to the number of intermediate coning maneuvers, as illustrated in Figure 8.1-13.

8.1.7 Attitude Control Thrusters

As mentioned above, thrusters are small jets mounted in pairs on a spacecraft to control its rotational motion about the center of mass. These thruster pairs may be mounted in principal planes (planes normal to the principal axes) passing through the center of mass. Figure 8.1-14 illustrates a pair of thrusters for producing a torque about the positive y axis. These would be accompanied by another pair of reaction motors pointing in the opposite directions to exert torque in the negative x direction. If the position vectors of the thrusters relative to the center of mass are $\mathbf{r}$ and $-\mathbf{r}$, and if $\mathbf{T}$ is their thrust, then the impulsive moment they exert during a brief time interval Δt is

$$\mathbf{M} = \mathbf{r}\times\mathbf{T}\Delta t + (-\mathbf{r})\times(-\mathbf{T}\Delta t) = 2\mathbf{r}\times\mathbf{T}\Delta t \tag{8.1.100}$$

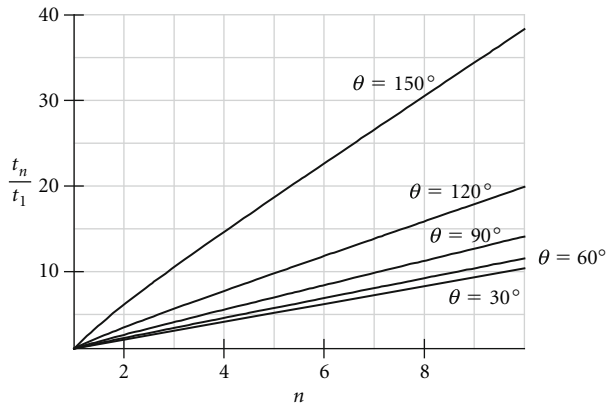


Figure 8.1-13 Time for a coning maneuver versus the number of intermediate steps.

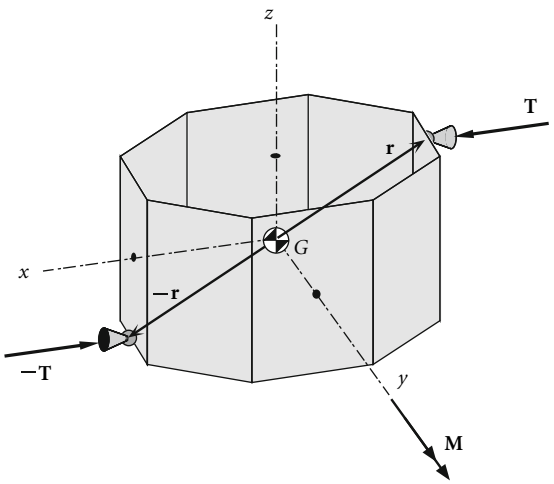


Figure 8.1-14 Pair of attitude control thrusters mounted in the *xz* plane of the principal body frame.

If the angular velocity was initially zero, then after the firing, according to Equation 8.1.31, the angular momentum becomes

$$\mathbf{H} = 2\mathbf{r} \times \mathbf{T}\Delta t \tag{8.1.101}$$

For $\mathbf{H}$ in the principal x direction, as in the figure, the corresponding angular velocity acquired by the vehicle is, from Equation 8.1.67,

$$\omega_y = \frac{\|\mathbf{H}\|}{B} \tag{8.1.102}$$

Example 8.1.7

A spacecraft of mass m and with the dimensions shown in Figure 8.1-15 is spinning without precession at the rate ω_0 about the z axis of the principal body frame. At the instant shown in part (a) of the figure, the spacecraft initiates a coning maneuver to swing its spin axis through 90°, so that at the end of the maneuver the vehicle is oriented as illustrated in Figure 8.1-15(b). Calculate the total delta-H required, and compare it with that required for the same reorientation without coning. Motion is to be controlled exclusively by the pairs of attitude thrusters shown, all of which have identical thrust T.

The moments of inertia about the principal body axes are

$$A = B = \frac{1}{12}m\left[w^2 + \left(\frac{w}{3}\right)^2\right] = \frac{5}{54}mw^2 \qquad C = \frac{1}{12}m(w^2 + w^2) = \frac{1}{6}mw^2$$

The initial angular momentum $\mathbf{H}_{G_1}$ points in the spin direction, along the positive z axis of the body frame,

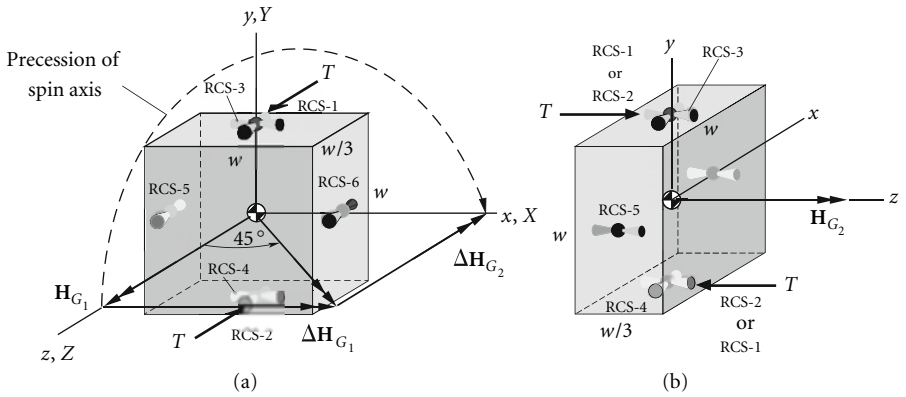


Figure 8.1-15 (a) Initial orientation of spinning spacecraft. (b) Final configuration, with spin axis rotated 90°.

$$\mathbf{H}_{G_1} = C\omega_z\hat{\mathbf{k}} = \frac{1}{6}mw^2\omega_0\hat{\mathbf{k}}$$

We can presume that in the initial orientation, the body frame happens to coincide instantaneously with inertial frame *XYZ*. The coning motion is initiated by briefly firing the pair of thrusters RCS-1 and RCS-2, aligned with the body z axis and lying in the yz plane. The impulsive torque will cause a change $\Delta\mathbf{H}_{G_1}$ in angular momentum directed normal to the plane of the thrusters, in the positive body x direction. The resultant angular momentum vector must lie at 45° to the x and z axes, bisecting the angle between the initial and final angular momenta. Thus,

$$\|\Delta\mathbf{H}_{G_1}\| = \|\mathbf{H}_{G_1}\| \tan 45^\circ = \frac{1}{6}mw^2\omega_0$$

After the coning is underway, the body axes of course move away from the *XYZ* frame. Since the spacecraft is oblate ($C > A$), the precession of the spin axis will be opposite to the spin direction, as indicated in Figure 8.1-15. When the spin axis, after 180° of precession, lines up with the x axis the thrusters must fire again for the same duration as before so as to produce the angular momentum change $\Delta\mathbf{H}_{G_2}$, equal in magnitude but perpendicular to $\Delta\mathbf{H}_{G_1}$, so that

$$\mathbf{H}_{G_1} + \Delta\mathbf{H}_{G_1} + \Delta\mathbf{H}_{G_2} = \mathbf{H}_{G_2}$$

where

$$\mathbf{H}_{G_2} = \|\mathbf{H}_{G_1}\|\,\hat{\mathbf{I}} = \frac{1}{6}mw^2\omega_0\hat{\mathbf{k}}$$

For this to work, the plane of thrusters RCS-1 and RCS-2 – the yz plane – must be parallel to the *XY* plane when they fire, as illustrated in Figure 8.1-15(b). Since the thrusters can fire fore or aft, it does not matter which of them ends up on top or bottom. The vehicle must therefore spin through an integral number n of half rotations while it precesses to the desired orientation. That is, the total spin angle ψ between the initial and final configurations is

$$\psi = n\pi = \omega_s t \tag{a}$$

where ω_s is the spin rate and t is the time for the proper final configuration to be achieved. In the meantime, the precession angle ϕ must be π or 3π or 5π, or, in general,

$$\phi = (2m-1)\pi = \omega_p t \tag{b}$$

where m is an integer and t is, of course, the same as that in (a). Eliminating t from both (a) and (b) yields

$$n\pi = (2m-1)\pi\frac{\omega_s}{\omega_p}$$

Substituting Equation 8.1.94, with $\theta = \pi/2$, gives

$$n = (1-2m)\frac{4}{9}\frac{1}{\sqrt{2}} \tag{c}$$

Obviously, this equation cannot be valid if both m and n are integers. However, by tabulating n as a function of m we find that when $m = 18$, $n = -10.999$. The minus sign simply reminds us that spin and precession are in opposite directions. Thus, the eighteenth time that the spin axis lines up with the x axis the thrusters may be fired to almost

perfectly align the angular momentum vector with the body z axis. The slight misalignment due to the fact that $||n||$ is not precisely 11 would probably occur in reality anyway. Passive or active nutation damping can drive this deviation to zero.

Since $||\mathrm{H}_{G_1}|| = ||\mathrm{H}_{G_2}||$, we conclude that

$$\Delta H_{\text{total}} = 2\left(\frac{1}{6}mw^2\omega_0\right) = \frac{2}{3}mw^2\omega_0 \tag{d}$$

An obvious alternative to the coning maneuver is to use thrusters RCS-3 and 4 to despin the craft completely, thrusters RCS-5 and 6 to initiate roll around the y axis and stop it after 90°, and then RCS-3 and 4 to respin the spacecraft to ω_0 around the z axis. The combined delta-H for the first and last steps equals that of (d). Additional fuel expenditure is required to start and stop the roll around the y axis. Hence, the coning maneuver is more fuel efficient.

8.1.8 Yo-Yo Despin Mechanism

A simple, inexpensive way to despin an axisymmetric satellite is to deploy small masses attached to cords wound around the girth of the satellite near the transverse plane through the center of mass. As the masses unwrap in the direction of the satellite's angular velocity, they exert centrifugal force through the cords on the periphery of the satellite, creating a moment opposite to the spin direction, thereby slowing down the rotational motion. The cord forces are internal to the system of satellite plus weights, so as the strings unwind, the total angular momentum must remain constant. Since the total moment of inertia increases as the yo-yo masses spiral further away, the angular velocity must drop. Not only angular momentum but also rotational kinetic energy is conserved during this process. Yo-yo despin devices were introduced early in unmanned space flight (e.g., 1959 Transit 1-A) and continue to be used today (e.g., 1996 Mars Pathfinder, 1998 Mars Climate Orbiter, 1999 Mars Polar Lander, 2003 Mars Exploration Rover).

We will use the conservation of energy and momentum to determine the length of cord required to reduce the satellite's angular velocity a specified amount. To maintain the position of the center of mass, two identical yo-yo masses are wound around the spacecraft in a symmetrical fashion, as illustrated in Figure 8.1-16. Both masses are released simultaneously by explosive bolts and unwrap in the manner shown (for only one of the weights) in the figure. In so doing, the point of tangency T moves around the circumference towards the split hinge device where the cord is attached to the spacecraft. When T and T' reach the hinges H and H', the cords automatically separate from the spacecraft.

Let each yo-yo weight have mass $m/2$. By symmetry, we need to track only one of the masses, to which we can ascribe the total mass m. Let the xyz system be a body frame rigidly attached to the satellite, as shown in Figure 8.1-16. As usual, the z axis lies in the spin direction, pointing out of the page. The x axis is directed from the center of mass of the system through the initial position of the yo-yo mass. The satellite and the yo-yo masses, prior to release, are rotating as a single rigid body with angular velocity $\boldsymbol{\omega}_0 = \omega_0\hat{\mathbf{k}}$. The moment of inertia of the satellite, excluding the yo-yo mass, is C, so that the angular momentum of the satellite by itself is $C\omega_0$. The concentrated yo-yo

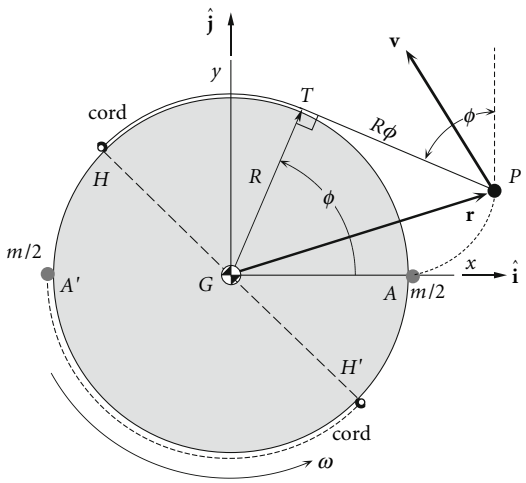


Figure 8.1-16 Two identical string and mass systems wrapped symmetrically around the periphery of an axisymmetric satellite. For simplicity, only one is shown being deployed.

masses are fastened a distance R from the spin axis, so their total moment of inertia is mR^2. Therefore, the initial angular momentum of the satellite plus yo-yo system is

$$H_{G_0} = C\omega_0 + mR^2\omega_0$$

It will be convenient to write this as

$$H_{G_0} = KmR^2\omega_0 \tag{8.1.103}$$

where the non-dimensional factor K is defined as

$$K = 1 + \frac{C}{mR^2} \tag{8.1.104}$$

$\sqrt{KR}$ is the initial radius of gyration of the system. The initial rotational kinetic energy of the system, before the masses are released, is

$$T_0 = \frac{1}{2}C\omega_0^2 + \frac{1}{2}mR^2\omega_0^2 = \frac{1}{2}KmR^2\omega_0^2 \tag{8.1.105}$$

At any state between the release of the weights and the release of the cords at the hinges, the velocity of the yo-yo mass must be found in order to compute the new angular momentum and kinetic energy. Observe that when the string has unwrapped an angle ϕ, the free length of string (between the point of tangency T and the yo-yo mass P) is $R\phi$. From the geometry shown in Figure 8.1-16, the position vector of the mass relative to the body frame is seen to be

$$\mathbf{r} = \overbrace{(R\cos\phi\hat{\mathbf{i}} + R\sin\phi\hat{\mathbf{j}})}^{\mathbf{r}_{T/G}} + \overbrace{(R\phi\sin\phi\hat{\mathbf{i}} - R\phi\cos\phi\hat{\mathbf{j}})}^{\mathbf{r}_{P/T}} = (R\cos\phi + R\phi\sin\phi)\hat{\mathbf{i}} + (R\sin\phi - R\phi\cos\phi)\hat{\mathbf{j}} \tag{8.1.106}$$

Since $\mathbf{r}$ is measured in the moving reference, the absolute velocity $\mathbf{v}$ of the yo-yo mass is

$$\mathbf{v} = \left(\frac{d\mathbf{r}}{dt}\right)_{\text{rel}} + \boldsymbol{\Omega}\times\mathbf{r} \tag{8.1.107}$$

where $\boldsymbol{\Omega}$ is the angular velocity of the xyz axes, which, of course, is the angular velocity $\boldsymbol{\omega}$ of the satellite at that instant,

$$\boldsymbol{\Omega} = \boldsymbol{\omega} \tag{8.1.108}$$

To calculate $(d\mathbf{r}/dt)_{\text{rel}}$, we hold $\hat{\mathbf{i}}$ and $\hat{\mathbf{j}}$ constant in Equation 8.1.106, obtaining

$$\begin{aligned}\left(\frac{d\mathbf{r}}{dt}\right)_{\text{rel}} &= (-R\dot{\phi}\sin\phi + R\dot{\phi}\sin\phi + R\dot{\phi}\cos\phi)\hat{\mathbf{i}} + (R\dot{\phi}\cos\phi - R\dot{\phi}\cos\phi + R\dot{\phi}\sin\phi)\hat{\mathbf{j}}\\ &= R\dot{\phi}\cos\phi\hat{\mathbf{i}} + R\dot{\phi}\sin\phi\hat{\mathbf{j}}\end{aligned}$$

Thus

$$\mathbf{v} = R\dot{\phi}\cos\phi\hat{\mathbf{i}} + R\dot{\phi}\sin\phi\hat{\mathbf{j}} + \begin{vmatrix}\hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}}\\ 0 & 0 & \omega\\ R\cos\phi + R\phi\sin\phi & R\sin\phi - R\phi\cos\phi & 0\end{vmatrix}$$

or

$$\mathbf{v} = [R\phi(\omega + \dot{\phi})\cos\phi - R\omega\sin\phi]\hat{\mathbf{i}} + [R\omega\cos\phi + R\phi(\omega + \dot{\phi})\sin\phi]\hat{\mathbf{j}} \tag{8.1.109}$$

From this we find the speed of the yo-yo weights,

$$v = \sqrt{\mathbf{v}\cdot\mathbf{v}} = R\sqrt{\omega^2 + (\omega + \dot{\phi})^2\phi^2} \tag{8.1.110}$$

The angular momentum of the satellite plus the weights at an intermediate stage of the despin process is

$$\begin{aligned}\mathbf{H}_G &= C\omega\hat{\mathbf{k}} + \mathbf{r}\times m\mathbf{v}\\ &= C\omega\hat{\mathbf{k}} + m\begin{vmatrix}\hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}}\\ R\cos\phi + R\phi\sin\phi & R\sin\phi - R\phi\cos\phi & \omega\\ R\phi(\omega + \dot{\phi})\cos\phi - R\omega\sin\phi & R\omega\cos\phi + R\phi(\omega + \dot{\phi})\sin\phi & 0\end{vmatrix}\end{aligned}$$

Carrying out the cross product, combining terms and simplifying, leads to

$$H_G = C\omega + mR^2[\omega + (\omega + \dot{\phi})\phi^2]$$

which, using Equation 8.1.104, can be written as

$$H_G = mR^2[K\omega + (\omega + \dot{\phi})\phi^2] \tag{8.1.111}$$

The kinetic energy of the satellite plus the yo-yo mass is

$$T = \frac{1}{2}C\omega^2 + \frac{1}{2}mv^2$$

Substituting the speed from Equation 8.1.110 and making use again of Equation 8.1.104, we find

$$T = \frac{1}{2}mR^2[K\omega^2 + (\omega + \dot{\phi})^2\phi^2] \tag{8.1.112}$$

By the conservation of angular momentum, $H_G = H_{G_0}$, we obtain from Equations 8.1.103 and 8.1.111,

$$mR^2[K\omega + (\omega + \dot{\phi})\phi^2] = KmR^2\omega_0$$

which we can write as

$$K(\omega_0 - \omega) = (\omega + \dot{\phi})\phi^2 \tag{8.1.113}$$

Equations 8.1.105 and 8.1.112 and the conservation of kinetic energy, $T = T_0$, combine to yield

$$\frac{1}{2}mR^2[K\omega^2 + (\omega + \dot{\phi})^2\phi^2] = \frac{1}{2}KmR^2\omega_0^2$$

or

$$K(\omega_0^2 - \omega^2) = (\omega + \dot{\phi})^2\phi^2 \tag{8.1.114}$$

Since $\omega_0^2 - \omega^2 = (\omega_0 - \omega)(\omega_0 + \omega)$, this can be written as

$$K(\omega_0 - \omega)(\omega_0 + \omega) = (\omega + \dot{\phi})^2\phi^2$$

Replacing the factor $K(\omega_0 - \omega)$ on the left using Equation 8.1.113 yields

$$(\omega + \dot{\phi})\phi^2(\omega_0 + \omega) = (\omega + \dot{\phi})^2\phi^2$$

After canceling terms, we find $\omega_0 + \omega = \omega + \dot{\phi}$, or, simply

$$\dot{\phi} = \omega_0 \tag{8.1.115}$$

In other words, the cord unwinds at a constant rate (relative to the satellite), equal to the satellite's initial angular velocity. Thus at any time t after the release of the weights,

$$\phi = \omega_0 t \tag{8.1.116}$$

By substituting Equation 8.1.115 into Equation 8.1.113,

$$K(\omega_0 - \omega) = (\omega + \omega_0)\phi^2$$

we find that

$$\phi = \sqrt{K\frac{\omega_0 - \omega}{\omega_0 + \omega}} \quad \text{Partial despin.} \tag{8.1.117}$$

Recall that the unwrapped length l of the cord is $R\phi$, which means

$$l = R\sqrt{K\frac{\omega_0 - \omega}{\omega_0 + \omega}} \quad \text{Partial despin.} \tag{8.1.118}$$

We use Equation 8.1.118 to find the length of cord required to despin the spacecraft from ω_0 to ω. To remove all of the spin ($\omega = 0$),

$$\phi = \sqrt{K} \quad \Rightarrow \quad l = R\sqrt{K} \qquad \text{Complete despin.} \tag{8.1.119}$$

Surprisingly, the length of cord required to reduce the angular velocity to zero is independent of the initial angular velocity.

We can solve Equation 8.1.117 for ω in terms of ϕ,

$$\omega = \left(\frac{2K}{K+\phi^2} - 1\right)\omega_0 \tag{8.1.120}$$

By means of Equation 8.1.116, this becomes an expression for the angular velocity as a function of time:

$$\omega = \left(\frac{2K}{K+\omega_0^2 t^2} - 1\right)\omega_0 \tag{8.1.121}$$

Alternatively, since $\phi = l/R$, Equation 8.1.120 yields the angular velocity as a function of cord length,

$$\omega = \left(\frac{2KR^2}{KR^2 + l^2} - 1\right)\omega_0 \tag{8.1.122}$$

Differentiating ω with respect to time in Equation 8.1.121 gives us an expression for the angular acceleration of the spacecraft,

$$\alpha = \frac{d\omega}{dt} = -\frac{4K\omega_0^3 t}{(K+\omega_0^2 t^2)^2} \tag{8.1.123}$$

whereas integrating ω with respect to time yields the angle rotated by the satellite since release of the yo-yo mass,

$$\theta = 2\sqrt{K}\tan^{-1}\frac{\omega_0 t}{\sqrt{K}} - \omega_0 t = 2\sqrt{K}\tan^{-1}\frac{\phi}{\sqrt{K}} - \phi \tag{8.1.124}$$

For complete despin, this expression, together with Equation 8.1.119, yields

$$\theta = \sqrt{K}\left(\frac{\pi}{2} - 1\right) \tag{8.1.125}$$

From the free-body diagram of the spacecraft shown in Figure 8.1-17, it is clear that the torque exerted by the yo-yo weights is

$$M_{G_z} = -2RN \tag{8.1.126}$$

where N is the tension in the cord. From Euler's equations of motion, Equation 8.1.72,

$$M_{G_z} = C\alpha \tag{8.1.127}$$

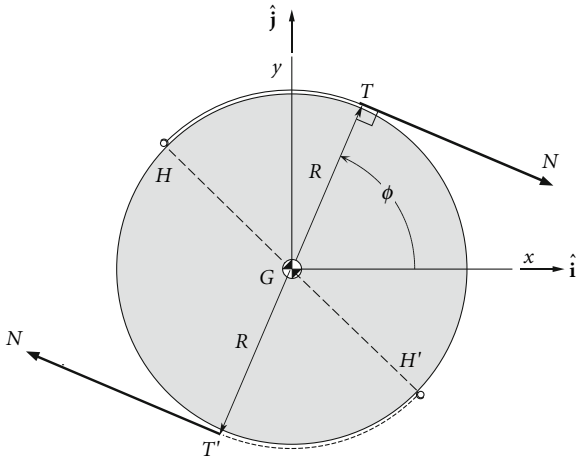


Figure 8.1-17 Free-body diagram of the satellite during the despin process.

Combining Equations 8.1.123, 8.1.126 and 8.1.127 leads to a formula for the tension in the yo-yo cables,

$$N = \frac{C}{R}\frac{2K\omega_0^3 t}{(K+\omega_0^2 t^2)^2} = \frac{C\omega_0^2}{R}\frac{2K\phi}{(K+\phi^2)^2} \tag{8.1.128}$$

Radial release

Finally, we note that instead of releasing the yo-yo masses when the cables are tangent at the split hinges (H and H'), they can be forced to pivot about the hinge and released when the string is directed radially outward, as illustrated in Figure 8.1-18. The above analysis must be then extended to include the pivoting of the cord around the hinges. It turns out that in this case, the length of the cord as a function of the final angular velocity is

$$l = R\left(\sqrt{\frac{[(\omega_0-\omega)K+\omega]^2}{(\omega_0^2-\omega^2)K+\omega^2}} - 1\right) \quad \text{Partial despin, radial release.} \tag{8.1.129}$$

so that for $\omega = 0$,

$$l = R(\sqrt{K} - 1) \quad \text{Complete despin, radial release.} \tag{8.1.130}$$

Example 8.1.8

A satellite is to be completely despun using a two-mass yo-yo device with tangential release. Assume the spin axis of moment of inertia of the satellite is $C = 200$ kg · m^2 and the initial spin rate is $\omega_0 = 5$ rad/s. The total yo-yo mass is 4 kg, and the radius of the spacecraft is 1 meter. Find (a) the required cord length l; (b) the time t to despin; (c) the maximum tension in the yo-yo cables; (d) the speed of the masses at release; (e) the angle rotated by the satellite during the despin; (f) the cord length required for radial release.

(a) From Equation 8.1.104,

$$K = 1 + \frac{C}{mR^2} = 1 + \frac{200}{4\cdot 1^2} = 51 \tag{a}$$

From Equation 8.1.118 it follows that the cord length required for complete despin is

$$l = R\sqrt{K} = 1\cdot\sqrt{51} = \underline{7.1414 \text{ m}} \tag{b}$$

(b) The time for complete despin is obtained from Equations 8.1.116 and 8.1.118,

$$\omega_0 t = \sqrt{K} \quad \Rightarrow \quad t = \frac{\sqrt{K}}{\omega_0} = \frac{\sqrt{51}}{5} = \underline{1.4283 \text{ s}}$$

(c) A graph of Equation 8.1.128 is shown in Figure 8.1-19. The maximum tension is 455 N, which occurs at 0.825 s.

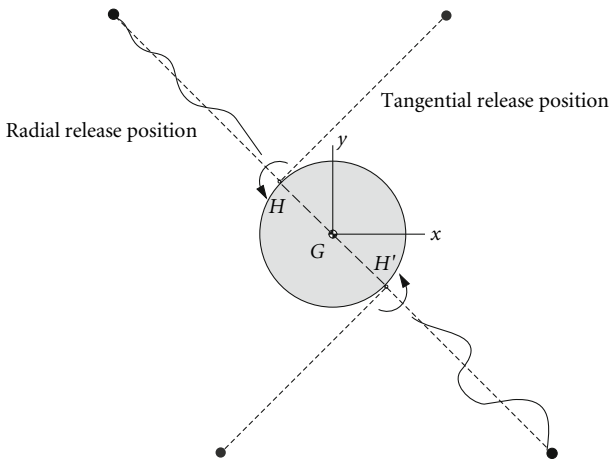


Figure 8.1-18 Radial versus tangential release of yo-yo masses.

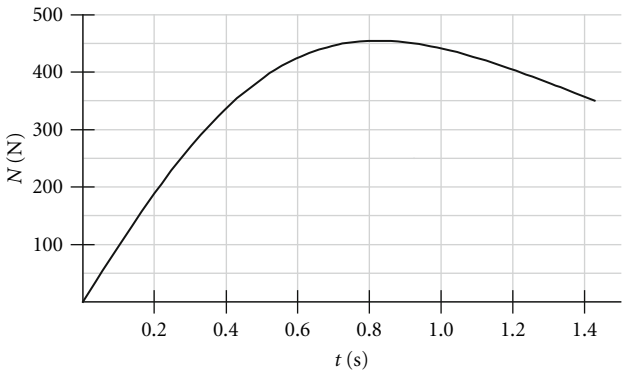


Figure 8.1-19 Variation of cable tension N up to point of release.

(d) From Equation 8.1.110, the speed of the yo-yo masses is

$$v = R\sqrt{\omega^2 + (\omega + \dot{\phi})^2\phi^2}$$

According to Equation 8.1.115, $\dot{\phi} = \omega_0$ and at the time of release ($\omega = 0$) Equation 8.1.118 states that $\phi = \sqrt{K}$. Thus

$$v = R\sqrt{\omega^2 + (\omega + \omega_0)^2\sqrt{K}^2} = 1\cdot\sqrt{0^2 + (0+5)^2\sqrt{51}^2} = \underline{35.71 \text{ m/s}}$$

(e) The angle through which the satellite rotates before coming to rotational rest is given by Equation 8.1.124,

$$\theta = \sqrt{K}\left(\frac{\pi}{2} - 1\right) = \sqrt{51}\left(\frac{\pi}{2} - 1\right) = \underline{4.076 \text{ rad } (233.5^\circ)}$$

(f) Allowing the cord to detach radially reduces the cord length required for complete despin from 7.141 m to

$$l = R(\sqrt{K} - 1) = 1\cdot(\sqrt{51} - 1) = \underline{6.141 \text{ m}}$$

8.1.9 Gyroscopic Attitude Control

Momentum exchange systems ('gyros') are used to control the attitude of a spacecraft without throwing consumable mass overboard, as occurs with the use of thruster jets. A momentum exchange system is illustrated schematically in Figure 8.1-20. n flywheels, labeled 1, 2, 3, etc., are attached to the body of the spacecraft at various locations. The mass of flywheel i is m_i. The mass of the body of the spacecraft is m_0.

The total mass of the entire system – the 'vehicle' – is m,

$$m = m_0 + \sum_{i=1}^{n} m_i$$

The vehicle's center of mass is G, through which pass the three axes xyz of the vehicle's body-fixed frame. The center of mass G_i of each flywheel is connected rigidly to the spacecraft, but the wheel, driven by electric motors, rotates more or less independently, depending on the type of gyro. The body of the spacecraft has an angular velocity $\boldsymbol{\omega}$. The angular velocity of the ith flywheel is $\boldsymbol{\omega}_i$, and differs from that of the body of the spacecraft unless the gyro is 'caged'. A caged gyro has no spin relative to the spacecraft, in which case $\boldsymbol{\omega}_i = \boldsymbol{\omega}$.

The angular momentum of the entire system about the vehicle's center of mass G is the sum of the angular momenta of the individual components of the system,

$$\mathbf{H}_G = \mathbf{H}_G^{(v)} + \mathbf{H}^{(w)} \tag{8.1.131}$$

$\mathbf{H}_G^{(v)}$ is the total angular momentum of the rigid body comprising the spacecraft and all of the flywheel masses concentrated at their centers of mass G_i. That system has the common vehicle angular velocity $\boldsymbol{\omega}$, which means that

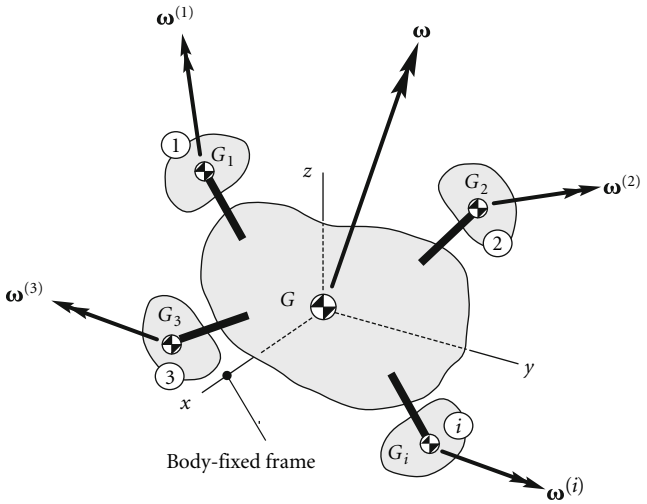


Figure 8.1-20 Several attitude control flywheels, each with their own angular velocity, attached to the body of a spacecraft.

$$\{\mathbf{H}_G^{(v)}\} = [\mathbf{I}_G^{(v)}]\{\boldsymbol{\omega}\} \tag{8.1.132}$$

where $[\mathbf{I}_G^{(v)}]$ is the moment of inertia found by adding the moments of inertia of all the concentrated flywheel masses about G to that of the body of the spacecraft. On the other hand, $\mathbf{H}^{(w)}$ is the net angular momentum of the n flywheels about each of their individual centers of mass,

$$\mathbf{H}^{(w)} = \sum_{i=1}^{n} \mathbf{H}_{G_i}^{(i)} \tag{8.1.133}$$

$\mathrm{H}_{G_i}^{(i)}$, the angular momentum of flywheel i about its center of mass G_i, is obtained as

$$\{\mathbf{H}_{G_i}^{(i)}\} = [\mathbf{I}_{G_i}^{(i)}]\{\boldsymbol{\omega}^{(i)}\} \tag{8.1.134}$$

$[\mathbf{I}_{G_i}^{(i)}]$ is the moment of inertia of flywheel i about G_i, relative to axes which are parallel to the body-fixed xyz axes. The mass distribution reflected in $[\mathbf{I}_G^{(v)}]$ is fixed relative to the body frame, which means this matrix does not vary with time. On the other hand, since a momentum wheel might be one that pivots on gimbals relative to the body frame, the inertia tensor $[\mathbf{I}_{G_i}^{(i)}]$ may be time dependent.

Using Equation 8.1.131 yields the equations of rotational motion of the gyro stabilized spacecraft,

$$\mathbf{M}_{G_{\text{net}}} = \dot{\mathbf{H}}_G^{(v)} + \dot{\mathbf{H}}^{(w)} \tag{8.1.135}$$

The angular momenta are computed in the non-inertial body-fixed frame, thus computing accordingly the time derivatives on the right-hand side of Equation 8.1.135, we get

$$\mathbf{M}_{G_{\text{net}}} = \left[\left(\frac{d\mathbf{H}_G^{(v)}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega} \times \mathbf{H}_G^{(v)}\right] + \left[\left(\frac{d\mathbf{H}^{(w)}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega} \times \mathbf{H}^{(w)}\right] \tag{8.1.136}$$

For torque-free motion, $\mathbf{M}_{G_{\text{net}}} = 0$, in which case we have the conservation of angular momentum about the vehicle center of mass,

$$\mathbf{H}_G^{(v)} + \mathbf{H}^{(w)} = \textbf{constant} \tag{8.1.137}$$

Example 8.1.9

Use Equation 8.1.136 to obtain the equations of motion of a torque-free, axisymmetric, dual-spin satellite, such as the one shown in Figure 8.1-21.

In the dual-spin satellite, we may arbitrarily choose the rotor as the body of the vehicle, to which the body frame is attached. The coaxial platform will play the role of the single reaction wheel. The center of mass G of the satellite lies

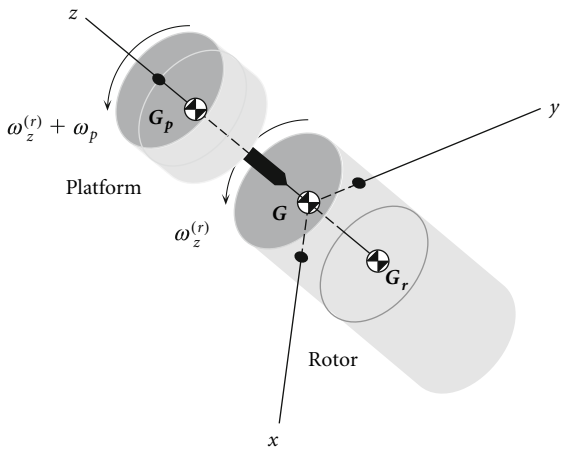


Figure 8.1-21 Dual-spin spacecraft.

on the axis of rotational symmetry (the z axis), between the center of mass of the rotor (G_r) and that of the platform (G_p). For this torque-free system, Equation 8.1.136 becomes

$$\left.\frac{d\mathbf{H}_G^{(v)}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega}^{(r)} \times \mathbf{H}_G^{(v)} + \left.\frac{d\mathbf{H}_{G_p}^{(p)}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega}^{(r)} \times \mathbf{H}_{G_p}^{(p)} = \mathbf{0} \tag{a}$$

in which r signifies the rotor and p the platform.

The vehicle angular momentum about G is that of the rotor plus that of the platform center of mass,

$$\{\mathbf{H}_G^{(v)}\} = [\mathbf{I}_G^{(r)}]\{\boldsymbol{\omega}^{(r)}\} + [\mathbf{I}_{m_G}^{(p)}]\{\boldsymbol{\omega}^{(r)}\} = ([\mathbf{I}_G^{(r)}] + [\mathbf{I}_{m_G}^{(p)}])\{\boldsymbol{\omega}^{(r)}\} \tag{b}$$

$[\mathbf{I}_{m_G}^{(p)}]$ is the moment of inertia tensor of the concentrated mass of the platform about the system center of mass. The components of $[\mathbf{I}_G^{(r)}]$ and $[\mathbf{I}_{m_G}^{(p)}]$ are constants, so from (b) we obtain

$$\left.\frac{d\{\mathbf{H}_G^{(v)}\}}{dt}\right)_{\text{rel}} = ([\mathbf{I}_G^{(r)}] + [\mathbf{I}_{m_G}^{(p)}])\{\dot{\boldsymbol{\omega}}^{(r)}\} \tag{c}$$

The angular momentum of the *platform* about its own center of mass is

$$\{\mathbf{H}_{G_p}^{(p)}\} = [\mathbf{I}_{G_p}^{(p)}]\{\boldsymbol{\omega}^{(p)}\} \tag{d}$$

For both the platform and the rotor, the z axis is an axis of rotational symmetry. Thus, even though the platform is not stationary in xyz, the moment of inertia matrix $[\mathbf{I}_{G_p}^{(p)}]$ is not time dependent. It follows that

$$\left.\frac{d\{\mathbf{H}_{G_p}^{(p)}\}}{dt}\right)_{\text{rel}} = [\mathbf{I}_{G_p}^{(p)}]\{\dot{\boldsymbol{\omega}}^{(p)}\} \tag{e}$$

Using (b) through (e), we can write the equation of motion (a) as

$$\begin{aligned}&([\mathbf{I}_G^{(r)}] + [\mathbf{I}_{m_G}^{(p)}])\{\dot{\boldsymbol{\omega}}^{(r)}\} + \{\boldsymbol{\omega}^{(r)}\} \times ([\mathbf{I}_G^{(r)}] + [\mathbf{I}_{m_G}^{(p)}])\{\boldsymbol{\omega}^{(r)}\} \\ &\quad + [\mathbf{I}_{G_p}^{(p)}]\{\dot{\boldsymbol{\omega}}^{(p)}\} + \{\boldsymbol{\omega}^{(r)}\} \times [\mathbf{I}_{G_p}^{(p)}]\{\boldsymbol{\omega}^{(p)}\} = \{\mathbf{0}\}\end{aligned} \tag{f}$$

The angular velocity $\boldsymbol{\omega}^{(p)}$ of the platform is that of the rotor, $\boldsymbol{\omega}^{(r)}$ plus the angular velocity of the platform relative to the rotor, $\boldsymbol{\omega}_{\text{rel}}^{(p)}$. Hence, we may replace $\{\boldsymbol{\omega}^{(p)}\}$ with $\{\boldsymbol{\omega}^{(r)}\} + \{\boldsymbol{\omega}_{\text{rel}}^{(p)}\}$, so that, after a little rearrangement, (f) becomes

$$\begin{aligned}&([\mathbf{I}_G^{(r)}] + [\mathbf{I}_G^{(p)}])\{\dot{\boldsymbol{\omega}}^{(r)}\} + \{\boldsymbol{\omega}^{(r)}\} \times ([\mathbf{I}_G^{(r)}] + [\mathbf{I}_G^{(p)}])\{\boldsymbol{\omega}^{(r)}\} \\ &\quad + [\mathbf{I}_{G_p}^{(p)}]\{\dot{\boldsymbol{\omega}}_{\text{rel}}^{(p)}\} + \{\boldsymbol{\omega}^{(r)}\} \times [\mathbf{I}_{G_p}^{(p)}]\{\boldsymbol{\omega}_{\text{rel}}^{(p)}\} = \{\mathbf{0}\}\end{aligned} \tag{g}$$

in which

$$[\mathbf{I}_G^{(p)}] = [\mathbf{I}_{m_G}^{(p)}] + [\mathbf{I}_{G_p}^{(p)}] \quad \text{(Parallel axis formula)} \tag{h}$$

The components of the matrices and vectors in (g) relative to the principal xyz body frame axes are

$$[\mathbf{I}_G^{(r)}] = \begin{bmatrix} A_r & 0 & 0 \\ 0 & A_r & 0 \\ 0 & 0 & C_r \end{bmatrix} \quad [\mathbf{I}_G^{(p)}] = \begin{bmatrix} A_p & 0 & 0 \\ 0 & A_p & 0 \\ 0 & 0 & C_p \end{bmatrix} \quad [\mathbf{I}_{G_p}^{(p)}] = \begin{bmatrix} \overline{A}_p & 0 & 0 \\ 0 & \overline{A}_p & 0 \\ 0 & 0 & \overline{C}_p \end{bmatrix} \tag{i}$$

and

$$\{\boldsymbol{\omega}^{(r)}\} = \begin{Bmatrix} \omega_x^{(r)} \\ \omega_y^{(r)} \\ \omega_z^{(r)} \end{Bmatrix} \quad \{\dot{\boldsymbol{\omega}}^{(r)}\} = \begin{Bmatrix} \dot{\omega}_x^{(r)} \\ \dot{\omega}_y^{(r)} \\ \dot{\omega}_z^{(r)} \end{Bmatrix} \quad \{\dot{\boldsymbol{\omega}}_{\text{rel}}^{(p)}\} = \begin{Bmatrix} 0 \\ 0 \\ \omega_p \end{Bmatrix} \quad \{\dot{\boldsymbol{\omega}}_{\text{rel}}^{(p)}\} = \begin{Bmatrix} 0 \\ 0 \\ \dot{\omega}_p \end{Bmatrix} \tag{j}$$

A_r, C_r, A_p and C_p are the rotor and platform principal moments of inertia about the vehicle center of mass G, whereas $\overline{A}_p$ is the moment of inertia of the platform about its own center of mass. We also used the fact that $\overline{C}_p = C_p$, which of course is due to the fact that G and G_p both lie on the z axis. This notation is nearly identical to that employed in our consideration of the stability of dual-spin satellites in Section 8.1.4 (wherein $\omega_r = \omega_z^{(r)}$ and $\boldsymbol{\omega}_\perp = \omega_x^{(r)}\hat{\mathbf{i}} + \omega_y^{(r)}\hat{\mathbf{j}}$). Substituting (i) and (j) into each of the four terms in (g), we get

$$([\mathbf{I}_G^{(r)}] + [\mathbf{I}_G^{(p)}])\{\dot{\boldsymbol{\omega}}^{(r)}\} = \begin{bmatrix} A_r + A_p & 0 & 0 \\ 0 & A_r + A_p & 0 \\ 0 & 0 & C_r + C_p \end{bmatrix} \begin{Bmatrix} \dot{\omega}_x^{(r)} \\ \dot{\omega}_y^{(r)} \\ \dot{\omega}_z^{(r)} \end{Bmatrix} = \begin{Bmatrix} (A_r + A_p)\dot{\omega}_x^{(r)} \\ (A_r + A_p)\dot{\omega}_y^{(r)} \\ (C_r + C_p)\dot{\omega}_z^{(r)} \end{Bmatrix} \tag{k}$$

$$\{\boldsymbol{\omega}^{(r)}\} \times ([\mathbf{I}_G^{(r)}] + [\mathbf{I}_G^{(p)}])\{\boldsymbol{\omega}^{(r)}\} = \begin{Bmatrix} \omega_x^{(r)} \\ \omega_y^{(r)} \\ \omega_z^{(r)} \end{Bmatrix} \times \begin{Bmatrix} (A_r + A_p)\omega_x^{(r)} \\ (A_r + A_p)\omega_y^{(r)} \\ (C_r + C_p)\omega_z^{(r)} \end{Bmatrix} = \begin{Bmatrix} [(C_p - A_p) + (C_r - A_r)]\omega_y^{(r)}\omega_z^{(r)} \\ [(A_p - C_p) + (A_r - C_r)]\omega_x^{(r)}\omega_z^{(r)} \\ 0 \end{Bmatrix} \tag{l}$$

$$[\mathbf{I}_{G_p}^{(p)}]\{\dot{\boldsymbol{\omega}}_{\text{rel}}^{(p)}\} = \begin{bmatrix} \overline{A}_p & 0 & 0 \\ 0 & \overline{A}_p & 0 \\ 0 & 0 & C_p \end{bmatrix} \begin{Bmatrix} 0 \\ 0 \\ \dot{\omega}_p \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ C_p\dot{\omega}_p \end{Bmatrix} \tag{m}$$

$$\{\boldsymbol{\omega}^{(r)}\} \times [\mathbf{I}_{G_p}^{(p)}]\{\boldsymbol{\omega}_{\text{rel}}^{(p)}\} = \begin{Bmatrix} \omega_x^{(r)} \\ \omega_y^{(r)} \\ \omega_z^{(r)} \end{Bmatrix} \times \begin{bmatrix} \overline{A}_p & 0 & 0 \\ 0 & \overline{A}_p & 0 \\ 0 & 0 & C_p \end{bmatrix} \begin{Bmatrix} 0 \\ 0 \\ \omega_p \end{Bmatrix} = \begin{Bmatrix} C_p\omega_y^{(r)}\omega_p \\ -C_p\omega_x^{(r)}\omega_p \\ 0 \end{Bmatrix} \tag{n}$$

With these four expressions, (g) becomes

$$\begin{Bmatrix} (A_r + A_p)\dot{\omega}_x^{(r)} \\ (A_r + A_p)\dot{\omega}_y^{(r)} \\ (C_r + C_p)\dot{\omega}_z^{(r)} \end{Bmatrix} + \begin{Bmatrix} [(C_p - A_p) + (C_r - A_r)]\omega_y^{(r)}\omega_z^{(r)} \\ [(A_p - C_p) + (A_r - C_r)]\omega_x^{(r)}\omega_z^{(r)} \\ 0 \end{Bmatrix} + \begin{Bmatrix} 0 \\ 0 \\ C_p\dot{\omega}_p \end{Bmatrix} + \begin{Bmatrix} C_p\omega_y^{(r)}\omega_p \\ -C_p\omega_x^{(r)}\omega_p \\ 0 \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \\ 0 \end{Bmatrix} \tag{o}$$

Combining the four vectors on the left-hand side, and then extracting the three components of the vector equation finally yields the three equations of motion of the dual-spin satellite in the body frame,

$$\begin{aligned} A\dot{\omega}_x^{(r)} + (C - A)\omega_y^{(r)}\omega_z^{(r)} + C_p\omega_y^{(r)}\omega_p &= 0 \\ A\dot{\omega}_y^{(r)} + (A - C)\omega_x^{(r)}\omega_z^{(r)} - C_p\omega_x^{(r)}\omega_p &= 0 \\ C\dot{\omega}_z^{(r)} + C_p\dot{\omega}_p &= 0 \end{aligned} \tag{p}$$

where A and C are the combined transverse and axial moments of inertia of the dual-spin vehicle about its center of mass,

$$A = A_r + A_p \qquad C = C_r + C_p \tag{q}$$

The three equations (p) involve four unknowns, $\omega_x^{(r)}, \omega_y^{(r)}, \omega_z^{(r)}$ and ω_p. A fourth equation is required to account for the means of providing the relative velocity ω_p between the platform and the rotor. Friction in the axle bearing between the platform and the rotor would eventually cause ω_p to go to zero, as pointed out in Section 8.1.4. We may assume that the electric motor in the bearing acts to keep ω_p constant at a specified value, so that $\dot{\omega}_p = 0$. Then Equation $(\text{p})_3$ implies that $\omega_z^{(r)} = \text{constant}$ as well. Thus, ω_p and $\omega_z^{(r)}$ are removed from our list of unknowns, leaving $\omega_x^{(r)}$ and $\omega_y^{(r)}$ to be governed by the first two equations in (p).

Example 8.1.10

A spacecraft in torque-free motion has three identical momentum wheels with their spin axes aligned with the vehicle's principal body axes. The spin axes of momentum wheels 1, 2 and 3 are aligned with the x,y and z axes, respectively. The inertia tensors of the rotationally symmetric momentum wheels about their centers of mass are, therefore,

$$[\mathbf{I}_{G_1}^{(1)}] = \begin{bmatrix} I & 0 & 0 \\ 0 & J & 0 \\ 0 & 0 & J \end{bmatrix} \qquad [\mathbf{I}_{G_2}^{(2)}] = \begin{bmatrix} J & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & J \end{bmatrix} \qquad [\mathbf{I}_{G_3}^{(3)}] = \begin{bmatrix} J & 0 & 0 \\ 0 & J & 0 \\ 0 & 0 & I \end{bmatrix} \tag{a}$$

The spacecraft moment of inertia tensor about the vehicle center of mass is

$$[\mathbf{I}_G^{(v)}] = \begin{bmatrix} A & 0 & 0 \\ 0 & B & 0 \\ 0 & 0 & C \end{bmatrix} \tag{b}$$

Calculate the spin accelerations of the momentum wheels in the presence of external torque.

The absolute angular velocity $\boldsymbol{\omega}$ of the spacecraft and the angular velocities $\boldsymbol{\omega}_{\text{rel}}^{(1)}, \boldsymbol{\omega}_{\text{rel}}^{(2)}, \boldsymbol{\omega}_{\text{rel}}^{(3)}$ of flywheels *relative* to the spacecraft are

$$\{\boldsymbol{\omega}\} = \begin{Bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} \qquad \{\boldsymbol{\omega}^{(1)}\}_{\text{rel}} = \begin{Bmatrix} \omega^{(1)} \\ 0 \\ 0 \end{Bmatrix} \qquad \{\boldsymbol{\omega}^{(2)}\}_{\text{rel}} = \begin{Bmatrix} 0 \\ \omega^{(2)} \\ 0 \end{Bmatrix} \qquad \{\boldsymbol{\omega}^{(3)}\}_{\text{rel}} = \begin{Bmatrix} 0 \\ 0 \\ \omega^{(3)} \end{Bmatrix} \tag{c}$$

Therefore, the angular momentum of the spacecraft and momentum wheels is

$$\begin{aligned} \{\mathbf{H}_G\} = {} & [\mathbf{I}_G^{(v)}]\{\boldsymbol{\omega}\} + [\mathbf{I}_{G1}^{(1)}](\{\boldsymbol{\omega}\} + \{\boldsymbol{\omega}^{(1)}\}\text{rel}) + [\mathbf{I}_{G_2}^{(2)}](\{\boldsymbol{\omega}\} + \{\boldsymbol{\omega}^{(2)}\}_{\text{rel}}) \\ & + [\mathbf{I}_{G_3}^{(3)}](\{\boldsymbol{\omega}\} + \{\boldsymbol{\omega}^{(3)}\}_{\text{rel}}) \end{aligned} \tag{d}$$

Substituting Equations (a), (b) and (c) into this expression yields

$$\{\mathbf{H}_G\} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{Bmatrix} \omega^{(1)} \\ \omega^{(2)} \\ \omega^{(3)} \end{Bmatrix} + \begin{bmatrix} A+I+2J & 0 & 0 \\ 0 & B+I+2J & 0 \\ 0 & 0 & C+I+2J \end{bmatrix} \begin{Bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} \tag{e}$$

In this case, Euler's equations are

$$\{\dot{\mathbf{H}}_G\}_{\text{rel}} + \{\boldsymbol{\omega}\} \times \{\mathbf{H}_G\} = \{\mathbf{M}_G\} \tag{f}$$

Substitutuing (e), we get

$$\begin{aligned} & \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{Bmatrix} \dot{\omega}^{(1)} \\ \dot{\omega}^{(2)} \\ \dot{\omega}^{(3)} \end{Bmatrix} + \begin{bmatrix} A+I+2J & 0 & 0 \\ 0 & B+I+2J & 0 \\ 0 & 0 & C+I+2J \end{bmatrix} \begin{Bmatrix} \dot{\omega}_x \\ \dot{\omega}_y \\ \dot{\omega}_z \end{Bmatrix} + \begin{Bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} \\ & \times \left(\begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \begin{Bmatrix} \omega^{(1)} \\ \omega^{(2)} \\ \omega^{(3)} \end{Bmatrix} + \begin{bmatrix} A+I+2J & 0 & 0 \\ 0 & B+I+2J & 0 \\ 0 & 0 & C+I+2J \end{bmatrix} \begin{Bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{Bmatrix} \right) \\ & = \begin{Bmatrix} M_{G_x} \\ M_{G_y} \\ M_{G_z} \end{Bmatrix} \end{aligned} \tag{g}$$

Expanding and collecting terms yields the time rates of change of the flywheel spins (relative to the spacecraft) in terms of those of the spacecraft's absolute angular velocity components,

$$\begin{aligned} \dot{\omega}^{(1)} &= \frac{M_{G_x}}{I} + \frac{B-C}{I}\omega_y\omega_z - \left(1 + \frac{A}{I} + 2\frac{J}{I}\right)\dot{\omega}_x + \omega^{(2)}\omega_z - \omega^{(3)}\omega_y \\ \dot{\omega}^{(2)} &= \frac{M_{G_y}}{I} + \frac{C-A}{I}\omega_x\omega_z - \left(1 + \frac{B}{I} + 2\frac{J}{I}\right)\dot{\omega}_y + \omega^{(3)}\omega_x - \omega^{(1)}\omega_z \\ \dot{\omega}^{(3)} &= \frac{M_{G_z}}{I} + \frac{A-B}{I}\omega_x\omega_y - \left(1 + \frac{C}{I} + 2\frac{J}{I}\right)\dot{\omega}_z + \omega^{(1)}\omega_y - \omega^{(2)}\omega_x \end{aligned} \tag{h}$$

Example 8.1.11

The communications satellite is in a circular earth orbit of period T. The body z axis always points towards the earth, so the angular velocity about the body y axis is $2\pi/T$ (Figure 8.1-22). The angular velocities about the body x and z axes are zero. The attitude control system consists of three momentum wheels 1, 2 and 3 aligned with the principal x, y and z axes of the satellite. Variable torque is applied to each wheel by its own electric motor. At time $t = 0$ the angular velocities of the three wheels relative to the spacecraft are all zero. A small, constant environmental torque $\mathbf{M_0}$ acts on the spacecraft.

Determine the axial torques $C^{(1)}$, $C^{(2)}$ and $C^{(3)}$ that the three motors must exert on their wheels so that the angular velocity $\boldsymbol{\omega}$ of the satellite will remain constant. The moment of inertia of each reaction wheel about its spin axis is I.

The absolute angular velocity of the xyz frame is given by

$$\boldsymbol{\omega} = \omega_0 \hat{\mathbf{j}} \tag{a}$$

where $\omega_0 = 2\pi/T$, a constant. At any instant, the absolute angular velocities of the three reaction wheels are, accordingly,

$$\begin{aligned}
\boldsymbol{\omega}^{(1)} &= \omega^{(1)}\hat{\mathbf{i}} + \omega_0\hat{\mathbf{j}} \\
\boldsymbol{\omega}^{(2)} &= [\omega^{(2)} + \omega_0]\hat{\mathbf{j}} \\
\boldsymbol{\omega}^{(3)} &= \omega_0\hat{\mathbf{j}} + \omega^{(3)}\hat{\mathbf{k}}
\end{aligned} \tag{b}$$

From (a) it is clear that $\omega_x = \omega_z = \dot{\omega}_x = \dot{\omega}_y = \dot{\omega}_z = 0$. Therefore, Equations (h) of Example 8.1.10 become, for the case at hand,

$$\begin{aligned}
\dot{\omega}^{(1)} &= \frac{M_{G_x}}{I} + \frac{B-C}{I}\omega_0(0) - \left(1 + \frac{A}{I} + 2\frac{J}{I}\right)(0) + \omega^{(2)}(0) - \omega^{(3)}\omega_0 \\
\dot{\omega}^{(2)} &= \frac{M_{G_y}}{I} + \frac{C-A}{I}(0)(0) - \left(1 + \frac{B}{I} + 2\frac{J}{I}\right)(0) + \omega^{(3)}(0) - \omega^{(1)}(0) \\
\dot{\omega}^{(3)} &= \frac{M_{G_z}}{I} + \frac{A-B}{I}(0)\omega_0 - \left(1 + \frac{C}{I} + 2\frac{J}{I}\right)(0) + \omega^{(1)}\omega_0 - \omega^{(2)}(0)
\end{aligned}$$

which reduce to the following set of three first order differential equations,

$$\begin{aligned}
\dot{\omega}^{(1)} + \omega_0\omega^{(3)} &= \frac{M_{G_x}}{I} \\
\dot{\omega}^{(2)} &= \frac{M_{G_y}}{I} \\
\dot{\omega}^{(3)} - \omega_0\omega^{(1)} &= \frac{M_{G_z}}{I}
\end{aligned} \tag{c}$$

Equation $(c)_2$ implies that $\omega^{(2)} = M_{G_y}t/I +$ constant, and since $\omega^{(2)} = 0$ at $t = 0$, this means that for time thereafter,

$$\omega^{(2)} = \frac{M_{G_y}}{I}t \tag{d}$$

Differentiating $(c)_3$ with respect to t and solving for $\dot{\omega}^{(1)}$ yields $\dot{\omega}^{(1)} = \ddot{\omega}^{(3)}/\omega_0$. Substituting this result into $(c)_1$ we get

$$\ddot{\omega}^{(3)} + \omega_0^2\omega^{(3)} = \frac{\omega_0 M_{G_x}}{I}$$

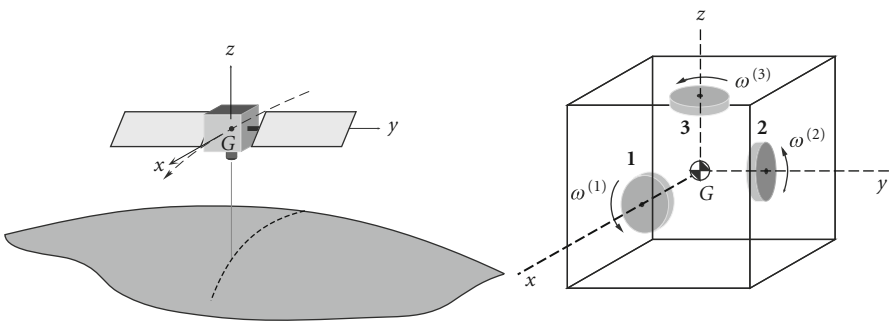


Figure 8.1-22 Three-axis stabilized satellite.

The well-known solution of this differential equation is

$$\omega^{(3)} = a \cos \omega_0 t + b \sin \omega_0 t + \frac{M_{G_x}}{I\omega_0}$$

where a and b are constants of integration. According to the problem statement, $\omega^{(3)} = 0$ when $t = 0$. This initial condition requires $a = -M_{G_x}/\omega_0 I$, so that

$$\omega^{(3)} = b \sin \omega_0 t + \frac{M_{G_x}}{I\omega_0}(1 - \cos \omega_0 t) \tag{e}$$

From this we obtain $\dot{\omega}^{(3)} = b\omega_0 \cos \omega_0 t + \frac{M_{G_x}}{I} \sin \omega_0 t$, which, when substituted into $(c)_3$, yields

$$\omega^{(1)} = b \cos \omega_0 t + \frac{M_{G_x}}{I\omega_0} \sin \omega_0 t - \frac{M_{G_z}}{I\omega_0} \tag{f}$$

Since $\omega^{(1)} = 0$ at $t = 0$, this implies $b = M_{G_z}/\omega_0 I$. In summary, therefore, the angular velocities of wheels 1, 2 and 3 relative to the satellite are

$$\omega^{(1)} = \frac{M_{G_x}}{I\omega_0} \sin \omega_0 t + \frac{M_{G_z}}{I\omega_0}(\cos \omega_0 t - 1) \tag{g1}$$

$$\omega^{(2)} = \frac{M_{G_y}}{I} t \tag{g2}$$

$$\omega^{(3)} = \frac{M_{G_z}}{I\omega_0} \sin \omega_0 t + \frac{M_{G_x}}{I\omega_0}(1 - \cos \omega_0 t) \tag{g3}$$

The angular momenta of the reaction wheels are

$$\begin{aligned}
\mathbf{H}_{G_1}^{(1)} &= I_x^{(1)}\omega_x^{(1)}\hat{\mathbf{i}} + I_y^{(1)}\omega_y^{(1)}\hat{\mathbf{j}} + I_z^{(1)}\omega_z^{(1)}\hat{\mathbf{k}} \\
\mathbf{H}_{G_2}^{(2)} &= I_x^{(2)}\omega_x^{(2)}\hat{\mathbf{i}} + I_y^{(2)}\omega_y^{(2)}\hat{\mathbf{j}} + I_z^{(2)}\omega_z^{(2)}\hat{\mathbf{k}} \\
\mathbf{H}_{G_3}^{(3)} &= I_x^{(3)}\omega_x^{(3)}\hat{\mathbf{i}} + I_y^{(3)}\omega_y^{(3)}\hat{\mathbf{j}} + I_z^{(3)}\omega_z^{(3)}\hat{\mathbf{k}}
\end{aligned} \tag{h}$$

According to (b), the components of the flywheels' angular velocities are

$$\begin{aligned}
&\omega_x^{(1)} = \omega^{(1)} \qquad \omega_y^{(1)} = \omega_0 \qquad \omega_z^{(1)} = 0 \\
&\omega_x^{(2)} = 0 \qquad \omega_y^{(2)} = \omega^{(2)} + \omega_0 \qquad \omega_z^{(2)} = 0 \\
&\omega_x^{(3)} = 0 \qquad \omega_y^{(3)} = \omega_0 \qquad \omega_z^{(3)} = \omega^{(3)}
\end{aligned}$$

Furthermore, $I_x^{(1)} = I_y^{(2)} = I_z^{(3)} = I$, so that (h) becomes

$$\begin{aligned}
\mathbf{H}_{G_1}^{(1)} &= I\omega^{(1)}\hat{\mathbf{i}} + I_y^{(1)}\omega_0\hat{\mathbf{j}} \\
\mathbf{H}_{G_2}^{(2)} &= I(\omega^{(2)} + \omega_0)\hat{\mathbf{j}} \\
\mathbf{H}_{G_3}^{(3)} &= I_y^{(3)}\omega_0\hat{\mathbf{j}} + I\omega^{(3)}\hat{\mathbf{k}}
\end{aligned} \tag{i}$$

Substituting (g) into these expressions yields the angular momenta of the wheels as a function of time,

$$\begin{aligned}
\mathbf{H}_{G_1}^{(1)} &= \left[\frac{M_{G_x}}{\omega_0} \sin \omega_0 t + \frac{M_{G_z}}{\omega_0}(\cos \omega_0 t - 1)\right]\hat{\mathbf{i}} + I_y^{(1)}\omega_0\hat{\mathbf{j}} \\
\mathbf{H}_{G_2}^{(2)} &= (M_{G_y} t + I\omega_0)\hat{\mathbf{j}} \\
\mathbf{H}_{G_3}^{(3)} &= I_y^{(3)}\omega_0\hat{\mathbf{j}} + \left[\frac{M_{G_z}}{\omega_0} \sin \omega_0 t + \frac{M_{G_x}}{\omega_0}(1 - \cos \omega_0 t)\right]\hat{\mathbf{k}}
\end{aligned} \tag{j}$$

The torque on the reaction wheels is found by applying Euler's equation to each one. Thus, for wheel 1

$$\begin{aligned}
\mathbf{M}G_{1_{\text{net}}} &= \left(\frac{d\mathbf{H}_{G_1}^{(1)}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega} \times \mathbf{H}_{G_1}^{(1)} \\
&= (M_{G_x} \cos \omega_0 t - M_{G_z} \sin \omega_0 t)\hat{\mathbf{i}} + [M_{G_z}(1 - \cos \omega_0 t) - M_{G_x} \sin \omega_0 t]\hat{\mathbf{k}}
\end{aligned}$$

Since the axis of wheel 1 is in the x direction, the torque is the x component of this moment (the z component being a gyroscopic bending moment),

$$\underline{C^{(1)} = M_{G_x}\cos\omega_0 t - M_{G_z}\sin\omega_0 t}$$

For wheel 2,

$$\mathbf{M}_{G_{2_{\text{net}}}} = \left.\frac{d\mathbf{H}_{G_2}^{(2)}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega}\times\mathbf{H}_{G_2}^{(2)} = \mathbf{M}_{G_y}\hat{\mathbf{j}}$$

Thus

$$\underline{C^{(2)} = M_{G_y}}$$

Finally, for wheel 3

$$\begin{aligned}\mathbf{M}_{G_{3_{\text{net}}}} &= \left.\frac{d\mathbf{H}_{G_3}^{(3)}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega}\times\mathbf{H}_{G_3}^{(3)}\\ &= [M_{G_x}(1-\cos\omega_0 t) + M_{G_z}\sin\omega_0 t]\hat{\mathbf{i}} + (M_{G_x}\sin\omega_0 t + M_{G_z}\cos\omega_0 t)\hat{\mathbf{k}}\end{aligned}$$

For this wheel, the torque direction is the z axis, so

$$\underline{C^{(3)} = M_{G_x}\sin\omega_0 t + M_{G_z}\cos\omega_0 t}$$

The external torques on the spacecraft of the previous example may be due to thruster misalignment or they may arise from environmental effects such as gravity gradients or solar pressure. The example assumed that these torques were constant, which is the simplest means of introducing their effects, but they actually vary with time. In any case, their magnitudes are extremely small, typically less than 10^{-3} N·m for ordinary-sized, unmanned spacecraft. Equation (g_2) of the example reveals that a small torque normal to the satellite's orbital plane will cause the angular velocity of momentum wheel 2 to slowly but constantly increase. Over a long enough period of time, the angular velocity of the gyro might approach its design limits, whereupon it is said to be *saturated*. At that point, attitude jets on the satellite would have to be fired to produce a torque around the y axis while the wheel is 'caged', i.e., its angular velocity is reduced to zero or to its non-zero *bias* value. Finally, note that if all of the external torques were zero, none of the momentum wheels in the example would be required. The constant angular velocity $\boldsymbol{\omega} = (2\pi/T)\hat{\mathbf{j}}$ of the vehicle, once initiated, would continue unabated.

So far we have dealt with momentum wheels, which are characterized by the fact that their axes are rigidly aligned with the principal axes of the spacecraft, as shown in Figure 8.1-23. The speed of the electrically driven wheels is varied to produce the required rotation rates of the vehicle in response to external torques. Depending on the spacecraft, the nominal speed of a momentum wheel may be from zero to several thousand rpm.

Momentum wheels that are free to pivot on one or more gimbals are called control moment gyros. Figure 8.1-24 illustrates a double-gimbaled control moment gyro. These gyros spin at several thousand rpm. The motor-driven

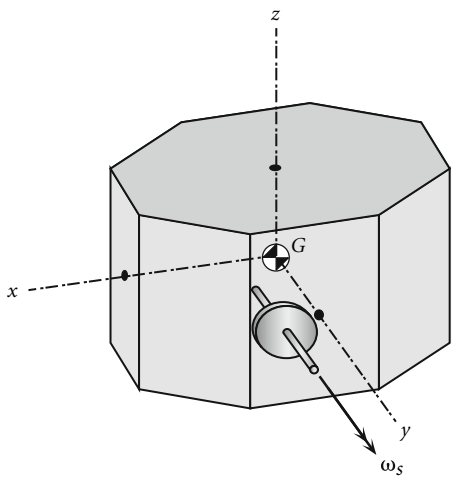


Figure 8.1-23 Momentum wheel aligned with a principal body axis

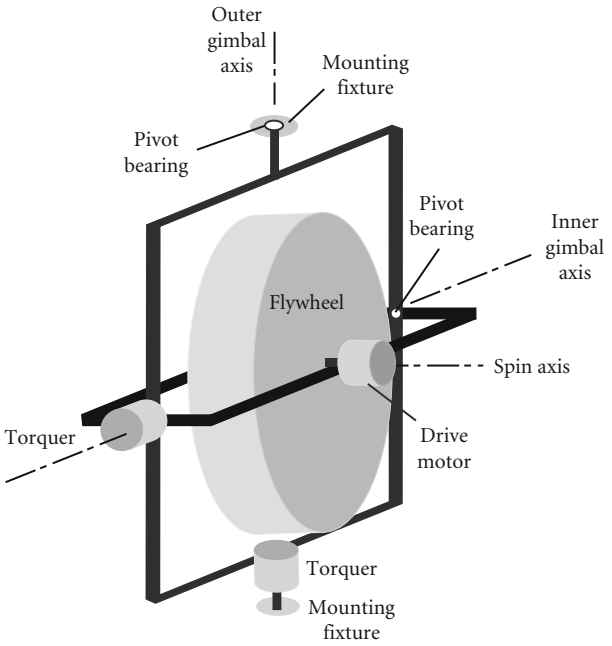


Figure 8.1-24 Two-gimbal control moment gyro.

speed of the flywheel is constant, and moments are exerted on the vehicle when torquers (electric motors) tilt the wheel about a gimbal axis. The torque direction is normal to the gimbal axis. To simplify the analysis of high-rpm gyros, we can assume that the angular momentum is directed totally along the spin axis. That is, in calculating the angular momentum $\mathrm{H}_{G_w}^{(w)}$ of a momentum wheel about its center of mass, we use the formula

$$\{\mathbf{H}_{G_w}^{(w)}\} = [\mathbf{I}_{G_w}^{(w)}]\{\boldsymbol{\omega}^{(w)}\}$$

where $\boldsymbol{\omega}^{(w)}$ is the absolute angular velocity of the spinning flywheel, which may be written as

$$\boldsymbol{\omega}^{(w)} = \boldsymbol{\omega}^{(v)} + \boldsymbol{\omega}_p^{(w)} + \boldsymbol{\omega}_n^{(w)} + \boldsymbol{\omega}_s^{(w)}$$

$\boldsymbol{\omega}^{(v)}$ is the angular velocity of the vehicle to which the gyro is attached, while $\boldsymbol{\omega}_p^{(w)}$, $\boldsymbol{\omega}_n^{(w)}$ and $\boldsymbol{\omega}_s^{(w)}$ are the precession, nutation and spin rates of the gyro relative to the vehicle. The spin rate of the gyro is three or more orders of magnitude greater than any of the other rates. That is, under conditions in which a control moment gyro is designed to operate,

$$\|\boldsymbol{\omega}_s^{(w)}\| \gg \|\boldsymbol{\omega}^{(v)}\| \qquad \|\boldsymbol{\omega}_s^{(w)}\| \gg \|\boldsymbol{\omega}_p^{(w)}\| \qquad \|\boldsymbol{\omega}_s^{(w)}\| \gg \|\boldsymbol{\omega}_n^{(w)}\|$$

We may therefore accurately express the angular momentum of any high-rpm gyro as

$$\{\mathbf{H}_{G_w}^{(w)}\} = [\mathbf{I}_{G_w}^{(w)}]\{\boldsymbol{\omega}_s^{(w)}\} \tag{8.1.138}$$

Since the spin axis of a gyro is an axis of symmetry, about which the moment of inertia is $C^{(w)}$, this can be written as

$$\mathbf{H}_{G_w}^{(w)} = C^{(w)}\omega_s^{(w)}\hat{\mathbf{n}}_s^{(w)}$$

where $\hat{\mathbf{n}}_s^{(w)}$ is the unit vector along the spin axis, as illustrated in Figure 8.1-25. Relative to the body frame axes of the spacecraft, the components of $\hat{\mathbf{n}}_s^{(w)}$ appear as follows,

$$\hat{\mathbf{n}}_s^{(w)} = \sin\theta\cos\phi\hat{\mathbf{i}} + \sin\theta\ \sin\phi\hat{\mathbf{j}} + \cos\theta\hat{\mathbf{k}} \tag{8.1.139}$$

If we let

$$H^{(w)} = C^{(w)}\omega_s^{(w)}$$

then Equation 8.1.138 becomes, simply,

$$\mathbf{H}_{G_w}^{(w)} = H^{(w)}\hat{\mathbf{n}}_s^{(w)} \tag{8.1.140}$$

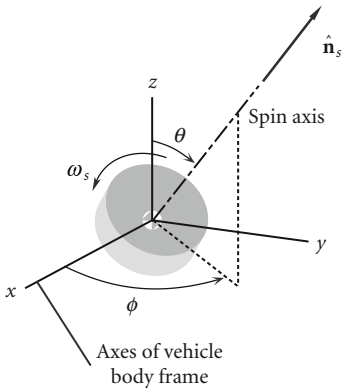


Figure 8.1-25 Inclination angles of the spin vector of a gyro.

Let us consider the equation of motion of a spacecraft with a single gyro. From Equation 8.1.136,

$$\left.\frac{d\mathbf{H}_G^{(v)}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega} \times \mathbf{H}_G^{(v)} + \left.\frac{d\mathbf{H}_{G_w}^{(w)}}{dt}\right)_{\text{rel}} + \boldsymbol{\omega} \times \mathbf{H}_{G_w}^{(w)} = \mathbf{M}_{G_{\text{net}}} \tag{8.1.141}$$

Calculating each term on the left, we have, for the vehicle,

$$\begin{aligned} \mathbf{H}_G^{(v)} &= A\omega_x\hat{\mathbf{i}} + B\omega_y\hat{\mathbf{j}} + C\omega_z\hat{\mathbf{k}} \\ \left.\frac{d\mathbf{H}_G^{(v)}}{dt}\right)_{\text{rel}} &= A\dot{\omega}_x\hat{\mathbf{i}} + B\dot{\omega}_y\hat{\mathbf{j}} + C\dot{\omega}_z\hat{\mathbf{k}} \end{aligned} \tag{8.1.142}$$

$$\boldsymbol{\omega} \times \mathbf{H}_G^{(v)} = \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \omega_x & \omega_y & \omega_z \\ A\omega_x & B\omega_y & C\omega_z \end{vmatrix} = (C-B)\omega_y\omega_z\hat{\mathbf{i}} + (A-C)\omega_x\omega_z\hat{\mathbf{j}} + (B-A)\hat{\mathbf{k}} \tag{8.1.143}$$

For the gyro,

$$\begin{aligned} H_{G_w}^{(w)} &= H^{(w)}\hat{\mathbf{n}}_s^{(w)} = H^{(w)}\sin\theta\cos\phi\hat{\mathbf{i}} + H^{(w)}\sin\theta\sin\phi\hat{\mathbf{j}} + H^{(w)}\cos\theta\hat{\mathbf{k}} \\ \left.\frac{d\mathbf{H}_{G_w}^{(w)}}{dt}\right)_{\text{rel}} &= (\dot{H}^{(w)}\sin\theta\cos\phi + H^{(w)}\dot{\theta}\cos\theta\cos\phi - H^{(w)}\dot{\phi}\sin\theta\sin\phi)\hat{\mathbf{i}} \\ &\quad + (\dot{H}^{(w)}\sin\theta\sin\phi + H^{(w)}\dot{\theta}\cos\theta\sin\phi + H^{(w)}\dot{\phi}\sin\theta\cos\phi)\hat{\mathbf{j}} \\ &\quad + (\dot{H}^{(w)}\cos\theta - H^{(w)}\dot{\theta}\sin\theta)\hat{\mathbf{k}} \end{aligned} \tag{8.1.144}$$

$$\begin{aligned} \boldsymbol{\omega} \times \mathbf{H}_{G_w}^{(w)} &= \begin{vmatrix} \hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}} \\ \omega_x & \omega_y & \omega_z \\ H^{(w)}\sin\theta\cos\phi & H^{(w)}\sin\theta\sin\phi & H^{(w)}\cos\theta \end{vmatrix} \\ &= (H^{(w)}\omega_y\cos\theta - H^{(w)}\omega_z\sin\phi\sin\theta)\hat{\mathbf{i}} \\ &\quad + (-H^{(w)}\omega_x\cos\theta + H^{(w)}\omega_z\cos\phi\sin\theta)\hat{\mathbf{j}} \\ &\quad + (-H^{(w)}\omega_y\cos\phi\sin\theta + H^{(w)}\omega_x\sin\phi\sin\theta)\hat{\mathbf{k}} \end{aligned} \tag{8.1.145}$$

Substituting Equations 8.1.142 through 8.1.145 into Equation 8.1.141 yields a vector equation with the following three components:

$$\begin{aligned} &A\dot{\omega}_x + H^{(w)}\dot{\theta}\cos\phi\cos\theta - H^{(w)}\dot{\phi}\sin\phi\sin\theta + \dot{H}^{(w)}\cos\phi\sin\theta \\ &\quad + (H^{(w)}\cos\theta + C\omega_z)\omega_y - (H^{(w)}\sin\phi\sin\theta + B\omega_y)\omega_z = M_{G_{\text{net}x}} \end{aligned} \tag{8.1.146a}$$

$$\begin{aligned} &B\dot{\omega}_y + H^{(w)}\dot{\theta}\sin\phi\cos\theta + H^{(w)}\dot{\phi}\cos\phi\sin\theta + \dot{H}^{(w)}\sin\phi\sin\theta \\ &\quad - (H^{(w)}\cos\theta + C\omega_z)\omega_x + (H^{(w)}\cos\phi\sin\theta + A\omega_x)\omega_z = M_{G_{\text{net}y}} \end{aligned} \tag{8.1.146b}$$

$$C\dot{\omega}_z - H^{(w)}\dot{\theta}\sin\theta + \dot{H}^{(w)}\cos\theta - (H^{(w)}\cos\phi\sin\theta + A\omega_x)\omega_y + (H^{(w)}\sin\phi\sin\theta + B\omega_y)\omega_x = M_{G_{\text{net}z}} \tag{8.1.146c}$$

Additional gyros are accounted for by adding the components of Equations 8.1.144 and 8.1.145 for each additional unit.

Example 8.1.12

A satellite is in torque-free motion ($\mathbf{M}_{G_{\text{net}}} = \mathbf{0}$). A non-gimbaled gyro (momentum wheel) is aligned with the vehicle's x axis and is spinning at the rate ω_{S_0}. The spacecraft angular velocity is $\omega = \omega_x\hat{\mathbf{i}}$. If the spin of the gyro is increased at the rate $\dot{\omega}_s$, find the angular cceleration of the spacecraft.

Using Figure 8.1-25 as a guide, we set $\phi = 0$ and $\theta = 90°$ to align the spin axis with the x axis. Since there is no gimbaling, $\dot{\theta} = \dot{\phi} = 0$. Equations 8.1.146 then yield

$$\begin{aligned} A\dot{\omega}_x + \dot{H}^{(w)} &= 0 \\ B\dot{\omega}_y &= 0 \\ C\dot{\omega}_z &= 0 \end{aligned}$$

Clearly, the angular velocities around the y and z axes remain zero, whereas

$$\underline{\dot{\omega}_x = -\frac{\dot{H}^{(w)}}{A} = -\frac{C^{(w)}}{A}\dot{\omega}_s}$$

Thus, a change in the vehicle's roll rate around the x axis can be initiated by accelerating the momentum wheel in the opposite direction.

Example 8.1.13

A satellite is in torque-free motion. A control moment gyro, spinning at the constant rate ω_s, is gimbaled about the spacecraft y and z axes, with $\phi = 0$ and $\theta = 90°$ (cf. Figure 8.1-25). The spacecraft angular velocity is $\boldsymbol{\omega} = \omega_z\hat{\mathbf{k}}$. If the spin axis of the gyro, initially along the x direction, is rotated around the y axis at the rate $\dot{\theta}$, what is the resulting angular acceleration of the spacecraft?

Substituting $\omega_x = \omega_y = \dot{H}^{(w)} = \phi = 0$ and $\theta = 90°$ into Equations 8.1.146 gives

$$\begin{aligned} A\dot{\omega}_x &= 0 \\ B\dot{\omega}_y + H^{(w)}(\omega_z + \dot{\phi}) &= 0 \\ C\dot{\omega}_z - H^{(w)}\dot{\theta} &= 0 \end{aligned}$$

where $H^{(w)} = C^{(w)}\omega_s$. Thus, the components of vehicle angular acceleration are

$$\underline{\dot{\omega}_x = 0 \quad \dot{\omega}_y = -\frac{C^{(w)}}{B}\omega_s(\omega_z + \dot{\phi}) \qquad \dot{\omega}_z = \frac{C^{(w)}}{C}\omega_s\dot{\theta}}$$

We see that pitching the gyro at the rate $\dot{\theta}$ around the vehicle y axis alters only ω_z, leaving ω_x unchanged. However, to keep $\omega_y = 0$ clearly requires $\dot{\phi} = -\omega_z$. In other words, for the control moment gyro to control the angular velocity about only one vehicle axis, it must therefore be able to precess around that axis (the z axis in this case). That is why the control moment gyro must have two gimbals.

8.1.10 Gravity-Gradient Stabilization

Consider a satellite in circular orbit, as shown in Figure 8.1-26. Let $\mathbf{r}$ be the position vector of a mass element dm relative to the center of attraction, $\mathbf{r}_0$ the position vector of the center of mass G, and $\boldsymbol{\varrho}$ the position of dm relative to G. The force of gravity on dm is

$$d\mathbf{F}_g = -G\frac{M\,dm}{r^3}\mathbf{r} = -\mu\frac{\mathbf{r}}{r^3}dm \tag{8.1.147}$$

where M is the mass of the central body, and $\mu = GM$. The net moment of the gravitational force around G is

$$\mathbf{M}_{G_{\text{net}}} = \int_m \boldsymbol{\varrho} \times d\mathbf{F}_g\,dm \tag{8.1.148}$$

Since $\mathbf{r} = \mathbf{r}_0 + \boldsymbol{\varrho}$, and

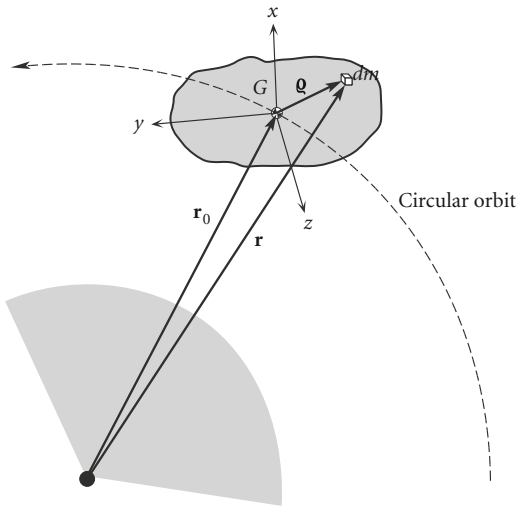


Figure 8.1-26 Rigid satellite in a circular orbit is the principal body frame.

$$\begin{aligned}\mathbf{r}_0 &= r_{0_x}\hat{\mathbf{i}} + r_{0_y}\hat{\mathbf{j}} + r_{0_z}\hat{\mathbf{k}}\\ \boldsymbol{\varrho} &= x\hat{\mathbf{i}} + y\hat{\mathbf{j}} + z\hat{\mathbf{k}}\end{aligned} \tag{8.1.149}$$

we have

$$\boldsymbol{\varrho} \times d\mathbf{F}_g = -\mu\frac{dm}{r^3}\boldsymbol{\varrho} \times (\mathbf{r}_0 + \boldsymbol{\varrho}) = -\mu\frac{dm}{r^3}\boldsymbol{\varrho} \times \mathbf{r}_0 = -\mu\frac{dm}{r^3}\begin{vmatrix}\hat{\mathbf{i}} & \hat{\mathbf{j}} & \hat{\mathbf{k}}\\ x & y & z\\ r_{0_x} & r_{0_y} & r_{0_z}\end{vmatrix}$$

Thus,

$$\boldsymbol{\varrho} \times d\mathbf{F}_g = -\mu\frac{dm}{r^3}(r_{0_z}y - r_{0_y}z)\hat{\mathbf{i}} - \mu\frac{dm}{r^3}(r_{0_x}z - r_{0_z}x)\hat{\mathbf{j}} - \mu\frac{dm}{r^3}(r_{0_y}x - r_{0_x}y)\hat{\mathbf{k}}$$

Substituting this back into Equation 8.1.148 yields

$$\begin{aligned}\mathbf{M}_{G_{\text{net}}} = &\left(-\mu r_{0_x}\int_m \frac{y}{r^3}\,dm + \mu r_{0_y}\int_m \frac{z}{r^3}\,dm\right)\hat{\mathbf{i}} + \left(-\mu r_{0_x}\int_m \frac{z}{r^3}\,dm + \mu r_{0_z}\int_m \frac{x}{r^3}\,dm\right)\hat{\mathbf{j}}\\ &+\left(-\mu r_{0_y}\int_m \frac{x}{r^3}\,dm + \mu r_{0_x}\int_m \frac{y}{r^3}\,dm\right)\hat{\mathbf{k}}\end{aligned}$$

or

$$\begin{aligned}M_{G_{\text{net}_x}} &= -\mu r_{0_z}\int_m \frac{y}{r^3}\,dm + \mu r_{0_y}\int_m \frac{z}{r^3}\,dm\\ M_{G_{\text{net}_y}} &= -\mu r_{0_x}\int_m \frac{z}{r^3}\,dm + \mu r_{0_z}\int_m \frac{x}{r^3}\,dm\\ M_{G_{\text{net}_z}} &= -\mu r_{0y}\int_m \frac{x}{r^3}\,dm + \mu r_{0_x}\int_m \frac{y}{r^3}\,dm\end{aligned} \tag{8.1.150}$$

Now, since $\|\boldsymbol{\varrho}\| \ll \|\mathbf{r}_0\|$, it follows that

$$\frac{1}{r^3} = \frac{1}{r_0^3} - \frac{3}{r_0^5}\mathbf{r}_0 \cdot \boldsymbol{\varrho}$$

or

$$\frac{1}{r^3} = \frac{1}{r_0^3} - \frac{3}{r_0^5}(r_{0_x}x + r_{0_y}y + r_{0_z}z)$$

Therefore,

$$\int_m \frac{x}{r^3}\,dm = \frac{1}{r_0^3}\int_m x\,dm - \frac{3r_{0_x}}{r_0^5}\int_m x^2\,dm - \frac{3r_{0_x}}{r_0^5}\int_m xy\,dm - \frac{3r_{0_x}}{r_0^5}\int_m xz\,dm$$

But the center of mass lies at the origin of the xyz axes, which are principal moment of inertia directions. That means

$$\int_m x\,dm = \int_m xy\,dm = \int_m xz\,dm = 0$$

so that

$$\int_m \frac{x}{r^3}\,dm = -\frac{3r_{0_x}}{r_0^5}\int_m x^2\,dm \tag{8.1.151}$$

In a similar fashion, we can show that

$$\int_m \frac{y}{r^3}\,dm = -\frac{3r_{0_y}}{r_0^5}\int_m y^2\,dm \tag{8.1.152}$$

and

$$\int_m \frac{z}{r^3}\,dm = -\frac{3r_{0_y}}{r_0^5}\int_m z^2\,dm \tag{8.1.153}$$

Substituting these last three expressions into Equations 8.1.150 leads to

$$\begin{aligned}
M_{G_{\text{net}_x}} &= \frac{3\mu r_{0_y} r_{0_z}}{r_0^5}\left(\int_m y^2 dm - \int_m z^2 dm\right)\\
M_{G_{\text{net}_y}} &= \frac{3\mu r_{0_x} r_{0_z}}{r_0^5}\left(\int_m z^2 dm - \int_m x^2 dm\right)\\
M_{\text{net}_z} &= \frac{3\mu r_{0_x} r_{0_y}}{r_0^5}\left(\int_m x^2 dm - \int_m y^2 dm\right)
\end{aligned} \tag{8.1.154}$$

The moments of inertia are defined as

$$A = \int_m y^2 dm + \int_m z^2 dm \quad B = \int_m x^2 dm + \int_m z^2 dm \quad C = \int_m x^2 dm + \int_m y^2 dm \tag{8.1.155}$$

from which we may write

$$\begin{aligned}
B - A &= \int_m x^2 dm - \int_m y^2 dm \qquad A - C = \int_m z^2 dm - \int_m x^2 dm\\
C - B &= \int_m y^2 dm - \int_m z^2 dm
\end{aligned}$$

It follows that Equations 8.1.154 reduce to

$$\begin{aligned}
M_{G_{\text{net}_x}} &= \frac{3\mu r_{0_y} r_{0_z}}{r_0^5}(C - B)\\
M_{G_{\text{net}_y}} &= \frac{3\mu r_{0_x} r_{0_z}}{r_0^5}(A - C)\\
M_{G_{\text{net}_z}} &= \frac{3\mu r_{0_x} r_{0_y}}{r_0^5}(B - A)
\end{aligned} \tag{8.1.156}$$

These are the components, in the spacecraft body frame, of the gravitational torque produced by the variation of the earth's gravitational field over the volume of the spacecraft. To get an idea of these torque magnitudes, note first of all that $r_{0_x}/r_0, r_{0_y}/r_0$ and r_{0_z}/r_0 are the direction cosines of the position vector of the center of mass, so that their magnitudes do not exceed 1. For a satellite in a low earth orbit of radius 6700 km, $3\mu/r_0^3 \cong 4 \times 10^{-6}\text{s}^{-2}$, which is therefore the maximum order of magnitude of the coefficients of the inertia terms in Equation 8.1.156. The moments of inertia of the space shuttle are on the order of 10^6 kg $\cdot$ m^2, so the gravitational torques on this large vehicle are on the order of 1 N $\cdot$ m.

Substituting Equations 8.1.156 into Euler's equations of motion, we get

$$\begin{aligned}
A\dot{\omega}_x + (C - B)\omega_y\omega_z &= \frac{3\mu r_{0_y} r_{0_z}}{r_0^5}(C - B)\\
B\dot{\omega}_y + (A - C)\omega_z\omega_x &= \frac{3\mu r_{0_x} r_{0_z}}{r_0^5}(A - C)\\
C\dot{\omega}_z + (B - A)\omega_x\omega_y &= \frac{3\mu r_{0_x} r_{0_y}}{r_0^5}(B - A)
\end{aligned} \tag{8.1.157}$$

Now consider the orbital reference frame shown in Figure 8.1-27. The z' axis points radially outward from the center of the earth, the x' axis is in the direction of the local horizon, and the y' axis completes the right-handed triad by pointing in the direction of the orbit normal. This frame rotates around the y' axis with an angular velocity equal to the mean motion n of the circular orbit. Suppose we align the satellite's principal body frame axes xyz with $x'y'z'$, respectively. When the body x axis is aligned with the x' direction, it is called the *roll* axis. The body y axis, when aligned with the y' direction, is the *pitch* axis. The body z axis, pointing outward from the earth in the z' direction, is the *yaw* axis. These directions are illustrated in Figure 8.1-28. With the spacecraft aligned in this way, the body frame components of the inertial angular velocity $\boldsymbol{\omega}$ are $\omega_x = \omega_z = 0$ and $\omega_y = n$. The components of the position vector $\mathbf{r}_0$ are $r_{0_x} = r_{0_y} = 0$ and $r_{0_z} = r_0$. Substituting this data into Equations 8.1.157 yields

$$\dot{\omega}_x = \dot{\omega}_y = \dot{\omega}_z = 0$$

That is, the spacecraft will orbit the planet with its principal axes remaining aligned with the orbital frame. If this motion is stable under the influence of gravity alone, without the use of thrusters, gyros or other devices, then the spacecraft is *gravity gradient* stabilized. We need to assess the stability of this motion so we can determine how to orient a spacecraft to take advantage of this type of passive attitude stabilization.

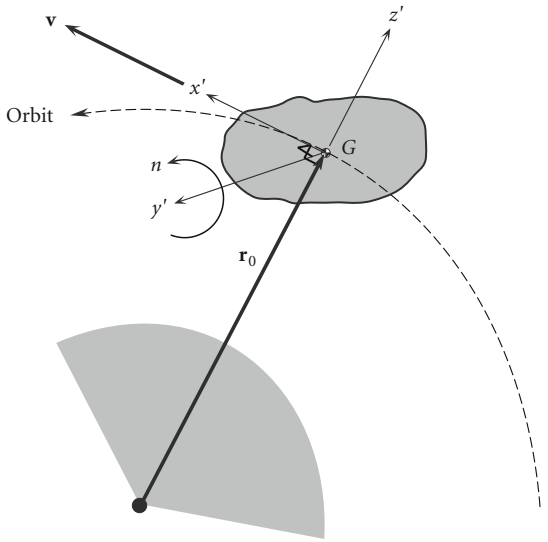


Figure 8.1-27 Orbital reference frame $x'y'z'$ attached to the center of mass of the satellite.

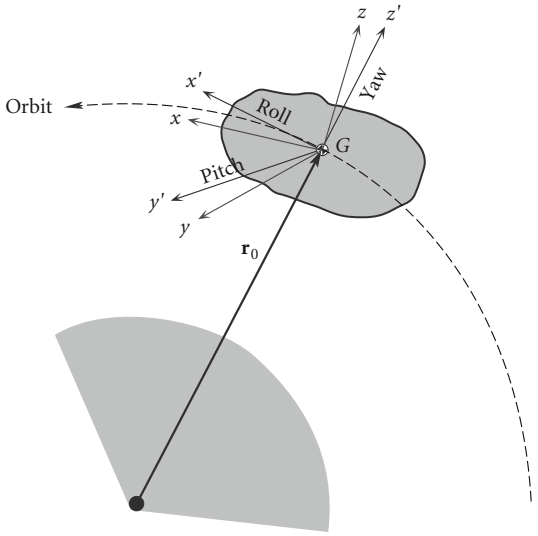


Figure 8.1-28 Satellite body frame slightly misaligned with the orbital frame $x'y'z'$.

Let the body frame xyz be slightly misaligned with the orbital reference frame, so that the yaw, pitch and roll angles between the xyz axes and the $x'y'z'$ axes, respectively, are very small, as suggested in Figure 8.1-28. The absolute angular velocity $\boldsymbol{\omega}$ of the spacecraft is the angular velocity $\boldsymbol{\omega}_{\text{rel}}$ relative to the orbital reference frame plus the inertial angular velocity $\boldsymbol{\Omega}$ of the $x'y'z'$ frame,

$$\boldsymbol{\omega} = \boldsymbol{\omega}_{\text{rel}} + \boldsymbol{\Omega}$$

The components of $\boldsymbol{\omega}_{\text{rel}}$ in the body frame are found using the yaw, pitch and roll relations. In so doing, it must be kept in mind that all angles and rates are assumed to be so small that their squares and products may be neglected. Recalling that $\sin\alpha = \alpha$ and $\cos\alpha = 1$ when $\alpha \ll 1$, we therefore obtain

$$\omega_{x_{\text{rel}}} = \omega_{\text{roll}} - \omega_{\text{yaw}} \overbrace{\sin\theta_{\text{pitch}}}^{=\theta_{\text{pitch}}} = \dot{\psi}_{\text{roll}} - \overbrace{\dot{\phi}_{\text{yaw}}\theta_{\text{pitch}}}^{\text{neglect product}} = \dot{\psi}_{\text{roll}} \tag{8.1.158}$$

$$\omega_{y_{\text{rel}}} = \omega_{\text{yaw}} \overbrace{\cos\theta_{\text{pitch}}}^{=1} \overbrace{\sin\psi_{\text{roll}}}^{=\psi_{\text{roll}}} + \omega_{\text{pitch}} \overbrace{\cos\psi_{\text{roll}}}^{=1} = \overbrace{\dot{\phi}_{\text{yaw}}\psi_{\text{roll}}}^{\text{neglect product}} + \dot{\theta}_{\text{pitch}} = \dot{\theta}_{\text{pitch}} \tag{8.1.159}$$

$$\omega_{z_{\text{rel}}} = \omega_{\text{yaw}} \overbrace{\cos\theta_{\text{pitch}}}^{=1} \overbrace{\cos\psi_{\text{roll}}}^{=1} - \omega_{\text{pitch}} \overbrace{\sin\psi_{\text{roll}}}^{=\psi_{\text{roll}}} = \dot{\phi}_{\text{yaw}} - \overbrace{\dot{\theta}_{\text{pitch}}\psi_{\text{roll}}}^{\text{neglect product}} = \dot{\phi}_{\text{yaw}} \tag{8.1.160}$$

The orbital frame's angular velocity is the mean motion n of the circular orbit, so that

$$\boldsymbol{\Omega} = n\hat{\mathbf{j}}'$$

To obtain the orbital frame's angular velocity components along the body frame, we must use the transformation rule

$$\{\boldsymbol{\Omega}\}_x = [\mathbf{Q}]_{x'x}\{\boldsymbol{\Omega}\}_{x'} \tag{8.1.161}$$

$$[\mathbf{Q}]_{x'x} = \begin{bmatrix} 1 & \phi_{\text{yaw}} & -\theta_{\text{pitch}} \\ -\phi_{\text{yaw}} & 1 & \psi_{\text{roll}} \\ \theta_{\text{pitch}} & -\psi_{\text{roll}} & 1 \end{bmatrix}$$

With this, Equation 8.1.161 becomes

$$\begin{Bmatrix} \Omega_x \\ \Omega_y \\ \Omega_z \end{Bmatrix} = \begin{bmatrix} 1 & \phi_{\text{yaw}} & -\theta_{\text{pitch}} \\ -\phi_{\text{yaw}} & 1 & \psi_{\text{roll}} \\ \theta_{\text{pitch}} & -\psi_{\text{roll}} & 1 \end{bmatrix} \begin{Bmatrix} 0 \\ n \\ 0 \end{Bmatrix} = \begin{Bmatrix} n\phi_{\text{yaw}} \\ n \\ -n\psi_{\text{roll}} \end{Bmatrix}$$

Now we can calculate the components of the satellite's inertial angular velocity along the body frame axes,

$$\begin{aligned} \omega_x &= \omega_{x_{\text{rel}}} + \Omega_x = \dot{\psi}_{\text{roll}} + n\phi_{\text{yaw}} \\ \omega_y &= \omega_{y_{\text{rel}}} + \Omega_y = \dot{\theta}_{\text{pitch}} + n \\ \omega_z &= \omega_{z_{\text{rel}}} + \Omega_z = \dot{\phi}_{\text{yaw}} - n\psi_{\text{roll}} \end{aligned} \tag{8.1.162}$$

Differentiating these with respect to time, remembering that n is constant for a circular orbit, gives the components of inertial angular acceleration in the body frame,

$$\begin{aligned} \dot{\omega}_x &= \ddot{\psi}_{\text{roll}} + n\dot{\phi}_{\text{yaw}} \\ \dot{\omega}_y &= \ddot{\theta}_{\text{pitch}} \\ \dot{\omega}_z &= \ddot{\phi}_{\text{yaw}} - n\dot{\psi}_{\text{roll}} \end{aligned} \tag{8.1.163}$$

The position vector of the satellite's center of mass lies along the z' axis of the orbital frame,

$$\mathbf{r}_0 = r_0\hat{\mathbf{k}}'$$

To obtain the components of $\mathbf{r}_0$ in the body frame we once again use the transformation matrix $[\mathbf{Q}]_{x'x}$

$$\begin{Bmatrix} r_{0_x} \\ r_{0_y} \\ r_{0_z} \end{Bmatrix} = \begin{bmatrix} 1 & \phi_{\text{yaw}} & -\theta_{\text{pitch}} \\ -\phi_{\text{yaw}} & 1 & \psi_{\text{roll}} \\ \theta_{\text{pitch}} & -\psi_{\text{roll}} & 1 \end{bmatrix} \begin{Bmatrix} 0 \\ 0 \\ r_0 \end{Bmatrix} = \begin{Bmatrix} -r_0\theta_{\text{pitch}} \\ r_0\psi_{\text{roll}} \\ r_0 \end{Bmatrix} \tag{8.1.164}$$

Substituting Equations 8.1.162, 8.1.163 and 8.1.164 into Equations 8.1.157, and setting

$$A = I_{\text{roll}} \qquad B = I_{\text{pitch}} \qquad C = I_{\text{yaw}} \tag{8.1.165}$$

yields

$$I_{\text{roll}}(\ddot{\psi}_{\text{roll}} + n\dot{\phi}_{\text{yaw}}) + (I_{\text{yaw}} - I_{\text{pitch}})(\dot{\theta}_{\text{pitch}} + n)(\dot{\phi}_{\text{yaw}} - n\psi_{\text{roll}}) = 3(I_{\text{yaw}} - I_{\text{pitch}})n^2\psi_{\text{roll}}$$
$$I_{\text{pitch}}\ddot{\theta}_{\text{pitch}} + (I_{\text{roll}} - I_{\text{yaw}})(\dot{\psi}_{\text{roll}} + n\phi_{\text{yaw}})(\dot{\phi}_{\text{yaw}} - n\psi_{\text{roll}}) = -3(I_{\text{roll}} - I_{\text{yaw}})n^2\theta_{\text{pitch}}$$
$$I_{\text{yaw}}(\ddot{\phi}_{\text{yaw}} - n\dot{\psi}_{\text{roll}}) + (I_{\text{pitch}} - I_{\text{roll}})(\dot{\theta}_{\text{pitch}} + n)(\dot{\psi}_{\text{roll}} + n\phi_{\text{yaw}}) = -3(I_{\text{pitch}} - I_{\text{roll}})n^2\theta_{\text{pitch}}\psi_{\text{roll}}$$

Expanding terms and retaining terms at most linear in all angular quantities and their rates yields

$$I_{\text{yaw}}\ddot{\phi}_{\text{yaw}} + (I_{\text{pitch}} - I_{\text{roll}})n^2\phi_{\text{yaw}} + (I_{\text{pitch}} - I_{\text{roll}} - I_{\text{yaw}})n\dot{\psi}_{\text{roll}} = 0 \tag{8.1.166}$$

$$I_{\text{roll}}\ddot{\psi}_{\text{roll}} + (I_{\text{roll}} - I_{\text{pitch}} + I_{\text{yaw}})n\dot{\phi}_{\text{yaw}} + 4(I_{\text{pitch}} - I_{\text{yaw}})n^2\psi_{\text{roll}} = 0 \tag{8.1.167}$$

$$I_{\text{pitch}}\ddot{\theta}_{\text{pitch}} + 3(I_{\text{roll}} - I_{\text{yaw}})n^2\theta_{\text{pitch}} = 0 \tag{8.1.168}$$

These are the differential equations governing the influence of gravity gradient torques on the small angles and rates of misalignment of the body frame with the orbital frame.

Equation 8.1.168, governing the pitching motion around the y' axis, is not coupled to the other two equations. We make the classical assumption that the solution is of the form

$$\theta_{\text{pitch}} = Pe^{pt} \tag{8.1.169}$$

where P and p are constants. P is the amplitude of the small disturbance that initiates the pitching motion. Substituting Equation 8.1.169 into Equation 8.1.168 yields $[I_{\text{pitch}}p^2 + 3(I_{\text{roll}} - I_{\text{yaw}})n^2]Pe^{pt} = 0$ for all t, which implies that the bracketed term must vanish, and that means p must have either of the two values

$$p_{1,2} = \pm i\sqrt{3\frac{(I_{\text{roll}} - I_{\text{yaw}})n^2}{I_{\text{pitch}}}} \quad (i = \sqrt{-1})$$

Thus

$$\theta_{\text{pitch}} = P_1e^{p_1t} + P_2e^{p_2t}$$

yields the stable, small-amplitude, steady-state harmonic oscillator solution only if p_1 and p_2 are imaginary, that is, if

$$I_{\text{roll}} > I_{\text{yaw}} \qquad \text{For stability in pitch.} \tag{8.1.170}$$

The stable pitch oscillation frequency is

$$\omega_{f_{\text{pitch}}} = n\sqrt{3\frac{(I_{\text{roll}} - I_{\text{yaw}})}{I_{\text{pitch}}}} \tag{8.1.171}$$

(If $I_{\text{yaw}} > I_{\text{roll}}$, then p_1 and p_2 are both real, one positive, the other negative. The positive root causes $\theta_{\text{pitch}} \to \infty$, which is the undesirable, unstable case.)

Let us now turn our attention to Equations 8.1.166 and 8.1.167, which govern yaw and roll motion under gravity gradient torque. Again, we assume the solution is exponential in form,

$$\phi_{\text{yaw}} = Ye^{qt} \qquad \psi_{\text{roll}} = Re^{qt} \tag{8.1.172}$$

Substituting these into Equations 8.1.166 and 8.1.167 yields

$$[(I_{\text{pitch}} - I_{\text{roll}})n^2 + I_{\text{yaw}}q^2]Y + (I_{\text{pitch}} - I_{\text{roll}} - I_{\text{yaw}})nqR = 0$$
$$(I_{\text{roll}} - I_{\text{pitch}} + I_{\text{yaw}})nqY + [4(I_{\text{pitch}} - I_{\text{yaw}})n^2 + I_{\text{roll}}q^2]R = 0$$

In the interest of simplification, we can factor I_{yaw} out of the first equation and I_{roll} out of the second one to get

$$\left(\frac{I_{\text{pitch}} - I_{\text{roll}}}{I_{\text{yaw}}}n^2 + q^2\right)Y + \left(\frac{I_{\text{pitch}} - I_{\text{roll}}}{I_{\text{yaw}}} - 1\right)nqR = 0$$
$$\left(1 - \frac{I_{\text{pitch}} - I_{\text{yaw}}}{I_{\text{roll}}}\right)nqY + \left(4\frac{I_{\text{pitch}} - I_{\text{yaw}}}{I_{\text{roll}}}n^2 + q^2\right)R = 0 \tag{8.1.173}$$

Let

$$k_Y = \frac{I_{\text{pitch}} - I_{\text{roll}}}{I_{\text{yaw}}} \qquad k_R = \frac{I_{\text{pitch}} - I_{\text{yaw}}}{I_{\text{roll}}} \tag{8.1.174}$$

It is easy to show from Equations 8.1.155, 8.1.165 and 8.1.174 that

$$k_Y = \frac{\left(\int_m x^2 dm \Big/ \int_m y^2 dm\right) - 1}{\left(\int_m x^2 dm \Big/ \int_m y^2 dm\right) + 1} \qquad k_R = \frac{\left(\int_m z^2 dm \Big/ \int_m y^2 dm\right) - 1}{\left(\int_m z^2 dm \Big/ \int_m y^2 dm\right) + 1}$$

which means

$$|k_Y| < 1 \quad |k_R| < 1$$

Using the definitions in Equation 8.1.174, we can write Equations 8.1.173 more compactly as

$$(k_Y n^2 + q^2)Y + (k_Y - 1)nqR = 0$$
$$(1 - k_R)nqY + (4k_R n^2 + q^2)R = 0$$

or, using matrix notation,

$$\begin{bmatrix} k_Y n^2 + q^2 & (k_Y - 1)nq \\ (1 - k_R)nq & 4k_R n^2 + q^2 \end{bmatrix} \begin{Bmatrix} Y \\ R \end{Bmatrix} = \begin{Bmatrix} 0 \\ 0 \end{Bmatrix} \tag{8.1.175}$$

In order to avoid the trivial solution *(Y = R = 0)*, the determinant of the coefficient matrix must be zero. Expanding the determinant and collecting terms yields the characteristic equation for q,

$$q^4 + bn^2q^2 + cn^4 = 0 \tag{8.1.176}$$

where

$$b = 3k_R + k_Y k_R + 1 \qquad c = 4k_Y k_R \tag{8.1.177}$$

This quadratic equation has four roots which, when substituted back into Equation 8.1.172, yield

$$\phi_{\text{yaw}} = Y_1 e^{q_1 t} + Y_2 e^{q_2 t} + Y_3 e^{q_3 t} + Y_4 e^{q_4 t}$$
$$\psi_{\text{roll}} = R_1 e^{q_1 t} + R_2 e^{q_2 t} + R_3 e^{q_3 t} + R_4 e^{q_4 t}$$

In order for these solutions to remain finite in time, the roots $q_1,\ldots,q_4$ must be negative (solution decays to zero) or imaginary (steady oscillation at initial small amplitude).

To reduce Equation 8.1.176 to a quadratic equation, let us introduce a new variable λ and write,

$$q = \pm n\sqrt{\lambda} \tag{8.1.178}$$

Then Equation 8.1.176 becomes

$$\lambda^2 + b\lambda + c = 0 \tag{8.1.179}$$

the familiar solution of which is

$$\lambda_1 = -\frac{1}{2}(b + \sqrt{b^2 - 4c}) \qquad \lambda_2 = -\frac{1}{2}(b - \sqrt{b^2 - 4c}) \tag{8.1.180}$$

To guarantee that q in Equation 8.1.178 does not take a positive value, we must require that λ be real and negative (so q will be imaginary). For λ to be real requires that $b > 2\sqrt{c}$, or

$$3k_R + k_Y k_R + 1 > 4\sqrt{k_Y k_R} \tag{8.1.181}$$

For λ to be negative requires $b^2 > b^2 - 4c$, which will be true if $c > 0$; i.e.,

$$k_Y k_R > 0 \tag{8.1.182}$$

Equations 8.1.181 and 8.1.182 are the conditions required for yaw and roll stability under gravity gradient torques, to which we must add Equation 8.1.170 for pitch stability. Observe that we can solve Equations 8.1.174 to obtain

$$I_{\text{yaw}} = \frac{1 - k_R}{1 - k_Y k_R} I_{\text{pitch}} \qquad I_{\text{roll}} = \frac{1 - k_Y}{1 - k_Y k_R} I_{\text{pitch}}$$

By means of these relationships, the pitch stability criterion, $I_{\text{roll}}/I_{\text{yaw}} > 1$, becomes

$$\frac{1 - k_Y}{1 - k_R} > 1$$

In view of the fact that $|k_R| < 1$, this means

$$k_Y < k_R \tag{8.1.183}$$

Figure 8.1-29 shows those regions *I* and *II* on the k_y–k_R plane in which all three stability criteria (Equations 8.1.181, 8.1.182 and 8.1.183) are simultaneously satisfied, along with the requirement that the three moments of inertia I_{pitch}, I_{roll} and I_{yaw} are positive.

In the small sliver of region *I*, and $k_Y < 0$ and $k_R < 0$; therefore, according to Equations 8.11.174, $I_{\text{yaw}} > I_{\text{pitch}}$ and $I_{\text{roll}} > I_{\text{pitch}}$ which together with Equation 8.11.170 yield $I_{\text{roll}} > I_{\text{yaw}} > I_{\text{pitch}}$. Remember that the gravity gradient spacecraft is slowly 'spinning' about the minor pitch axis (normal to the orbit plane) at an angular velocity equal to the mean motion of the orbit. So this criterion makes the spacecraft a 'minor axis spinner', the roll axis (flight direction) being the major axis of inertia. With energy dissipation, we know this orientation is not stable in the long run. On the other hand, in region *II*, k_Y and k_R are both positive, so that Equations 8.11.174 imply $I_{\text{pitch}} > I_{\text{yaw}}$ and $I_{\text{pitch}} > I_{\text{roll}}$. Thus, along with the pitch criterion ($I_{\text{roll}} > I_{\text{yaw}}$), we have $I_{\text{pitch}} > I_{\text{roll}} > I_{\text{yaw}}$. In this, the preferred, configuration, the gravity gradient spacecraft is a 'major axis spinner' about the pitch axis, and the minor yaw axis is the minor axis of inertia. It turns out that all of the known gravity-gradient stabilized moons of the solar system, like the earth's, whose 'captured' rate of rotation equals the orbital period, are major axis spinners.

In Equation 8.1.171 we presented the frequency of the gravity gradient pitch oscillation. For completeness we should also point out that the coupled yaw and roll motions have two oscillation frequencies, which are obtained from Equations 8.1.178 and 8.1.180,

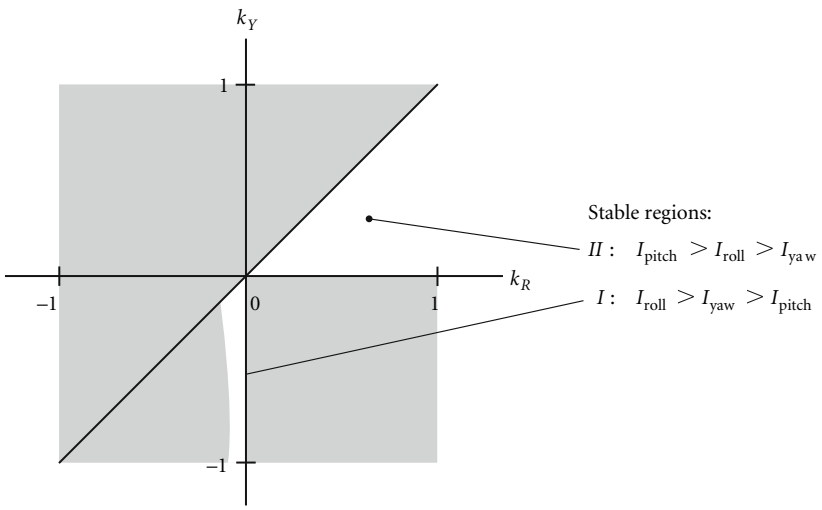


Figure 8.1-29 Regions in which the values of k_Y and k_R yield neutral stability in yaw, pitch and roll of a gravity gradient satellite.

$$\omega f_{\text{yaw/roll}})1, 2 = n\sqrt{\frac{1}{2}(b \pm \sqrt{b^2 - 4c})} \tag{8.1.184}$$

Recall that b and c are found in Equation 8.1.177.

We have assumed throughout this discussion that the orbit of the gravity gradient satellite is circular. Kaplan (1976) has shown that the effect of a small eccentricity turns up only in the pitching motion. In particular, the natural oscillation expressed by Equation 8.1.170 is augmented by a forced oscillation term,

$$\theta_{\text{pitch}} = P_1 e^{p_1 t} + P_2 e^{p_2 t} + \frac{2e\sin nt}{3\left(\dfrac{I_{\text{roll}} - I_{\text{yaw}}}{I_{\text{pitch}}}\right) - 1} \tag{8.1.185}$$

where e is the (small) eccentricity of the orbit. From this we see that there is a pitch resonance. When $(I_{\text{roll}} - I_{\text{yaw}})/I_{\text{pitch}}$ approaches 1/3, the amplitude of the last term grows without bound.

Example 8.1.14

The uniform, monolithic 10 000 kg slab, having the dimensions shown in Figure 8.1-30, is in a circular LEO. Determine the orientation of the satellite in its orbit for gravity gradient stabilization, and compute the periods of the pitch and yaw/roll oscillations in terms of the orbital period T.

The principal moments of inertia around the xyz axes through the center of mass are

$$A = \frac{10\,000}{12}(1^2 + 9^2) = \ 68\,333\ \text{kg}\cdot\text{m}^2$$

$$B = \frac{10\,000}{12}(3^2 + 9^2) = 75\,000\ \text{kg}\cdot\text{m}^2$$

$$C = \frac{10\,000}{12}(3^2 + 1^2) = \ 8333.3\ \text{kg}\cdot\text{m}^2$$

Let us first determine whether we can stabilize this object as a minor axis spinner. In that case,

$$I_{\text{pitch}} = C = 8333.3\ \text{kg}\cdot\text{m}^2 \qquad I_{\text{yaw}} = A = 68\,333\ \text{kg}\cdot\text{m}^2 \qquad I_{\text{roll}} = \text{B} = 75\,000\ \text{kg}\cdot\text{m}^2$$

Since $I_{\text{roll}} > I_{\text{yaw}}$, the satellite would be stable in pitch. To check yaw/roll stability, we first compute

$$k_Y = \frac{I_{\text{pitch}} - I_{\text{roll}}}{I_{\text{yaw}}} = -0.97561 \qquad k_R = \frac{I_{\text{pitch}} - I_{\text{yaw}}}{I_{\text{roll}}} = -0.8000$$

We see that $k_Y k_R > 0$, which is one of the two requirements. The other one is found in Equation 8.1.181, but in this case

$$1 + 3k_R + k_Y k_R - 4\sqrt{k_Y k_R} = -4.1533 < 0$$

so that condition is not met. Hence, the object cannot be gravity-gradient stabilized as a minor axis spinner.

As a major axis spinner, we must have

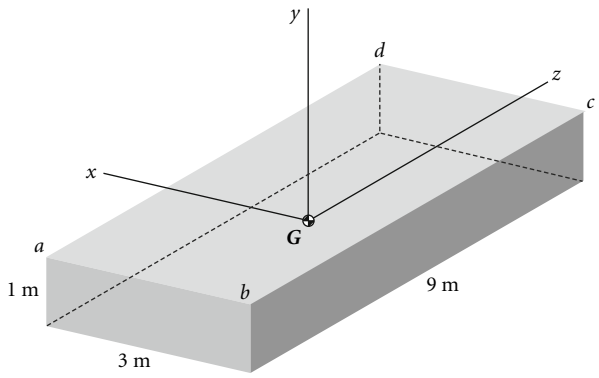


Figure 8.1-30 Parallelepiped satellite.

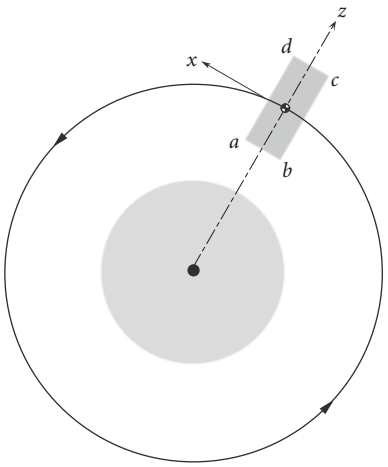


Figure 8.1-31 Orientation of the parallelepiped for gravity-gradient stabilization.

$$I_{\text{pitch}} = B = 75\,000 \text{ kg}\cdot\text{m}^2 \qquad I_{\text{yaw}} = C = 8333.3 \text{ kg}\cdot\text{m}^2 \qquad I_{\text{roll}} = A = 68\,333 \text{ kg}\cdot\text{m}^2$$

Then $I_{\text{roll}} > I_{\text{yaw}}$, so the pitch stability condition is satisfied. Furthermore, since

$$k_Y = \frac{I_{\text{pitch}} - I_{\text{roll}}}{I_{\text{yaw}}} = 0.8000 \qquad k_R = \frac{I_{\text{pitch}} - I_{\text{yaw}}}{I_{\text{roll}}} = 0.97561$$

we have

$$k_Y k_R = 0.7805 > 0$$
$$1 + 3k_R + k_Y k_R - 4\sqrt{k_Y k_R} = 1.1735 > 0$$

which means the two criteria for stability in the yaw and roll modes are met. The satellite should therefore be orbited as shown in Figure 8.1-31, with its minor axis aligned with the radial from the earth's center, the plane *abcd* lying in the orbital plane, and the body x axis aligned with the local horizon.

According to Equation 8.1.171, the frequency of the pitch oscillation is

$$\omega_{f_{\text{pitch}}} = n\sqrt{3\frac{I_{\text{roll}} - I_{\text{yaw}}}{I_{\text{pitch}}}} = n\sqrt{3\frac{68\,333 - 8333.3}{75\,000}} = 1.5492n$$

where n is the mean motion. Hence, the period of this oscillation, in terms of that of the orbit, is

$$T_{\text{pitch}} = \frac{2\pi}{\omega_{f_{\text{pitch}}}} = 0.6455\frac{2\pi}{n} = \underline{0.6455T}$$

For the yaw/roll frequencies, we use Equation 8.1.184,

$$\omega_{f_{\text{yaw/roll}}}\Big)_1 = n\sqrt{\frac{1}{2}\left(b + \sqrt{b^2 - 4c}\right)}$$

where

$$b = 1 + 3k_R + k_Y k_R = 4.7073 \quad \text{and} \quad c = 4k_Y k_R = 3.122$$

Thus,

$$\omega_{f_{\text{yaw/roll}}}\Big)_1 = 2.3015n$$

Likewise,

$$\omega_{f_{\text{yaw/roll}}}\Big)_2 = \sqrt{\frac{1}{2}\left(b - \sqrt{b^2 - 4c}\right)} = 1.977n$$

From these we obtain

$$T_{\text{yaw/roll}_1} = \underline{0.5058T} \qquad T_{\text{yaw/roll}_2} = \underline{0.4345T}$$

Finally, observe that

$$\frac{I_{\text{roll}} - I_{\text{yaw}}}{I_{\text{pitch}}} = 0.8$$

so that we are far from the pitch resonance condition that exists if the orbit has a small eccentricity.

Problems

8.1.1 The axisymmetric satellite has axial and transverse mass moments of inertia about axes through the mass center G of $C = 1200$ kg · m^2 and $A = 2600$ kg · m^2, respectively. If it is spinning at $\omega_s = 6$ rad/s when it is launched, determine its angular momentum. Precession occurs about the inertial Z axis.
{Ans.: $||\mathbf{H}_G|| = 13\,450$ kg · m^2/s}

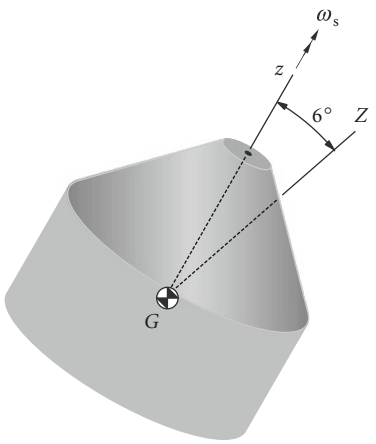


Figure P.8.1.1

8.1.2 A spacecraft is symmetrical about its body-fixed z axis. Its principal mass moments of inertia are $A = B = 300$ kg · m^2 and $C = 500$ kg · m^2. The z axis sweeps out a cone with a total vertex angle of 10° as it precesses around the angular momentum vector. If the spin velocity is 6 rad/s, compute the period of precession.
{Ans.: 0.417 s}

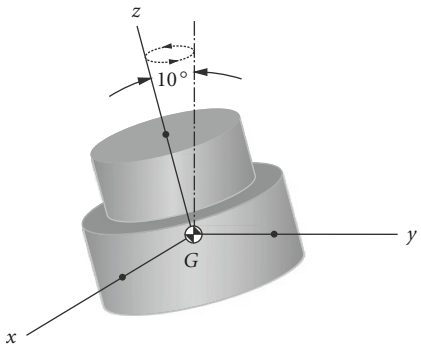


Figure P.8.1.2

8.1.3 A thin ring tossed into the air with a spin velocity of ω_s has a very small nutation angle θ (in radians). What is the precession rate ω_p?

{Ans.: $\omega_p = 2\omega_s\ (1 + \theta^2/2)$, retrograde}

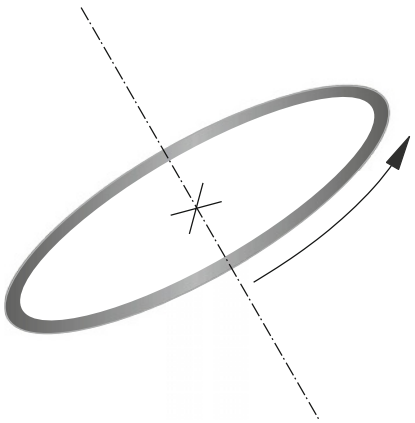


Figure P.8.1.3

8.1.4 For an axisymmetric rigid satellite,

$$[\mathrm{I}_G] = \begin{bmatrix} I_{xx} & 0 & 0 \\ 0 & I_{yy} & 0 \\ 0 & 0 & I_{zz} \end{bmatrix} \begin{bmatrix} 1000 & 0 & 0 \\ 0 & 1000 & 0 \\ 0 & 0 & 5000 \end{bmatrix} \mathrm{kg \cdot m^2}$$

It is spinning about the body z axis in torque-free motion, precessing around the angular momentum vector $\mathbf{H}$ at the rate of 2 rad/s. Calculate the magnitude of $\mathbf{H}$.

{Ans.: 2000 N · m · s}

8.1.5 At a given instant the box-shaped 500 kg satellite (in torque-free motion) has an absolute angular velocity $\omega = 0.01\hat{\mathbf{i}} - 0.03\hat{\mathbf{j}} + 0.2\hat{\mathbf{k}}$ (rad/s). Its moments of inertia about the principal body axes xyz are $A = 385.4$ kg · m^2, $B = 416.7$ kg · m^2 and $C = 52.08$ kg · m^2, respectively. Calculate the magnitude of its absolute angular acceleration.

{Ans.: 6.167×10^{-4} m/s^2}

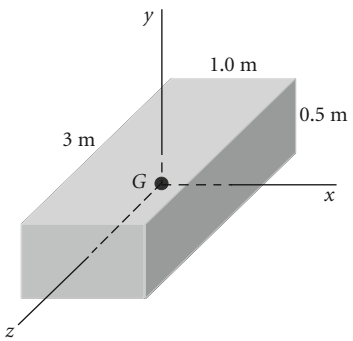


Figure P.8.1.5

8.1.6 An 8 kg thin ring in torque-free motion is spinning with an angular velocity of 30 rad/s and a constant nutation angle of 15°. Calculate the rotational kinetic energy if $A = B = 0.36$ kg · m^2, $C = 0.72$ kg · m^2.

{Ans.: 370.5 J}

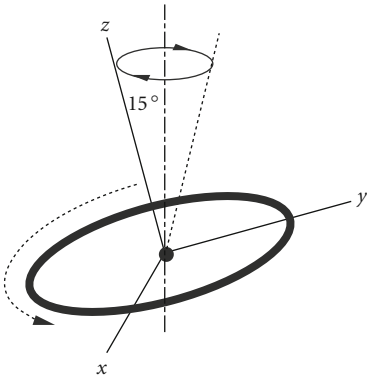


Figure P.8.1.6

8.1.7 The rectangular block has an angular velocity $\boldsymbol{\omega} = 1.5\omega_0\hat{\mathbf{i}} + 0.8\omega_0\hat{\mathbf{j}} + 0.6\omega_0\hat{\mathbf{k}}$,where ω_0 has units of rad/s.

(a) Determine the angular velocity ω of the block if it spins around the body z axis with the same rotational kinetic energy.

(b) Determine the angular velocity ω of the block if it spins around the body z axis with the same angular momentum.

{Ans.: (a) $\omega = 1.31\omega_0$, (b) $\omega = 1.04\omega_0$}

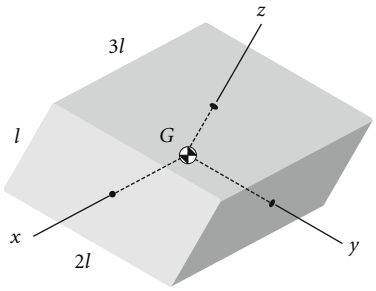


Figure P.8.1.7

8.1.8 For a rigid axisymmetric satellite, the mass moment of inertia about its long axis is 1000 kg · m^2, and the moment of inertia about transverse axes through the centroid is 5000 kg · m^2. It is spinning about the minor principal body axis in torque-free motion at 6 rad/s with the angular velocity lined up with the angular momentum vector **H**. Over time, the energy degrades due to internal effects and the satellite is eventually spinning about a major principal body axis with the angular velocity lined up with the angular momentum vector **H**. Calculate the change in rotational kinetic energy between the two states.

{Ans.: –14.4 kJ}

8.1.9 Consider a highly dissipative torque-free satellite, whose angular velocity at the instant shown is $\boldsymbol{\omega} = 10\hat{\mathbf{i}}$ rad/s. Calculate the decrease in kinetic energy after it becomes, as eventually it must, a major axis spinner.

{Ans.: –0.487 J}

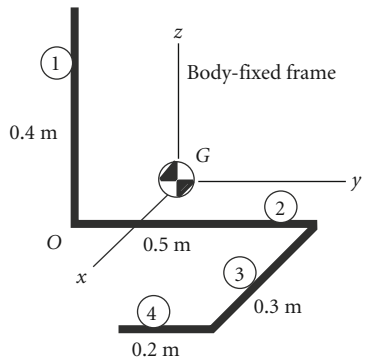


Figure P.8.1.9

8.1.10 For a non-precessing, dual-spin satellite, $C_r = 1000$ kg · m^2 and $C_p = 500$ kg · m^2. The angular velocity of the rotor is 3 $\widehat{\mathbf{k}}$rad/s and the angular velocity of the platform relative to the rotor is $1\widehat{\mathbf{k}}$ rad/s. If the relative angular velocity of the platform is reduced to $0.5\widehat{\mathbf{k}}$ rad/s, what is the new angular velocity of the rotor?
{Ans.: 3.17 rad/s}

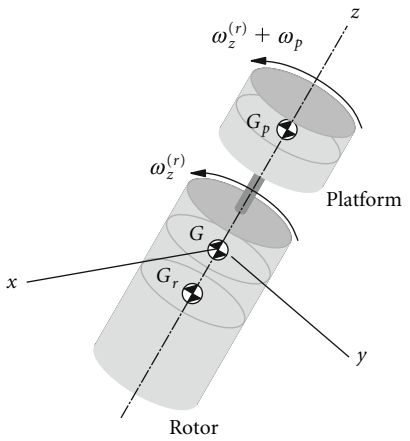


Figure P.8.1.10

8.1.11 For a rigid axisymmetric satellite, the mass moment of inertia about its long axis is 1000 kg · m^2, and the moment of inertia about transverse axes through the center of mass is 5000 kg · m^2. It is initially spinning about the *minor* principal body axis in torque-free motion at $\omega_s = 0.1$ rad/s, with the angular velocity lined up with the angular momentum vector $\mathbf{H}_0$. A pair of thrusters exert an external impulsive torque on the satellite, causing an instantaneous change $\mathbf{\Delta H}$ of angular momentum in the direction normal to $\mathbf{H}_0$ (no change in spin rate), so that the new angular momentum is $\mathbf{H}_1$, at an angle of 20° to $\mathbf{H}_0$, as shown in the figure. How long does it take the satellite to precess ('cone') through an angle of 180° around $\mathbf{H}_1$?
{Ans.: 118 s}

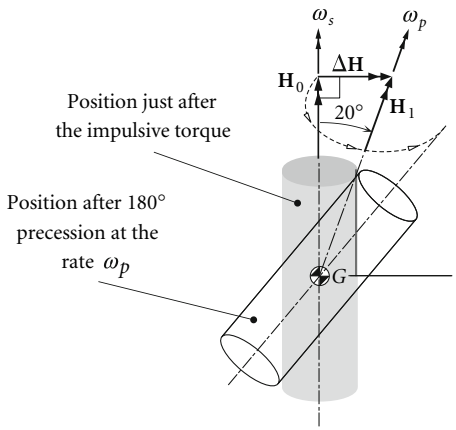


Figure P.8.1.11

8.1.12 The solid right-circular cylinder of mass 500 kg is set into torque-free motion with its symmetry axis initially aligned with the fixed spatial line a–a. Due to an injection error, the vehicle's angular velocity vector ω is misaligned 5° (the wobble angle) from the symmetry axis. Calculate to three significant figures the maximum angle ϕ between fixed line a–a and the axis of the cylinder.
{Ans.: 31°}

8.1.13 A satellite is spinning at 0.01 rev/s. The moment of inertia of the satellite about the spin axis is 2000 kg · m^2. Paired thrusters are located at a distance of 1.5 m from the spin axis. They deliver their thrust in pulses, each thruster producing an impulse of 15 N · s per pulse. At what rate will the satellite be spinning after 30 pulses?
{Ans.: 0.0637 rev/s}

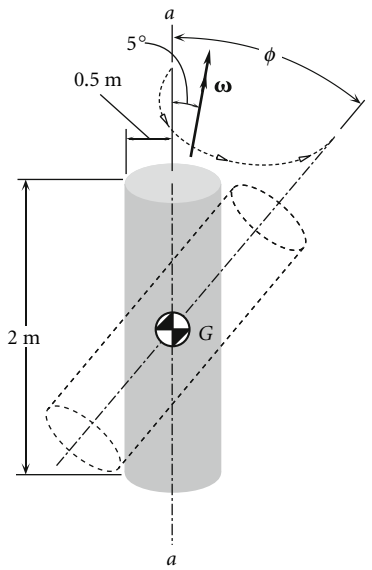


Figure P.8.1.12

8.1.14 A satellite has moments of inertia $A = 2000$ kg · m^2, $B = 4000$ kg · m^2 and $C = 6000$ kg · m^2 about its principal body axes xyz. Its angular velocity is $\boldsymbol{\omega} = 0.1\hat{\mathbf{i}} + 0.3\hat{\mathbf{j}} + 0.5\hat{\mathbf{k}}$ (rad/s). If thrusters cause the angular momentum vector to undergo the change $\Delta\mathbf{H}_G = 50\hat{\mathbf{i}} - 100\hat{\mathbf{j}} + 3000\hat{\mathbf{k}}$(kg · m^2/s), what is the magnitude of the new angular velocity?

{Ans.: 0.628 rad/s}

8.1.15 The body-fixed xyz axes are principal axes of inertia passing through the center of mass of the 300 kg cylindrical satellite, which is spinning at 1 revolution per second about the z axis. What impulsive torque about the y axis must the thrusters impart to cause the satellite to precess at 0.1 revolution per second?

{Ans.: 137 N · m · s}

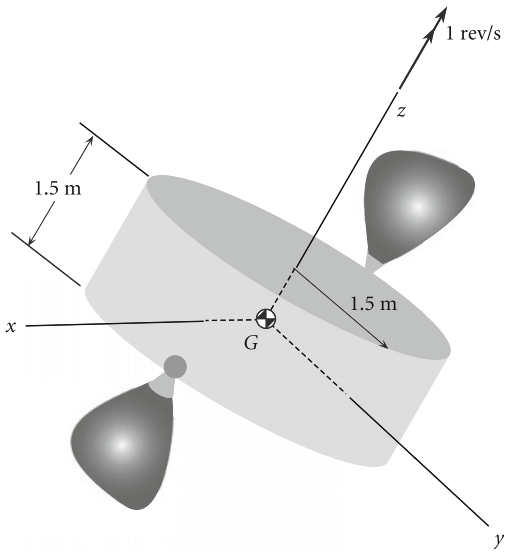


Figure P.8.1.15

8.1.16 A satellite is to be despun by means of a tangential-release yo-yo mechanism consisting of two masses, 3 kg each, wound around the mid plane of the satellite. The satellite is spinning around its axis of symmetry with an angular velocity $\omega_s = 5$ rad/s. The radius of the cylindrical satellite is 1.5 m and the moment of inertia about the spin axis is $C = 300$ kg · m^2.

(a) Find the cord length and the deployment time to reduce the spin rate to 1 rad/s.
(b) Find the cord length and time to reduce the spin rate to zero.

{Ans.: (a) $l = 5.902$ m, $t = 0.787$ s; (b) $l = 7.228$ m, $t = 0.964$ s}

8.1.17 A cylindrical satellite of radius 1 m is initially spinning about the axis of symmetry at the rate of 2 revolutions per second with a nutation angle of 15°. The principal moments of inertia are $I_x = I_y = 30$ kg · m^2, $I_z = 60$ kg · m^2. An energy dissipation device is built into the satellite, so that it eventually ends up in pure spin around the z axis.

(a) Calculate the final spin rate about the z axis.
(b) Calculate the loss of kinetic energy.
(c) A tangential release yo-yo despin device is also included in the satellite. If the two yo-yo masses are each 7 kg, what cord length is required to completely despin the satellite? Is it wrapped in the proper direction in the figure?

{Ans.: (a) 2.071 rad/s; (b) 8.62 J; (c) 2.3 m}

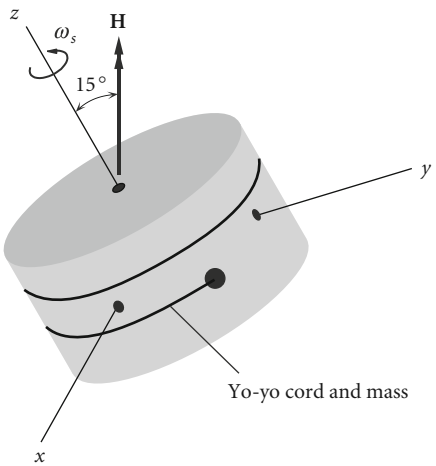


Figure P.8.1.17

8.1.18 A communications satellite is in a geostationary equatorial orbit (GEO) with a period of 24 hours. The spin rate ω_s about its axis of symmetry is 1 revolution per minute, and the moment of inertia about the spin axis is 550 kg · m^2. The moment of inertia about transverse axes through the mass center G is 225 kg · m^2. If the spin axis is initially pointed towards the earth, calculate the magnitude and direction of the applied torque $\mathbf{M}_G$ required to keep the spin axis pointed always towards the earth.
{Ans.: 0.00420 N · m, about the negative x axis}

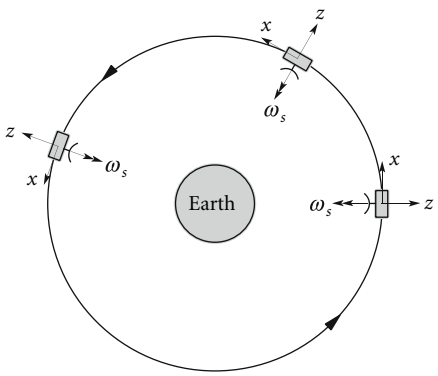


Figure P.8.1.18

8.1.19 The moments of inertia of a satellite about its principal body axes xyz are $A = 1000$ kg · m^2, $B = 600$ kg · m^2 and $C = 500$ kg · m^2, respectively. The moments of inertia of a momentum wheel at the center of mass of the satellite and aligned with the x axis are $I_x = 20$ kg · m^2 and $I_y = I_z = 6$ kg · m^2. The absolute angular velocity of the satellite with the momentum wheel locked is $\boldsymbol{\omega}_0 = 0.1\hat{\mathbf{i}} + 0.5\hat{\mathbf{j}}$ (rad/s). Calculate the angular velocity ω_f of the momentum wheel (relative to the satellite) required to reduce the x component of the absolute angular velocity of the satellite to 0.003 rad/s.
{Ans.: 4.95 rad/s}

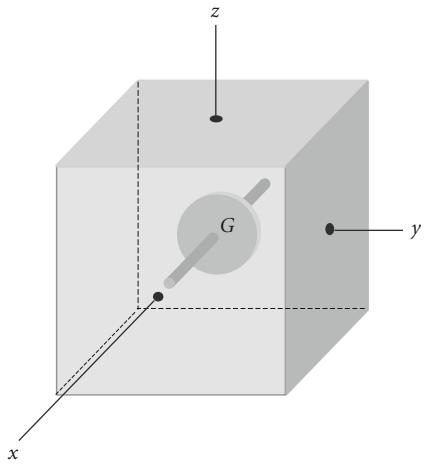


Figure P.8.1.19

8.1.20 A satellite has principal moments of inertia $I_1 = 300$ kg · m^2, $I_2 = 400$ kg · m^2, $I_3 = 500$ kg · m^2. Determine the permissible orientations in a circular orbit for gravity-gradient stabilization. Specify which axes may be aligned in the pitch, roll and yaw directions. (Recall that, relative to a Clohessy–Wiltshire frame at the center of mass of the satellite, yaw is about the x axis (outward radial from earth's center); roll is about the y axis (velocity vector); pitch is about the z axis (normal to orbital plane).

Aerospace Engineering Desk Reference; ISBN 9781856175753

8.2 Rocket Vehicle Dynamics

Howard D. Curtis

8.2.1 Introduction

The delta-v maneuvers of spacecraft require a propulsion system of some sort whose job it is to throw vehicle mass (in the form of propellants) overboard. Newton's balance of momentum principle dictates that when mass is ejected from a system in one direction, the mass left behind must acquire a velocity in the opposite direction. The familiar and oft-quoted example is the rapid release of air from an inflated toy balloon. Another is that of a diver leaping off a small boat at rest in the water, causing the boat to acquire a motion of its own. The unfortunate astronaut who becomes separated from his ship in the vacuum of space cannot with any amount of flailing of arms and legs 'swim' back to safety. If he has tools or other expendable objects of equipment, accurately throwing them in the direction opposite to his spacecraft may do the trick. Spewing compressed gas from a tank attached to his back through to a nozzle pointed away from the spacecraft would be a better solution.

The purpose of a rocket motor is to use the chemical energy of solid or liquid pro-pellants to steadily and rapidly produce a large quantity of hot, high pressure gas which is then expanded and accelerated through a nozzle. This large mass of combustion products flowing out of the nozzle at supersonic speed possesses a lot of momentum and, leaving the vehicle behind, causes the vehicle itself to acquire a momentum in the opposite direction. This is represented as the action of the force we know as thrust. The design and analysis of rocket propulsion systems is well beyond our scope.

This chapter contains a necessarily brief introduction to some of the fundamentals of rocket vehicle dynamics. The equations of motion of a launch vehicle in a gravity turn trajectory are presented first. This is followed by a simple development of the thrust equation, which brings in the concept of specific impulse. The thrust equation and the equations of motion are then combined to produce the rocket equation, which relates delta-v to propellant expenditure and specific impulse. The sounding rocket provides an important but relatively simple application of the concepts introduced to this point. The chapter concludes with an elementary consideration of multi-stage launch vehicles.

Those seeking a more detailed introduction to the subject of rockets and rocket performance will find the texts by Wiesel (1997) and Hale (1994), as well as references cited therein, useful.

8.2.2 Equations of Motion

Figure 8.2-1 illustrates the trajectory of a satellite launch vehicle and the forces acting on it during the powered ascent. Rockets at the base of the booster produce the thrust $\mathbf{T}$ which acts along the vehicle's axis in the direction of the velocity vector $\mathbf{v}$. The aerodynamic drag force $\mathbf{D}$ is directed opposite to the velocity, as shown. Its magnitude is given by

$$D = qAC_D \tag{8.2.1}$$

where $q = \frac{1}{2}\varrho v^2$ is the dynamic pressure, in which ϱ is the density of the atmosphere and v is the speed, i.e., the magnitude of $\mathbf{v}$. A is the frontal area of the vehicle and C_D is the coefficient of drag. C_D depends on the speed and the external geometry of the rocket. The force of gravity on the booster is $m\boldsymbol{g}$, where m is its mass and $\boldsymbol{g}$ is the local gravitational acceleration, pointing towards the center of the earth. At any point of the trajectory, the velocity $\mathbf{v}$ defines the direction of the unit tangent $\hat{\mathbf{u}}_t$ to the path. The unit normal $\hat{\mathbf{u}}_n$ is perpendicular to $\mathbf{v}$ and points towards the center of curvature C. The distance of point C from the path is ϱ (not to be confused with density), ϱ is the radius of curvature.

In Figure 8.2-1 the vehicle and its flight path are shown relative to the earth. In the interest of simplicity we will ignore the earth's spin and write the equations of motion relative to a non-rotating earth. The small acceleration terms required to account for the earth's rotation can be added for a more refined analysis. Let us resolve Newton's second law, $\mathbf{F}_{\text{net}} = m\mathbf{a}$, into components along the path directions $\hat{\mathbf{u}}_t$ and $\hat{\mathbf{u}}_n$. The acceleration along the path is

$$a_t = \frac{dv}{dt} \tag{8.2.2}$$

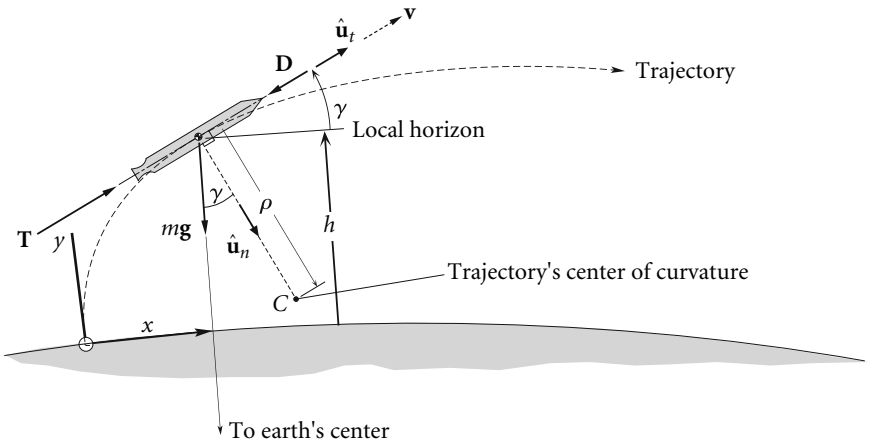


Figure 8.2-1 Launch vehicle boost trajectory, γ is the flight path angle.

and the normal acceleration is $a_n = v^2/\varrho$ (where ϱ is the radius of curvature). For flight over a flat surface, $v/\varrho = -d\gamma/dt$, in which case the normal acceleration can be expressed in terms of the flight path angle as

$$a_n = -v\frac{d\gamma}{dt}$$

To account for the curvature of the earth, one can use polar coordinates with origin at the earth's center to show that a term must be added to this expression, so that it becomes

$$a_n = -v\frac{d\gamma}{dt} + \frac{v^2}{R_E + h}\cos\gamma \tag{8.2.3}$$

where R_E is the radius of the earth and h is the altitude of the rocket. Thus, in the direction of $\hat{\mathrm{u}}_t$ Newton's second law requires

$$T - D - mg\sin\gamma = ma_t \tag{8.2.4}$$

whereas in the $\hat{\mathbf{u}}_n$ direction

$$mg\cos\gamma = ma_n \tag{8.2.5}$$

After substituting Equations 8.2.2 and 8.2.3, these latter two expressions may be written

$$\frac{dv}{dt} = \frac{T}{m} - \frac{D}{m} - g\sin\gamma \tag{8.2.6}$$

$$v\frac{d\gamma}{dt} = -\left(g - \frac{v^2}{R_E + h}\right)\cos\gamma \tag{8.2.7}$$

To these we must add the equations for downrange distance x and altitude h,

$$\frac{dx}{dt} = \frac{R_E}{R_E + h}v\cos\gamma \qquad \frac{dh}{dt} = v\sin\gamma \tag{8.2.8}$$

Numerical methods must be used to solve Equations 8.2.6, 8.2.7 and 8.2.8. To do so, one must account for the variation of the thrust, booster mass, atmospheric density, the drag coefficient, and the acceleration of gravity. Of course, the vehicle mass continuously decreases as propellants are consumed to produce the thrust, which we shall discuss in the following section.

The free-body diagram in Figure 8.2-1 does not include a lifting force, which, if the vehicle were an airplane, would act normal to the velocity vector. Launch vehicles are designed to be strong in lengthwise compression, like a column. To save weight they are, unlike an airplane, made relatively weak in bending, shear and torsion, which are the kinds of loads induced by lifting surfaces. Transverse lifting loads are held closely to zero during powered ascent through the atmosphere by maintaining zero angle of attack, i.e., by keeping the axis of the booster aligned with its velocity vector (the relative wind). Pitching maneuvers are done early in the launch, soon after the rocket clears the launch tower, when its speed is still low. At the high speeds acquired within a minute or so after launch, the slightest angle of attack

can produce destructive transverse loads in the vehicle. The space shuttle orbiter has wings so it can act as a glider after re-entry into the atmosphere. However, the launch configuration of the orbiter is such that its wings are at the zero-lift angle of attack throughout the ascent.

Satellite launch vehicles take off vertically and, at injection into orbit, must be flying parallel to the earth's surface. During the initial phase of the ascent, the rocket builds up speed on a nearly vertical trajectory taking it above the dense lower layers of the atmosphere. While it transitions the thinner upper atmosphere, the trajectory bends over, trading vertical speed for horizontal speed so the rocket can achieve orbital perigee velocity at burnout. The gradual transition from vertical to horizontal flight, illustrated in Figure 8.2-1, is caused by the force of gravity, and it is called a gravity turn trajectory.

At lift off, the rocket is vertical and the flight path angle γ is 90°. After clearing the tower and gaining speed, vernier thrusters or gimbaling of the main engines produce a small, programmed pitchover, establishing an initial flight path angle γ_0, slightly less than 90°. Thereafter, γ will continue to decrease at a rate dictated by Equation 8.2.7. (For example, if $\gamma = 85°$, $v = 110$ m/s (250 mph), and $h = 2$ km, then $d\gamma/dt = -0.44°/s$.) As the speed v of the vehicle increases, the coefficient of cos γ in Equation 8.2.7 decreases, which means the rate of change of the flight path angle becomes increasingly smaller, tending towards zero as the booster approaches orbital speed, $v_{\text{circular orbit}} = \sqrt{g(R+h)}$. Ideally, the vehicle is flying horizontally ($\gamma = 0$) at that point.

The gravity turn trajectory is just one example of a practical trajectory, tailored for satellite boosters. On the other hand, sounding rockets fly straight up from launch through burnout. Rocket-powered guided missiles must execute high-speed pitch and yaw maneuvers as they careen towards moving targets, and require a rugged structure to withstand the accompanying side loads.

8.2.3 The Thrust Equation

To discuss rocket performance requires an expression for the thrust T in Equation 8.2.6. It can be obtained by a simple one-dimensional momentum analysis. Figure 8.2-2(a) shows a system consisting of a rocket and its propellants. The exterior of the rocket is surrounded by the static pressure p_a of the atmosphere everywhere except at the rocket nozzle exit where the pressure is p_e. p_e acts over the nozzle exit area A_e. The value of p_e depends on the design of the nozzle. For simplicity, we assume no other forces act on the system. At time t the mass of the system is m and the absolute velocity in its axial direction is v. The propellants combine chemically in the rocket's combustion chamber, and during the small time interval Δt, a small mass Δm of combustion products is forced out of the nozzle, to the left. As a result of this expulsion, the velocity of the rocket changes by the small amount Δv, to the right. The absolute velocity of Δm is v_e, assumed to be to the left. According to Newton's second law of motion,

$$(\text{momentum of the system at } t+\Delta t) - (\text{momentum of the system at } t) = \text{net external impulse}$$

or

$$[(m-\Delta m)(v+\Delta v)\hat{\mathbf{i}} + \Delta m(-v_e\hat{\mathbf{i}})] - mv\hat{\mathbf{i}} = (p_e - p_a)A_e\Delta t\hat{\mathbf{i}} \tag{8.2.9}$$

Let $\dot{m}_e$ (a positive quantity) be the rate at which exhaust mass flows across the nozzle exit plane. The mass m of the rocket decreases at the rate dm/dt, and conservation of mass requires the decrease of mass to equal the mass flow rate out of the nozzle. Thus,

$$\frac{dm}{dt} = -\dot{m}_e \tag{8.2.10}$$

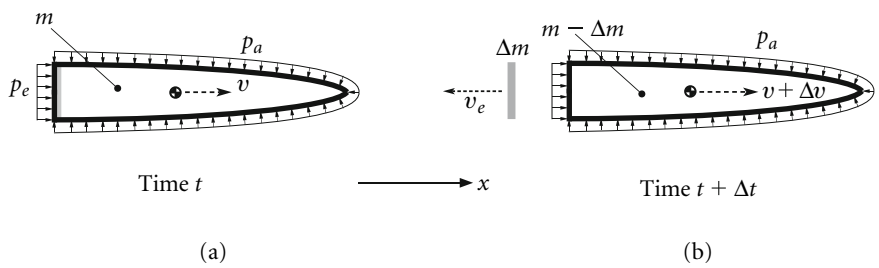


Figure 8.2-2 (a) System of rocket and propellant at time t. (b) The system an instant later, after ejection of a small element Δm of combustion products.

Assuming $\dot{m}_e$ is constant, the vehicle mass as a function of time (from $t = 0$) may therefore be written as

$$m(t) = m_0 - \dot{m}_e t \tag{8.2.11}$$

where m_0 is the initial mass of the vehicle. Since Δm is the mass which flows out in the time interval Δt, we have

$$\Delta m = \dot{m}_e \Delta t \tag{8.2.12}$$

Let us substitute this expression into Equation 8.2.9 to obtain

$$\left[(m - \dot{m}_e \Delta t)(v + \Delta v)\hat{\mathbf{i}} + \dot{m}_e \Delta t\left(-v_e \hat{\mathbf{i}}\right)\right] - mv\hat{\mathbf{i}} = (p_e - p_a)A_e \Delta t \hat{\mathbf{i}}$$

Collecting terms, we get

$$m\Delta v\hat{\mathbf{i}} - \dot{m}_e \Delta t(v_e + v)\hat{\mathbf{i}} - \dot{m}_e \Delta t \Delta v\hat{\mathbf{i}} = (p_e - p_a)A_e \Delta t \hat{\mathbf{i}}$$

Dividing through by Δt, taking the limit as $\Delta t \to 0$, and canceling the common unit vector leads to

$$m\frac{dv}{dt} - \dot{m}_e c_a = (p_e - p_a)A_e \tag{8.2.13}$$

where c_a is the speed of the exhaust relative to the rocket,

$$c_a = v_e + v \tag{8.2.14}$$

Rearranging terms, Equation 8.2.13 maybe written as

$$\dot{m}_e c_a + (p_e - p_a)A_e = m\frac{dv}{dt} \tag{8.2.15}$$

The left-hand side of this equation is the unbalanced force responsible for the acceleration dv/dt of the system in Figure 8.2-2. This unbalanced force is the thrust T,

$$T = \dot{m}_e c_a + (p_e - p_a)A_e \tag{8.2.16}$$

where $\dot{m}_e c_a$ is the jet thrust and $(p_e - p_a)A_e$ is the pressure thrust. We can write Equation 8.2.16 as

$$T = \dot{m}_e\left[c_a + \frac{(p_e - p_a)A_e}{\dot{m}_e}\right] \tag{8.2.17}$$

The term in brackets is called the effective exhaust velocity c,

$$c = c_a + \frac{(p_e - p_a)A_e}{\dot{m}_e} \tag{8.2.18}$$

In terms of the effective exhaust velocity, the thrust may be expressed simply as

$$T = \dot{m}_e c \tag{8.2.19}$$

The specific impulse I_{sp} is defined as the thrust per sea-level weight rate (per second) of propellant consumption. That is,

$$I_{sp} = \frac{T}{\dot{m}_e g_0} \tag{8.2.20}$$

where g_0 is the standard sea-level acceleration of gravity. The unit of specific impulse is force divided by (force/second) or seconds. Together, Equations 8.2.19 and 8.2.20 imply that

$$c = I_{sp} g_0 \tag{8.2.21}$$

Obviously, one can infer the jet velocity directly from the specific impulse. Specific impulse is an important performance parameter for a given rocket engine and propellant combination. However, large specific impulse equates to large thrust only if the mass flow rate is large, which is true of chemical rocket engines. The specific impulses of chemical rockets typically lie in the range 200–300 s for solid fuels and 250–450 s for liquid fuels. Ion propulsion systems have very high specific impulse ($>10^4$ s), but their very low mass flow rates produce much smaller thrust than chemical rockets.

8.2.4 Rocket Performance

From Equations 8.2.10 and 8.2.20 we have

$$T = -I_{sp}g_0\frac{dm}{dt} \tag{8.2.22}$$

or

$$\frac{dm}{dt} = -\frac{T}{I_{sp}g_0}$$

If the thrust and specific impulse are constant, then the integral of this expression over the burn time Δt is

$$\Delta m = -\frac{T}{I_{sp}g_0}\Delta t$$

from which we obtain

$$\Delta t = \frac{I_{sp}g_0}{T}(m_0 - m_f) = \frac{I_{sp}g_0}{T}m_0\left(1 - \frac{m_f}{m_0}\right) \tag{8.2.23}$$

where m_0 and m_f are the mass of the vehicle at the beginning and end of the burn, respectively. The mass ratio is defined as the ratio of the initial mass to final mass,

$$n = \frac{m_0}{m_f} \tag{8.2.24}$$

Clearly, the mass ratio is always greater than unity. In terms of the initial mass ratio, Equation 8.2.23 may be written as

$$\Delta t = \frac{n-1}{n}\frac{I_{sp}}{T/m_0g_0} \tag{8.2.25}$$

T/m_0g_0 is the thrust-to-weight ratio. The thrust-to-weight ratio for a launch vehicle at lift-off is typically in the range 1.3 to 2.

Substituting Equation 8.2.22 into Equation 8.2.6, we get

$$\frac{dv}{dt} = -I_{sp}g_0\frac{dm/dt}{m} - \frac{D}{m} - g\sin\gamma$$

Integrating with respect to time, from t_0 to t_f, yields

$$\Delta v = I_{sp}g_0\ln\frac{m_0}{m_f} - \Delta v_D - \Delta v_G \tag{8.2.26}$$

where the drag loss Δv_D and the gravity loss Δv_g are given by the integrals

$$\Delta v_D = \int_{t_0}^{t_f}\frac{D}{m}dt \qquad \Delta v_G = \int_{t_0}^{t_f} g\sin\gamma\, dt \tag{8.2.27}$$

Since the drag D, acceleration of gravity g, and flight path angle γ are unknown functions of time, these integrals cannot be computed. (Equations 8.2.6 through 8.2.8, together with 8.2.3, must be solved numerically to obtain $v(t)$ and $\gamma(t)$; but then Δv would follow from those results.) Equation 8.2.26 can be used for rough estimates where previous data and experience provide a basis for choosing conservative values of Δv_D and Δv_G. Obviously, if drag can be neglected, then $\Delta v_D = 0$. This would be a good approximation for the last stage of a satellite booster, for which it can also be said that $\Delta v_G = 0$, since $\gamma \cong 0^\circ$ when the satellite is injected into orbit.

Sounding rockets are launched vertically and fly straight up to their maximum altitude before falling back to earth, usually by parachute. Their purpose is to measure remote portions of the earth's atmosphere. ('Sound' in this context means to measure or investigate.) If for a sounding rocket $\gamma = 90^\circ$, then $\Delta v_G \approx g_0\,(t_f - t_0)$, since g is within 90 percent of g_0 out to 300 km altitude.

Example 8.2.1

A sounding rocket of initial mass m_0 and mass m_f after all propellant is consumed is launched vertically ($\gamma = 90^\circ$). The propellant mass flow rate $\dot{m}_e$ is constant. Neglecting drag and the variation of gravity with altitude, calculate the maximum height h attained by the rocket. For what flow rate is the greatest altitude reached?

The vehicle mass as a function of time, up to burnout, is

$$m = m_0 - \dot{m}_e t \tag{a}$$

At burnout, $m = m_f$, so the burnout time t_{bo} is

$$t_{bo} = \frac{m_0 - m_f}{\dot{m}_e} \tag{b}$$

The drag loss is assumed to be zero, and the gravity loss is

$$\Delta v_G = \int_0^{t_{bo}} g_0 \sin(90^\circ) dt = g_0 t_{bo}$$

Recalling that $I_{sp} g_0 = c$ and using (a), it follows from Equation 8.2.26 that, up to burnout, the velocity as a function of time is

$$v = c \ln \frac{m_0}{m_0 - \dot{m}_e t} - g_0 t \tag{c}$$

Since $dh/dt = v$, the altitude as a function of time is

$$h = \int_0^t v\, dt = \int_0^t \left(c \ln \frac{m_0}{m_0 - \dot{m}_e t} - g_0 t \right) dt = \frac{c}{\dot{m}_e} \left[(m_0 - \dot{m}_e t) \ln \frac{m_0 - bt}{m_0} + \dot{m}_e t \right] - \frac{1}{2} g_0 t^2 \tag{d}$$

The height at burnout h_{bo} is found by substituting (b) into this expression,

$$h_{bo} = \frac{c}{\dot{m}_e} \left(m_f \ln \frac{m_f}{m_0} + m_0 - m_f \right) - \frac{1}{2} \left(\frac{m_0 - m_f}{\dot{m}_e} \right)^2 g \tag{e}$$

Likewise, the burnout velocity is obtained by substituting (b) into (c),

$$v_{bo} = c \ln \frac{m_0}{m_f} - \frac{g_0}{\dot{m}_e} (m_0 - m_f) \tag{f}$$

After burnout, the rocket coasts upward with the constant downward acceleration of gravity,

$$v = v_{bo} - g_0 (t - t_{bo})$$
$$h = h_{bo} + v_{bo} (t - t_{bo}) - \tfrac{1}{2} g_0 (t - t_{bo})^2$$

Substituting (b), (e) and (f) into these expressions yields, for $t > t_{bo}$,

$$\begin{aligned} v &= c \ln \frac{m_0}{m_f} - g_0 t \\ h &= \frac{c}{\dot{m}_e} \left(m_0 \ln \frac{m_f}{m_0} + m_0 - m_f \right) + ct \ln \frac{m_0}{m_f} - \frac{1}{2} g_0 t^2 \end{aligned} \tag{g}$$

The maximum height $h_{\max}$ is reached when $v = 0$,

$$c \ln \frac{m_0}{m_f} - g_0 t_{\max} = 0 \Rightarrow t_{\max} = \frac{c}{g_0} \ln \frac{m_0}{m_f}$$

Substituting $t_{\max}$ into (g) leads to our result,

$$\underline{h_{\max} = \frac{c m_0}{\dot{m}_e} (1 + \ln n - n) + \frac{1}{2} \frac{c^2}{g_0} \ln^2 n}$$

where n is the mass ratio ($n > 1$). Since $n > (1 + \ln n)$, it follows that $(1 + \ln n - n)$ is negative. Hence, $h_{\max}$ can be increased by increasing the mass flow rate $\dot{m}_e$. In fact, the greatest height is achieved when $\dot{m}_e \to \infty$, i.e., all of the propellant is expended at once, like a mortar shell.

8.2.5 Restricted Staging in Field-Free Space

In field-free space we neglect drag and gravitational attraction. In that case, Equation 8.2.26 becomes

$$\Delta v = I_{sp} g_0 \ln \frac{m_0}{m_f} \tag{8.2.28}$$

This is at best a poor approximation for high-thrust rockets, but it will suffice to shed some light on the rocket staging problem. Observe that we can solve this equation for the mass ratio to obtain

$$\frac{m_0}{m_f} = e^{\frac{\Delta v}{I_{sp} g_0}} \tag{8.2.29}$$

The amount of propellant expended to produce the velocity increment Δv is $m_0 - m_f$. If we let $\Delta m = m_0 - m_f$, then Equation 8.2.29 can be written as

$$\frac{\Delta m}{m_0} = 1 - e^{-\frac{\Delta v}{I_{sp} g_0}} \tag{8.2.30}$$

This relation is used to compute the propellant required to produce a given delta-v.

The gross mass m_0 of a launch vehicle consists of the empty mass m_E, the propellant mass m_p and the payload mass m_{PL},

$$m_0 = m_E + m_p + m_{PL} \tag{8.2.31}$$

The empty mass comprises the mass of the structure, the engines, fuel tanks, control systems, etc. m_E is also called the structural mass, although it embodies much more than just structure. Dividing Equation 8.2.31 through by m_0, we obtain

$$\pi_E + \pi_p + \pi_{PL} = 1 \tag{8.2.32}$$

where $\pi_E = m_E/m_0$, $\pi_p = m_p/m_0$ and $\pi_{PL} = m_{PL}/m_0$ are the structural fraction, propellant fraction and payload fraction, respectively. It is convenient to define the payload ratio

$$\lambda = \frac{m_{PL}}{m_E + m_p} = \frac{m_{PL}}{m_0 - m_{PL}} \tag{8.2.33}$$

and the structural ratio

$$\varepsilon = \frac{m_E}{m_E + m_p} = \frac{m_E}{m_0 - m_{PL}} \tag{8.2.34}$$

The mass ratio n was introduced in Equation 8.2.24. Assuming all of the propellant is consumed, that may now be written as

$$n = \frac{m_E + m_p + m_{PL}}{m_E + m_{PL}} \tag{8.2.35}$$

λ, ε and n are not independent. From Equation 8.2.34 we have

$$m_E = \frac{\varepsilon}{1 - \varepsilon} m_p \tag{8.2.36}$$

whereas Equation 8.2.33 gives

$$m_{PL} = \lambda(m_E + m_p) = \lambda\left(\frac{\varepsilon}{1 - \varepsilon} m_p + m_p\right) = \frac{\lambda}{1 - \varepsilon} m_p \tag{8.2.37}$$

Substituting Equations 8.2.36 and 8.2.37 into Equation 8.2.35 leads to

$$n = \frac{1 + \lambda}{\varepsilon + \lambda} \tag{8.2.38}$$

Thus, given any two of the ratios λ, ε and n, we obtain the third from Equation 8.2.38. Using this relation in Equation 8.2.28 and setting Δv equal to the burnout speed v_{b0}, when the propellants have been used up, yields

$$v_{bo} = I_{sp} g_0 \ln n = I_{sp} g_0 \ln \frac{1 + \lambda}{\varepsilon + \lambda} \tag{8.2.39}$$

This equation is plotted in Figure 8.2-3 for a range of structural ratios. Clearly, for a given empty mass, the greatest possible Δv occurs when the payload is zero. However, what we want to do is maximize the amount of payload while keeping the structural weight to a minimum. Of course, the mass of load-bearing structure, rocket motors, pumps, piping, etc., cannot be made arbitrarily small. Current materials technology places a lower limit on ε of about 0.1. For this value of the structural ratio and $\lambda = 0.05$, Equation 8.2.39 yields

$$v_{bo} = 1.94 I_{sp} g_0 = 0.019 I_{sp} \text{ (km/s)}$$

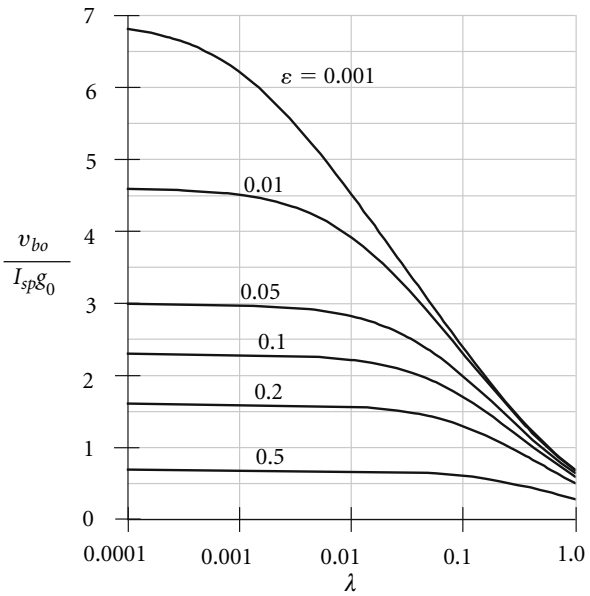


Figure 8.2-3 Dimensionless burnout speed versus payload ratio.

The specific impulse of a typical chemical rocket is about 300 s, which in this case would provide $\Delta v = 5.7$ km/s. However, the circular orbital velocity at the earth's surface is 7.905 km/s. So this booster by itself could not orbit the payload. The minimum specific impulse required for a single stage to orbit would be 416 s. Only today's most advanced liquid hydrogen/liquid oxygen engines, e.g., the space shuttle main engines, have this kind of performance. Practicality and economics would likely dictate going the route of a multi-stage booster.

Figure 8.2-4 shows a series or tandem two-stage rocket configuration, with one stage sitting on top of the other. Each stage has its own engines and propellant tanks. The dividing line between the stages is where they separate during flight. The first stage drops off first, the second stage next, etc. The payload of an N stage rocket is actually stage $N+1$. Indeed, satellites commonly carry their own propulsion systems into orbit. The payload of a given stage is everything

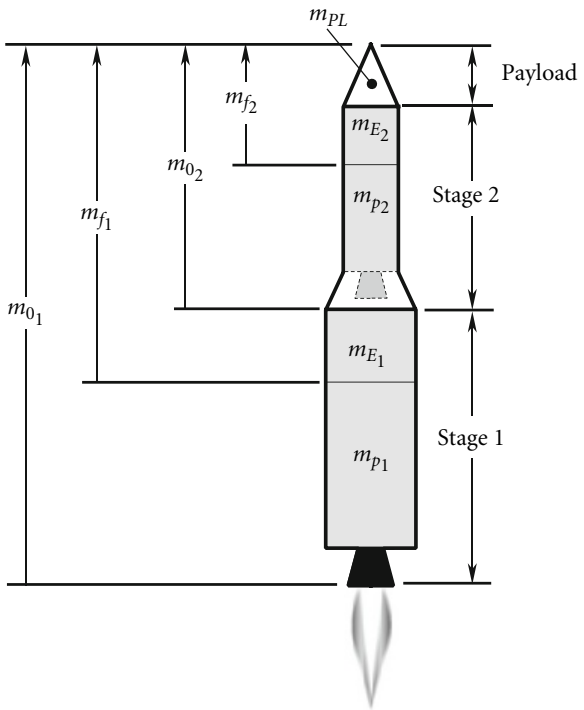


Figure 8.2-4 Tandem two-stage booster.

above it. Therefore, as illustrated in Figure 8.2-4, the initial mass m_{0_1} of stage 1 is that of the entire vehicle. After stage 1 expels all of its fuel, the mass m_{f_1} which remains is stage l's empty mass m_{E_1} plus the mass of stage 2 and the payload. After separation of stage 1, the process continues likewise for stage 2, with m_{0_2} being its initial mass.

Titan II, the launch vehicle for the US Gemini program, had the two-stage, tandem configuration. So did the Saturn 1 B, used to launch earth orbital flights early in the US Apollo program, as well as to send crews to Skylab and an Apollo spacecraft to dock with a Russian Soyuz spacecraft in 1975.

Figure 8.2-5 illustrates the concept of parallel staging. Two or more solid or liquid rockets are attached ('strapped on') to a core vehicle carrying the payload. Whereas in the tandem arrangement, the motors in a given stage cannot ignite until separation of the previous stage, all of the rockets ignite at once in the parallel-staged vehicle. The strap-on boosters fall away after they burn out early in the ascent. The space shuttle is the most obvious example of parallel staging. Its two solid rocket boosters are mounted on the external tank, which fuels the three 'main' engines built into the orbiter. The solid rocket boosters and the external tank are cast off after they are depleted. In more common use is the combination of parallel and tandem staging, in which boosters are strapped to the first stage of a multi-stage stack. Examples include the United States' Titan III and IV and Delta launchers, Europe's Ariane 4 and 5, Russia's Proton and Soyuz variants, Japan's H-2, and China's Long March launch vehicles.

The original Atlas, used in many variants, for among other things, to launch the orbital flights of the US Mercury program, had three main liquid-fuel engines at its base. They all fired simultaneously at launch, but several minutes into the flight, the outer two 'boosters' dropped away, leaving the central sustainer engine to burn the rest of the way to orbit. Since the booster engines shared the sustainer's propellant tanks, the Atlas exhibited partial staging, and is sometimes referred to as a one and a half stage rocket, the discarded boosters comprising the half stage.

We will, for simplicity, focus on tandem staging, although parallel-staged systems are handled in a similar way (Wiesel, 1997). Restricted staging involves the simple but unrealistic assumption that all stages are similar. That is, each stage has the same specific impulse I_{sp}, the same structural ratio ε, and the same payload ratio λ. From Equation 8.2.38 it follows that the mass ratios n are identical, too. Let us investigate the effect of restricted staging on the final burnout speed v_{b0} for a given payload mass m_{PL} and overall payload fraction

$$\pi_{PL} = \frac{m_{PL}}{m_0} \tag{8.2.40}$$

where m_0 is the total mass of the tandem-stacked vehicle.

For a single-stage vehicle, the payload ratio is

$$\lambda = \frac{m_{PL}}{m_0 - m_{PL}} = \frac{1}{\dfrac{m_0}{m_{PL}} - 1} = \frac{\pi_{PL}}{1 - \pi_{PL}} \tag{8.2.41}$$

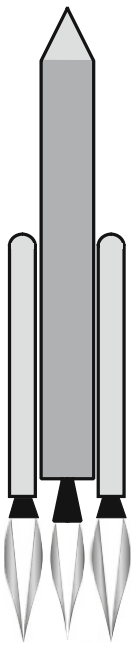


Figure 8.2-5 Parallel staging.

so that, from Equation 8.2.38, the mass ratio is

$$n = \frac{1}{\pi_{PL}(1-\varepsilon)+\varepsilon} \tag{8.2.42}$$

According to Equation 8.2.39, the burnout speed is

$$v_{bo} = I_{sp}g_0 \ln \frac{1}{\pi_{PL}(1-\varepsilon)+\varepsilon} \tag{8.2.43}$$

Let m_0 be the total mass of the two-stage rocket of Figure 8.2-4, i.e.,

$$m_0 = m_{0_1} \tag{8.2.44}$$

The payload of stage 1 is the entire mass m_{0_2} of stage 2. Thus, for stage 1 the payload ratio is

$$\lambda_1 = \frac{m_{0_2}}{m_{0_1} - m_{0_2}} = \frac{m_{0_2}}{m_0 - m_{0_2}} \tag{8.2.45}$$

The payload ratio of stage 2 is

$$\lambda_2 = \frac{m_{PL}}{m_{0_2} - m_{PL}} \tag{8.2.46}$$

By virtue of the two stages' being similar, $\lambda_1 = \lambda_2$, or

$$\frac{m_{0_2}}{m_0 - m_{0_2}} = \frac{m_{PL}}{m_{0_2} - m_{PL}}$$

Solving this equation for m_{0_2} yields

$$m_{0_2} = \sqrt{m_0}\sqrt{m_{PL}}$$

But $m_0 = m_{PL}/\pi_{PL}$, so the gross mass of the second stage is

$$m_{0_2} = \sqrt{\frac{1}{\pi_{PL}}} m_{PL} \tag{8.2.47}$$

Putting this back into Equation 8.2.45 (or 8.2.46), we obtain the common two-stage payload ratio $\lambda = \lambda_1 = \lambda_2$,

$$\lambda_{2\text{-stage}} = \frac{\pi_{PL}^{\frac{1}{2}}}{1 - \pi_{PL}^{\frac{1}{2}}} \tag{8.2.48}$$

This together with Equation 8.2.38 and the assumption that $\varepsilon_1 = \varepsilon_2 = \varepsilon$ leads to the common mass ratio for each stage,

$$n_{2\text{-stage}} = \frac{1}{\pi_{PL}^{\frac{1}{2}}(1-\varepsilon)+\varepsilon} \tag{8.2.49}$$

Assuming that stage 2 ignites immediately after burnout of stage 1, the final velocity of the two-stage vehicle is the sum of the burnout velocities of the individual stages,

$$v_{bo} = v_{bo_1} + v_{bo_2}$$

or

$$v_{bo_{2\text{-stage}}} = I_{sp}g_0 \ln n_{2\text{-stage}} + I_{sp}g_0 \ln n_{2\text{-stage}} = 2I_{sp}g_0 \ln n_{2\text{-stage}}$$

so that, with Equation 8.2.49, we get

$$v_{bo_{2\text{-stage}}} = I_{sp}g_0 \ln \left[\frac{1}{\pi_{PL}^{\frac{1}{2}}(1-\varepsilon)+\varepsilon}\right]^2 \tag{8.2.50}$$

The empty mass of each stage can be found in terms of the payload mass using the common structural ratio ε,

$$\frac{m_{E_1}}{m_{0_1} - m_{0_2}} = \varepsilon \qquad \frac{m_{E_2}}{m_{0_2} - m_{PL}} = \varepsilon$$

Substituting Equations 8.2.40 and 8.2.44 together with 8.2.47 yields

$$m_{E_1} = \frac{(1-\pi_{PL}^{\frac{1}{2}})\varepsilon}{\pi_{PL}} m_{PL} \qquad m_{E_2} = \frac{(1-\pi_{PL}^{\frac{1}{2}})\varepsilon}{\pi_{PL}^{\frac{1}{2}}} m_{PL} \tag{8.2.51}$$

Likewise, we can find the propellant mass for each stage from the expressions

$$m_{p_1} = m_{0_1} - (m_{E_1} + m_{0_2}) \qquad m_{p_2} = m_{0_2} - (m_{E_2} + m_{PL}) \tag{8.2.52}$$

Substituting Equations 8.2.40 and 8.2.44, together with 8.2.47, 8.2.51 and 8.2.52, we get

$$m_{p_1} = \frac{(1-\pi_{PL}^{\frac{1}{2}})(1-\varepsilon)}{\pi_{PL}} m_{PL} \qquad m_{p_2} = \frac{(1-\pi_{PL}^{\frac{1}{2}})(1-\varepsilon)}{\pi_{PL}^{\frac{1}{2}}} m_{PL} \tag{8.2.53}$$

Example 8.2.2

The following data is given

$$\begin{aligned} m_{PL} &= 10\,000 \text{ kg} \\ \pi_{PL} &= 0.05 \\ \varepsilon &= 0.15 \\ I_{sp} &= 350\text{s} \\ g_0 &= 0.00981 \text{ km/s}^2 \end{aligned} \tag{a}$$

Calculate the payload velocity v_{bo} at burnout, the empty mass of the launch vehicle and the propellant mass for (a) a single stage and (b) a restricted, two-stage vehicle.

(a) From Equation 8.2.43 we find

$$v_{bo} = 350 \cdot 0.00981 \ln \frac{1}{0.05(1+0.15)+0.15} = \underline{5.657 \text{ km/s}}$$

Equation 8.2.40 yields the gross mass

$$m_0 = \frac{10\,000}{0.05} = 200\,000 \text{ kg}$$

from which we obtain the empty mass using Equation 8.2.34,

$$m_E = \varepsilon(m_0 - m_{PL}) = 0.15(200\,000 - 10\,000) = \underline{28\,500 \text{ kg}}$$

The mass of propellant is

$$m_p = m_0 - m_E - m_{PL} = 200\,000 - 28\,500 - 10\,000 = \underline{161\,500 \text{ kg}}$$

(b) For a restricted two-stage vehicle, the burnout speed is given by Equation 8.2.50,

$$v_{bo_{2\text{-stage}}} = 350 \cdot 0.00981 \ln \left[\frac{1}{0.05^{\frac{1}{2}}(1-0.15)+0.15}\right]^2 = \underline{7.407 \text{ km/s}}$$

The empty mass of each stage is found using Equations 8.2.51,

$$m_{E_1} = \frac{(1-0.05^{\frac{1}{2}})\cdot 0.15}{0.05} \cdot 10\,000 = \underline{23\,292 \text{ kg}}$$

$$m_{E_2} = \frac{(1-0.05^{\frac{1}{2}})\cdot 0.15}{0.05^{\frac{1}{2}}} \cdot 10\,000 = \underline{5208 \text{ kg}}$$

For the propellant masses, we turn to Equations 8.2.53

$$m_{p_1} = \frac{(1-0.05^{\frac{1}{2}})\cdot(1-0.15)}{0.05} \cdot 10\,000 = \underline{131\,990 \text{ kg}}$$

$$m_{p_2} = \frac{(1-0.05^{\frac{1}{2}})\cdot(1-0.15)}{0.05^{\frac{1}{2}}} \cdot 10\,000 = \underline{29\,513 \text{ kg}}$$

The total empty mass, $m_E = m_{E_1} + m_{E_2}$, and the total propellant mass, $m_p = m_{p_1} + m_{p_2}$, are the same as for the single stage rocket. The mass of the second stage, including the payload, is 22.4 percent of the total vehicle mass.

Observe in the previous example that, although the total vehicle mass was unchanged, the burnout velocity increased 31 percent for the two-stage arrangement. The reason is that the second stage is lighter and can therefore be accelerated to a higher speed. Let us determine the velocity gain associated with adding another stage, as illustrated in Figure 8.2-6.

The payload ratios of the three stages are

$$\lambda_1 = \frac{m_{0_2}}{m_{0_1} - m_{0_2}} \quad \lambda_2 = \frac{m_{0_3}}{m_{0_2} - m_{0_3}} \quad \lambda_3 = \frac{m_{PL}}{m_{0_3} - m_{PL}}$$

Since the stages are similar, these payload ratios are all the same. Setting $\lambda_1 = \lambda_2$ and recalling that m_{01}=m_0, we find

$$m_{0_2}^2 - m_{0_3} m_0 = 0$$

Similarly, $\lambda_1 = \lambda_3$ yields

$$m_{0_2} m_{0_3} - m_0 m_{PL} = 0$$

These two equations imply that

$$m_{0_2} = \frac{m_{PL}}{\pi_{PL}^{\frac{2}{3}}} \qquad m_{0_3} = \frac{m_{PL}}{\pi_{PL}^{\frac{1}{3}}} \tag{8.2.54}$$

Substituting these results back into any one of the above expressions for λ_1, λ_2 or λ_3 yields the common payload ratio for the restricted three-stage rocket,

$$\lambda_{3\text{-stage}} = \frac{\pi_{PL}^{\frac{1}{3}}}{1 - \pi_{PL}^{\frac{1}{3}}}$$

With this result and Equation 8.2.38 we find the common mass ratio,

$$n_{3\text{-stage}} = \frac{1}{\pi_{PL}{}^{\frac{1}{3}}(1 - \varepsilon) + \varepsilon} \tag{8.2.55}$$

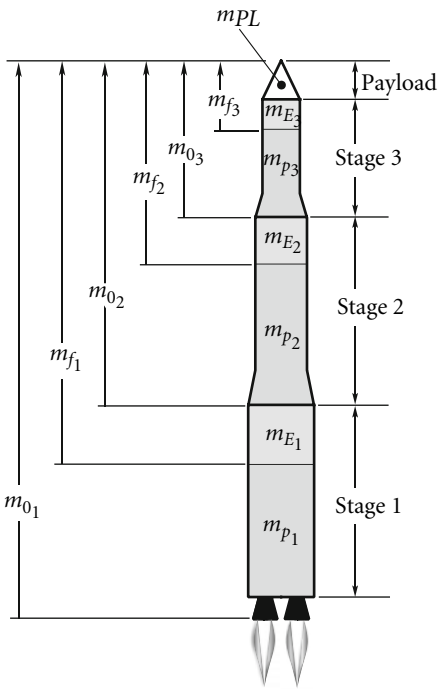


Figure 8.2-6 Tandem three-stage launch vehicle.

Since the payload burnout velocity is $v_{bo} = v_{bo_1} + v_{bo_2} + v_{bo_3}$, we have

$$v_{bo_{3\text{-stage}}} = 3I_{sp}g_o \ln n_{3\text{-stage}} = I_{sp}g_o \ln\left(\frac{1}{\pi_{PL}^{\frac{1}{3}}(1-\varepsilon)+\varepsilon}\right)^3 \tag{8.2.56}$$

Because of the common structural ratio across each stage,

$$\frac{m_{E_1}}{m_{0_1} - m_{0_2}} = \varepsilon \qquad \frac{m_{E_2}}{m_{0_2} - m_{0_3}} = \varepsilon \qquad \frac{m_{E_3}}{m_{0_3} - m_{PL}} = \varepsilon$$

Substituting Equations 8.2.40 and 8.2.54 and solving the resultant expressions for the empty stage masses yields

$$m_{E_1} = \frac{(1-\pi_{PL}^{\frac{1}{3}})\varepsilon}{\pi_{PL}} m_{PL} \quad m_{E_2} = \frac{(1-\pi_{PL}^{\frac{1}{3}})\varepsilon}{\pi_{PL}^{\frac{2}{3}}} m_{PL} \quad m_{E_3} = \frac{(1-\pi_{PL}^{\frac{1}{3}})\varepsilon}{\pi_{PL}^{\frac{1}{3}}} m_{PL} \tag{8.2.57}$$

The stage propellant masses are

$$m_{p_1} = m_{0_1} - (m_{E_1} + m_{0_2}) \quad m_{p_2} = m_{0_2} - (m_{E_2} + m_{0_3}) \quad m_{p_3} = m_{0_3} - (m_{E_3} + m_{PL})$$

Substituting Equations 8.2.40, 8.2.54 and 8.2.57 leads to

$$\begin{aligned} m_{p_1} &= \frac{(1-\pi_{PL}^{\frac{1}{3}})(1-\varepsilon)}{\pi_{PL}} m_{PL} \\ m_{p_2} &= \frac{(1-\pi_{PL}^{\frac{1}{3}})(1-\varepsilon)}{\pi_{PL}^{\frac{2}{3}}} m_{PL} \\ m_{p_3} &= \frac{(1-\pi_{PL}^{\frac{1}{3}})(1-\varepsilon)}{\pi_{PL}^{\frac{1}{3}}} m_{PL} \end{aligned} \tag{8.2.58}$$

Example 8.2.3

Repeat Example 8.2.2 for the restricted three-stage launch vehicle. Equation 8.2.56 gives the burnout velocity for three stages,

$$v_{bo} = 350 \cdot 0.00981 \cdot \ln\left(\frac{1}{0.05^{\frac{1}{3}}(1-0.15)+0.15}\right)^3 = \underline{7.928 \text{ km/s}}$$

Substituting $m_{PL} = 10\,000$ kg, $\pi_{PL} = 0.05$ and $\varepsilon = 0.15$ into Equations 8.2.57 and 8.2.58 yields

$$\underline{m_{E_1} = 18\,948 \text{ kg}} \quad \underline{m_{E_2} = 6980 \text{ kg}} \quad \underline{m_{E_3} = 2572 \text{ kg}}$$

$$\underline{m_{p_1} = 107\,370 \text{ kg}} \quad \underline{m_{p_2} = 39\,556 \text{ kg}} \quad \underline{m_{p_3} = 14\,573 \text{ kg}}$$

Again, the total empty mass and total propellant mass are the same as for the single and two-stage vehicles. Notice that the velocity increase over the two-stage rocket is just 7 percent, which is much less than the advantage the two-stage had over the single-stage vehicle.

Looking back over the velocity formulas for one-, two-, and three-stage vehicles (Equations 8.2.43, 8.2.50 and 8.2.56), we can induce that for an N-stage rocket,

$$\begin{aligned} v_{bo_{N\text{-stage}}} &= I_{sp}g_o \ln\left(\frac{1}{\pi_{PL}^{\frac{1}{N}}(1-\varepsilon)+\varepsilon}\right)^N \\ &= I_{sp}g_o N \ln\left(\frac{1}{\pi_{PL}^{\frac{1}{N}}(1-\varepsilon)+\varepsilon}\right) \end{aligned} \tag{8.2.59}$$

What happens as we let N become very large? First of all, it can be shown using Taylor series expansion that, for large N,

$$\pi_{PL}^{\frac{1}{N}} \approx 1 + \frac{1}{N}\ln \pi_{PL} \tag{8.2.60}$$

Substituting this into Equation 8.2.59, we find that

$$v_{bo_{N\text{-stage}}} \approx I_{sp}g_o N \ln\left[\frac{1}{1+\frac{1}{N}(1-\varepsilon)\ln \pi_{PL}}\right]$$

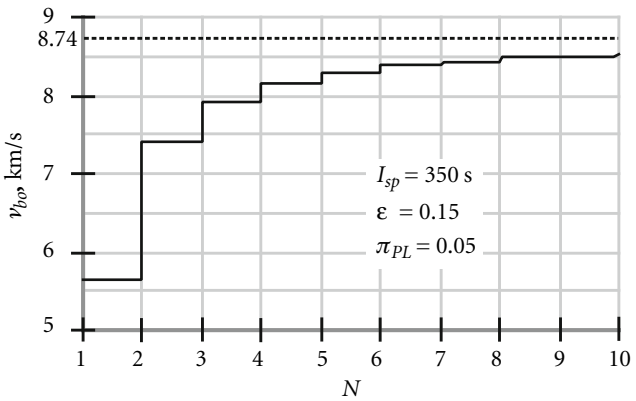


Figure 8.2-7 Burnout velocity versus number of stages (Equation 8.2.59).

Since the term $\frac{1}{N}(1-\varepsilon)\ln \pi_{PL}$ is arbitrarily small, we can use the fact that $1/(1+x) = 1 - x + x^2 - x^3 + \ldots$ to write

$$\frac{1}{1+\frac{1}{N}(1-\varepsilon)\ln \pi_{PL}} \approx 1 - \frac{1}{N}(1-\varepsilon)\ln \pi_{PL}$$

which means

$$v_{bo_{N\text{-stage}}} \approx I_{sp} g_o N \ln\left[1 - \frac{1}{N}(1-\varepsilon)\ln \pi_{PL}\right]$$

Finally, since $\ln(1-x) = -x - x^2/2 - x^3/3 - x^4/4 - \cdots$, we can write this as

$$v_{bo_{N\text{-stage}}} \approx I_{sp} g_0 N\left[-\frac{1}{N}(1-\varepsilon)\ln \pi_{PL}\right]$$

Therefore, as N, the number of stages, tends towards infinity, the burnout velocity approaches

$$v_{bo_\infty} = I_{sp} g_0 (1-\varepsilon)\ln\frac{1}{\pi_{PL}} \tag{8.2.61}$$

Thus, no matter how many similar stages we use, for a given specific impulse, payload fraction and structural ratio, we cannot exceed this burnout speed. For example, using $I_{sp} = 350$ s, $\pi_{PL} = 0.05$ and $\varepsilon = 0.15$ from the previous two examples yields $v_{b0\infty} = 8.743$ km/s, which is only 10 percent greater than v_{bo} of a three-stage vehicle. The trend of v_{bo} towards this limiting value is illustrated by Figure 8.2-7.

Our simplified analysis does not take into account the added weight and complexity accompanying additional stages. Practical reality has limited the number of stages of actual launch vehicles to rarely more than 3.

8.2.6 Optimal Staging

Let us now abandon the restrictive assumption that all stages of a tandem-stacked vehicle are similar. Instead, we will specify the specific impulse I_{sp_i} and structural ratio ε_i of each stage, and then seek the minimum-mass N-stage vehicle that will carry a given payload m_{PL} to a specified burnout velocity v_{bo}. To optimize the mass requires using the Lagrange multiplier method, which we shall briefly review.

8.2.6.1 Lagrange Multiplier

Consider a bivariate function f on the xy plane. Then $z = f(x, y)$ is a surface lying above or below the plane, or both. $f(x, y)$ is stationary at a given point if it takes on a local maximum or a local minimum, i.e., an extremum, at that point. For f to be stationary means $df = 0$; i.e.,

$$\frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy = 0 \tag{8.2.62}$$

where dx and dy are independent and not necessarily zero. It follows that for an extremum to exist,

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0 \tag{8.2.63}$$

Now let $g(x, y) = 0$ be a curve in the xy plane. Let us find the points on the curve $g = 0$ at which f is stationary. That is, rather than searching the entire xy plane for extreme values of f we confine our attention to the curve $g = 0$, which is therefore a constraint. Since $g = 0$, it follows that $dg = 0$, or

$$\frac{\partial g}{\partial x}dx + \frac{\partial g}{\partial y}dy = 0 \tag{8.2.64}$$

If Equations 8.2.62 and 8.2.64 are both valid at a given point, then

$$\frac{dy}{dx} = -\frac{\partial f/\partial x}{\partial f/\partial y} = \frac{\partial g/\partial x}{\partial g/\partial y}$$

That is,

$$\frac{\partial f/\partial x}{\partial g/\partial x} = \frac{\partial f/\partial y}{\partial g/\partial y} = -\eta$$

From this we obtain

$$\frac{\partial f}{\partial x} + \eta\frac{\partial g}{\partial x} = 0 \qquad \frac{\partial f}{\partial y} + \eta\frac{\partial g}{\partial y} = 0$$

But these, together with the constraint $g(x, y) = 0$, are the very conditions required for the function

$$h(x, y, \eta) = f(x, y) + \eta g(x, y) \tag{8.2.65}$$

to have an extremum, namely,

$$\begin{aligned}
\frac{\partial h}{\partial x} &= \frac{\partial f}{\partial x} + \eta\frac{\partial g}{\partial x} = 0 \\
\frac{\partial h}{\partial y} &= \frac{\partial f}{\partial y} + \eta\frac{\partial g}{\partial y} = 0 \\
\frac{\partial h}{\partial \eta} &= g = 0
\end{aligned} \tag{8.2.66}$$

η is the Lagrange multiplier. The procedure generalizes to functions of any number of variables.

One can determine mathematically whether the extremum is a maximum or a minimum by checking the sign of the second differential d^2h of the function h in Equation 8.2.65,

$$d^2h = \frac{\partial^2 h}{\partial x^2}dx^2 + 2\frac{\partial^2 h}{\partial x \partial y}dxdy + \frac{\partial^2 h}{\partial y^2}dy^2 \tag{8.2.67}$$

If $d^2h < 0$ at the extremum for all dx and dy satisfying the constraint condition, Equation 8.2.64, then the extremum is a local maximum. Likewise, if $d^2h > 0$, then the extremum is a local minimum.

Example 8.2.4

(a) Find the extrema of the function $z = -x^2 - y^2$. (b) Find the extrema of the same function under the constraint $y = 2x + 3$.

(a) To find the extrema we must use Equations 8.2.63. Since $\partial z/\partial x = -2x$ and $\partial z/\partial y = -2y$, it follows that $\partial z/\partial x = \partial z/\partial y = 0$ at $x = y = 0$, at which point $z = 0$. Since z is negative everywhere else (see Figure 8.2-8), it is clear that the extreme value is the maximum value.

(b) The constraint may be written as $g = y - 2x - 3$. Clearly, $g = 0$. Multiply the constraint by the Lagrange multiplier η and add the result (zero!) to the function $-(x^2 + y^2)$ to obtain

$$h = -(x^2 + y^2) + \eta(y - 2x - 3)$$

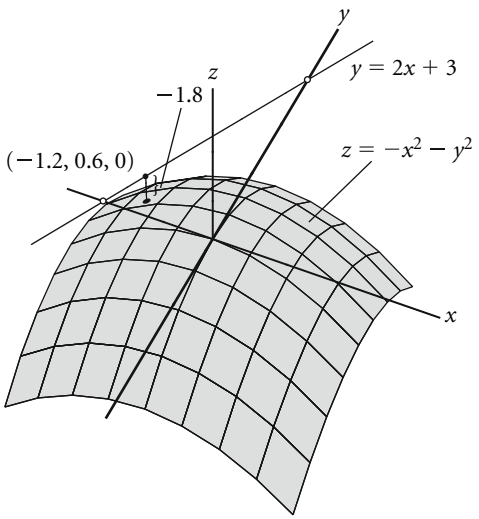


Figure 8.2-8 Location of the point on the line $y = 2x + 3$ at which the surface $z = -x^2 - y^2$ is closest to the xy plane.

This is a function of the three variables x, y and η. For it to be stationary, the partial derivatives with respect to all three of these variables must vanish. First we have

$$\frac{\partial h}{\partial x} = -2x - 2\eta$$

Setting this equal to zero yields

$$x = -\eta \tag{a}$$

Next,

$$\frac{\partial h}{\partial y} = -2y + \eta$$

For this to be zero means

$$y = \frac{\eta}{2} \tag{b}$$

Finally

$$\frac{\partial h}{\partial \eta} = y - 2x - 3$$

Setting this equal to zero gives us back the constraint condition,

$$y - 2x - 3 = 0 \tag{c}$$

Substituting (a) and (b) into (c) yields $\eta = 1.2$, from which (a) and (b) imply,

$$\underline{x = -1.2} \qquad \underline{y = 0.6} \tag{d}$$

These are the coordinates of the point on the line $y = 2x + 3$ at which $z = -x^2 - y^2$ is stationary. Using (d), we find that $z = -1.8$ at this point.

Figure 8.2-8 is an illustration of this problem, and it shows that the computed extremum (a maximum, in the sense that small negative numbers exceed large negative numbers) is where the surface $z = -x^2 - y^2$ is closest to the line $y = 2x + 3$, as measured in the z direction. Note that in this case, Equation 8.2.67 yields $d^2h = -2dx^2 - 2dy^2$, which is negative, confirming our conclusion that the extremum is a maximum.

Now let us return to the optimal staging problem. It is convenient to introduce the step mass m_i of the ith stage. The step mass is the empty mass plus the propellant mass of the stage, exclusive of all the other stages,

$$m_i = m_{E_i} + m_{p_i} \tag{8.2.68}$$

The empty mass of stage i can be expressed in terms of its step mass and its structural ratio ε_i as follows,

$$m_{E_i} = \varepsilon_i(m_{E_i} + m_{p_i}) = \varepsilon_i m_i \tag{8.2.69}$$

The total mass of the rocket excluding the payload is M, which is the sum of all of the step masses,

$$M = \sum_{i=1}^{N} m_i \tag{8.2.70}$$

Thus, recalling that m_0 is the total mass of the vehicle, we have

$$m_0 = M + m_{PL} \tag{8.2.71}$$

Our goal is to minimize m_0.

For simplicity, we will deal first with a two-stage rocket, and then generalize our results to N stages. For a two-stage vehicle, $m_0 = m_1 + m_2 + m_{PL}$, so we can write

$$\frac{m_0}{m_{PL}} = \frac{m_1 + m_2 + m_{PL}}{m_2 + m_{PL}} \frac{m_2 + m_{PL}}{m_{PL}} \tag{8.2.72}$$

The mass ratio of stage 1 is

$$n_1 = \frac{m_{0_1}}{m_{E_1} + m_2 + m_{PL}} = \frac{m_1 + m_2 + m_{PL}}{\varepsilon_1 m_1 + m_2 + m_{PL}} \tag{8.2.73}$$

where Equation 8.2.69 was used. Likewise, the mass ratio of stage 2 is

$$n_2 = \frac{m_{0_2}}{\varepsilon_2 m_2 + m_{PL}} = \frac{m_2 + m_{PL}}{\varepsilon_2 m_2 + m_{PL}} \tag{8.2.74}$$

We can solve Equations 8.2.73 and 8.2.74 to obtain the step masses from the mass ratios,

$$\begin{aligned} m_2 &= \frac{n_2 - 1}{1 - n_2\varepsilon_2} m_{PL} \\ m_1 &= \frac{n_1 - 1}{1 - n_1\varepsilon_1}(m_2 + m_{PL}) \end{aligned} \tag{8.2.75}$$

Now,

$$\frac{m_1 + m_2 + m_{PL}}{m_2 + m_{PL}} = \frac{1 - \varepsilon_1}{1 - \varepsilon_1} \frac{m_1 + m_2 + m_{PL}}{m_2 + m_{PL} + (\varepsilon_1 m_1 - \varepsilon_1 m_1)} \frac{\dfrac{1}{\varepsilon_1 m_1 + m_2 + m_{PL}}}{\dfrac{1}{\varepsilon_1 m_1 + m_2 + m_{PL}}}$$

These manipulations leave the right-hand side unchanged. Carrying out the multiplications proceed as follows:

$$\begin{aligned} \frac{m_1 + m_2 + m_{PL}}{m_2 + m_{PL}} &= \frac{(1 - \varepsilon_1)(m_1 + m_2 + m_{PL})}{\varepsilon_1 m_1 + m_2 + m_{PL} - \varepsilon_1(m_1 + m_2 + m_{PL})} \frac{\dfrac{1}{\varepsilon_1 m_1 + m_2 + m_{PL}}}{\dfrac{1}{\varepsilon_1 m_1 + m_2 + m_{PL}}} \\ &= \frac{(1 - \varepsilon_1)\dfrac{m_1 + m_2 + m_{PL}}{\varepsilon_1 m_1 + m_2 + m_{PL}}}{\dfrac{\varepsilon_1 m_1 + m_2 + m_{PL}}{\varepsilon_1 m_1 + m_2 + m_{PL}} - \varepsilon_1 \dfrac{m_1 + m_2 + m_{PL}}{\varepsilon_1 m_1 + m_2 + m_{PL}}} \end{aligned}$$

Finally, with the aid of Equation 8.2.73, this algebraic trickery reduces to

$$\frac{m_1 + m_2 + m_{PL}}{m_2 + m_{PL}} \frac{(1 - \varepsilon_1)n_1}{1 - \varepsilon_1 n_1} \tag{8.2.76}$$

Likewise,

$$\frac{m_2 + m_{PL}}{m_{PL}} = \frac{(1 - \varepsilon_2)n_2}{1 - \varepsilon_2 n_2} \tag{8.2.77}$$

so that Equation 8.2.72 may be written in terms of the stage mass ratios instead of the step masses,

$$\frac{m_0}{m_{PL}} = \frac{(1 - \varepsilon_1)n_1}{1 - \varepsilon_1 n_1} \frac{(1 - \varepsilon_2)n_2}{1 - \varepsilon_2 n_2} \tag{8.2.78}$$

Taking the natural logarithm of both sides, we get

$$\ln\frac{m_0}{m_{PL}} = \ln\frac{(1-\varepsilon_1)n_1}{1-\varepsilon_1 n_1} + \ln\frac{(1-\varepsilon_2)n_2}{1-\varepsilon_2 n_2}$$

Expanding the logarithms on the right side leads to

$$\begin{aligned}\ln\frac{m_0}{m_{PL}} &= [\ln(1-\varepsilon_1) + \ln n_1 - \ln(1-\varepsilon_1 n_1)] \\ &\quad + [\ln(1-\varepsilon_2) + \ln n_2 - \ln(1-\varepsilon_2 n_2)]\end{aligned} \tag{8.2.79}$$

Observe that for m_{PL}, fixed, $\ln(m_0/m_{PL})$ is a monotonically increasing function of m_0,

$$\frac{d}{dm_0}\left(\ln\frac{m_0}{m_{PL}}\right) = \frac{1}{m_0} > 0$$

Therefore, $\ln(m_0/m_{PL})$ is stationary when m_0 is stationary.

From Equations 8.2.21 and 8.2.39, the burnout velocity of the two-stage rocket is

$$v_{bo} = v_{bo_1} + v_{bo_2} = c_1 \ln n_1 + c_2 \ln n_2 \tag{8.2.80}$$

which means that, given v_{bo}, our constraint equation is

$$v_{bo} - c_1 \ln n_1 - c_2 \ln n_2 = 0 \tag{8.2.81}$$

Introducing the Lagrange multiplier η, we combine Equations 8.2.79 and 8.2.81 to obtain

$$\begin{aligned}h &= [\ln(1-\varepsilon_1) + \ln n_1 - \ln(1-\varepsilon_1 n_1)] + [\ln(1-\varepsilon_2) + \ln n_2 - \ln(1-\varepsilon_2 n_2)] \\ &\quad + \eta(v_{bo} - c_1 \ln n_1 - c_2 \ln n_2)\end{aligned} \tag{8.2.82}$$

Finding the values of n_1 and n_2 for which h is stationary will extremize $\ln(m_0/m_{PL})$ (and, hence, m_0) for the prescribed burnout velocity v_{bo}. h is stationary when $\partial h/\partial n_1 = \partial h/\partial n_2 = \partial h/\partial n = 0$. Thus,

$$\frac{\partial h}{\partial n_1} = \frac{1}{n_1} + \frac{\varepsilon_1}{1-\varepsilon_1 n_1} - \eta\frac{c_1}{n_1} = 0$$

$$\frac{\partial h}{\partial n_2} = \frac{1}{n_2} + \frac{\varepsilon_2}{1-\varepsilon_2 n_2} - \eta\frac{c_2}{n_2} = 0$$

$$\frac{\partial h}{\partial n} = v_{bo} - c_1 \ln n_1 - c_2 \ln n_2 = 0$$

These three equations yield, respectively,

$$n_1 = \frac{c_1\eta - 1}{c_1\varepsilon_1\eta} \qquad n_2 = \frac{c_2\eta - 1}{c_2\varepsilon_2\eta} \qquad v_{bo} = c_1 \ln n_1 + c_2 \ln n_2 \tag{8.2.83}$$

Substituting n_1 and n_2 into the expression for v_{bo}, we get

$$c_1 \ln\left(\frac{c_1\eta - 1}{c_1\varepsilon_1\eta}\right) + c_2 \ln\left(\frac{c_2\eta - 1}{c_2\varepsilon_2\eta}\right) = v_{bo} \tag{8.2.84}$$

This equation must be solved iteratively for η, after which η is substituted into Equations $8.2.83_{1,2}$ to obtain the stage mass ratios n_1 and n_2. These mass ratios are used in Equations 8.2.75 together with the assumed structural ratios, exhaust velocities, and payload mass to obtain the step masses of each stage.

We can now generalize the optimization procedure to an N-stage vehicle, for which Equation 8.2.82 becomes

$$h = \sum_{i=1}^{N}[\ln(1-\varepsilon_i) + \ln n_i - \ln(1-\varepsilon_i n_i)] - \eta\left(v_{bo} - \sum_{i=1}^{N} c_i \ln n_i\right) \tag{8.2.85}$$

At the outset, we know the required burnout velocity v_{bo}, the payload mass m_{PL}, and for every stage we have the structural ratio ε_i and the exhaust velocity c_i (i.e., the specific impulse). The first step is to solve for the Lagrange parameter η using Equation 8.2.84, which, for N stages, is written as

$$\sum_{i=1}^{N} c_i \ln\frac{c_i\eta - 1}{c_i\varepsilon_i\eta} = v_{bo}$$

Expanding the logarithm, this can be written

$$\sum_{i=1}^{N} c_i \ln(c_i\eta - 1) - \ln \eta \sum_{i=1}^{N} c_i - \sum_{i=1}^{N} c_i \ln c_i\varepsilon_i = v_{bo} \tag{8.2.86}$$

After solving this equation iteratively for η, we use that result to calculate the optimum mass ratio for each stage (cf. Equation 8.2.83),

$$n_i = \frac{c_i\eta - 1}{c_i\varepsilon_i\eta}, \qquad i = 1,\ 2,\ \ldots,\ N \tag{8.2.87}$$

Of course, each n_i must be greater than 1.

Referring to Equations 8.2.75, we next obtain the step masses of each stage, beginning with stage N and working our way down the stack to stage 1,

$$\begin{aligned}
m_N &= \frac{n_N - 1}{1 - n_N\varepsilon_N} m_{PL} \\
m_{N-1} &= \frac{n_{N-1} - 1}{1 - n_{N-1}\varepsilon_{N-1}} (m_N + m_{PL}) \\
m_{N-2} &= \frac{n_{N-2} - 1}{1 - n_{N-2}\varepsilon_{N-2}} (m_{N-1} + m_N + m_{PL}) \\
&\vdots \\
m_1 &= \frac{n_1 - 1}{1 - n_1\varepsilon_1} (m_2 + m_3 + \cdots m_{PL})
\end{aligned} \tag{8.2.88}$$

Having found each step mass, each empty stage mass is

$$m_{E_i} = \varepsilon_i m_i \tag{8.2.89}$$

and each stage propellant mass is

$$m_{p_i} = m_i - m_{E_i} \tag{8.2.90}$$

For the function h in Equation 8.2.85 it is easily shown that

$$\frac{\partial^2 h}{\partial n_i \partial n_j} = 0, \quad i, j = 1, \ldots, N\ (i \neq j)$$

It follows that the second differential of h is

$$d^2h = \sum_{i=1}^{N} \sum_{j=1}^{N} \frac{\partial^2 h}{\partial n_i \partial n_j} dn_i dn_j = \sum_{i=1}^{N} \frac{\partial^2 h}{\partial n_i^2} (dn_i)^2 \tag{8.2.91}$$

where it can be shown, again using Equation 8.2.85, that

$$\frac{\partial^2 h}{\partial n_i^2} = \frac{\eta c_i (\varepsilon_i n_i - 1)^2 + 2\varepsilon_i n_i - 1}{(\varepsilon_i n_i - 1)^2 n_i^2} \tag{8.2.92}$$

For h to be minimum at the mass ratios n_i given by Equation 8.2.87, it must be true that $d^2h > 0$. Equations 8.2.91 and 8.2.92 indicate that this will be the case if

$$\eta c_i (\varepsilon_i n_i - 1)^2 + 2\varepsilon_i n_i - 1 > 0, \qquad i = 1,\ \ldots,\ N \tag{8.2.93}$$

Example 8.2.5

Find the optimal mass for a three-stage launch vehicle which is required to lift a 5000 kg payload to a speed of 10 km/s. For each stage, we are given that

Stage 1	I_{sp_1}	= 400 s ($c_1 = 3.924$ km/s)	$\varepsilon_1 = 0.10$
Stage 2	I_{sp_2}	= 350 s ($c_2 = 3.434$ km/s)	$\varepsilon_2 = 0.15$
Stage 3	I_{sp_3}	= 300 s ($c_3 = 2.943$ km/s)	$\varepsilon_3 = 0.20$

Substituting this data into Equation 8.2.86, we get

$$3.924 \ln (3.924\eta - 1) + 3.434 \ln (3.434\eta - 1) + 2.943 \ln (2.943\eta - 1) - 10.30 \ln \eta + 7.5089 = 10$$

As can be checked by substitution, the iterative solution of this equation is

$$\eta = 0.4668$$

Substituting η into Equations 8.2.87 yields the optimum mass ratios,

$$n_1 = 4.541 \qquad n_2 = 2.507 \qquad n_3 = 1.361$$

For the step masses, we appeal to Equations 8.2.88 to obtain

$$m_1 = 165\,700 \text{ kg} \qquad m_2 = 18\,070 \text{ kg} \qquad m_3 = 2477 \text{ kg}$$

Using Equations 8.2.89 and 8.2.90, the empty masses and propellant masses are found to be

$$m_{E_1} = 16\,570 \text{ kg} \qquad m_{E_2} = 2710 \text{ kg} \qquad m_{E_3} = 495.4 \text{ kg}$$
$$m_{p_1} = 149\,100 \text{ kg} \qquad m_{p_2} = 15\,360 \text{ kg} \qquad m_{p_3} = 1982 \text{ kg}$$

The payload ratios for each stage are

$$\lambda_1 = \frac{m_2 + m_3 + m_{PL}}{m_1} = 0.1542$$
$$\lambda_2 = \frac{m_3 + m_{PL}}{m_2} = 0.4139$$
$$\lambda_3 = \frac{m_{PL}}{m_3} = 2.018$$

The total mass of the vehicle is

$$m_0 = m_1 + m_2 + m_3 + m_{PL} = 191\,200 \text{ kg}$$

and the overall payload fraction is

$$\pi_{PL} = \frac{m_{PL}}{m_0} = \frac{5000}{191\,200} = 0.0262$$

Finally, let us check Equation 8.2.93,

$$\eta c_1 (\varepsilon_1 n_1 - 1)^2 + 2\varepsilon_1 n_1 - 1 = 0.4541$$
$$\eta c_2 (\varepsilon_2 n_2 - 1)^2 + 2\varepsilon_2 n_2 - 1 = 0.3761$$
$$\eta c_3 (\varepsilon_3 n_3 - 1)^2 + 2\varepsilon_3 n_3 - 1 = 0.2721$$

A positive number in every instance means we have indeed found a local minimum of the function in Equation 8.2.85.

Problems

8.2.1 Suppose a spacecraft in permanent orbit around the earth is to be used for delivering payloads from low earth orbit (LEO) to geostationary equatorial orbit (GEO). Before each flight from LEO, the spacecraft is refueled with propellant which it uses up in its round trip to GEO. The outbound leg requires four times as much propellant as the inbound return leg. The delta-v for transfer from LEO to GEO is 4.22 km/s. The specific impulse of the propulsion system is 430 s. If the payload mass is 3500 kg, calculate the empty mass of the vehicle.
{Ans.: 2733 kg}

8.2.2 A two stage, solid-propellant sounding rocket has the following properties:

First stage :	$m_0 = 249.5$ kg	$m_f = 170.1$ kg	$\dot{m}_e = 10.61\ kg/s$	$I_{sp} = 235$ s
Second stage :	$m_0 = 113.4$ kg	$m_f = 58.97$ kg	$\dot{m}_e = 4.053\ kg/s$	$I_{sp} = 235$ s

Delay time between burnout of first stage and ignition of second stage: 3 seconds.
As a preliminary estimate, neglect drag and the variation of earth's gravity with altitude to calculate the maximum height reached by the second stage after burnout.
{Ans.: 322 km}

8.2.3 A two-stage launch vehicle has the following properties:
First stage: two solid propellant rockets. Each one has a total mass of 525 000 kg, 450 000 kg of which is propellant. $I_{sp} = 290$ s.
Second stage: two liquid rockets with $I_{sp} = 450$ s. Dry mass = 30 000 kg, propellant mass = 600 000 kg.
Calculate the payload mass to a 300 km orbit if launched due east from KSC. Let the total gravity and drag loss be 2 km/s.
{Ans.: 114 000 kg}

8.2.4 Consider a rocket comprising three similar stages (i.e., each stage has the same specific impulse, structural ratio and payload ratio). The common specific impulse is 310 s. The total mass of the vehicle is 150 000 kg, the total structural mass (empty mass) is 20 000 kg and the payload mass is 10 000 kg. Calculate
(a) The mass ratio n and the total Δv for the three-stage rocket.
{Ans.: $n = 2.04$, $\Delta\ v = 6.50$ km/s}
(b) m_{p_1}, m_{p_2}, and m_{p_3}.
(c) m_{E_1}, m_{E_2}, and m_{E_3}.
(d) m_{0_1}, m_{0_2} and m_{0_3}.

8.2.5 A small two-stage vehicle is to propel a 10 kg payload to a speed of 6.2 km/s. The properties of the stages are: for the first stage, $I_{sp} = 300$ s and $\varepsilon = 0.2$; for the second stage, $I_{sp} = 235$ s and $\varepsilon = 0.3$. Estimate the optimum mass of the vehicle.
{Ans.: 1125 kg}

8.2.6 Find the extrema of the function $z = x^2 + y^2 + 2xy$ subject to the constraint $x^2 - 2x + y^2 = 0$.
{Ans.: $z_{\min} = 0.1716$ at $(x, y) = (0.2929, -0.7071)$ and $z_{\max} = 5.828$ at $(x, y) = (1.707, 0.7071)$}

SECTION 9

Airworthiness

Aerospace Engineering Desk Reference; ISBN 9781856175753

9.1 The ICAO and the Civil Aviation Authorities

Filippo De Florio

9.1.1 The ICAO (International Civil Aviation Organization)

The first recorded flight by a heavier-than-air machine was by the Wright brothers on 17 December 1903 in North Carolina.

Since the earliest years of aviation, far-seeing people envisaged a new dimension of transport that would go beyond national boundaries. In 1910, the first conference on air navigation international law was hosted by France in Paris, with the attendance of 18 European states.

The First World War fostered considerable development of aeronautical techniques, also demonstrating the potential for transport of goods and people. After the war, it became increasingly evident that this advanced means of transport would require international attention.

These problems were debated at the Paris Conference of Peace in 1919, and the discussions led to the establishment of an Aeronautical Commission. To succeed in the purpose of making aviation an instrument of peace, an International Air Convention was written and ratified by 38 states. The Convention contemplated all aspects of civil aviation and also the establishment of an International Commission for Air Navigation (ICAN) in order to monitor the development of civil aviation and to propose measures for this development.

The years between the two world wars marked a continuous development of civil aviation both in the technical and the commercial fields.

The Second World War, apart from the horrors also caused by the operations of progressively more sophisticated military aeroplanes, had a major effect upon the technical development of the aeroplane, compressing a quarter of a century of normal peacetime development into six years.

The possibility of carrying a great number of people and a large quantity of goods over long distances became a reality. For these reasons, the Government of the United States conducted exploratory discussions with other allied nations from the early months of 1944. On the basis of these talks, invitations were sent to 55 allied and neutral states to meet in Chicago in November 1944. Of these 55 states, 52 attended. The outcome of 5 weeks of meetings was the Convention on International Civil Aviation, consisting of a preamble and 96 articles.

The ICAO officially came into existence on 4 April 1947. At the invitation of the Government of Canada, Montreal was chosen as the site for its headquarters. Presently, the Contracting States number more than 180.

The aims and objectives of the ICAO are to develop the principles and techniques of international air navigation and to foster the planning and development of international air transport so as to:

1. Ensure the safe and orderly growth of international civil aviation throughout the world.
2. Encourage the arts of aircraft design and operation for peaceful purposes.
3. Encourage the development of airways, airports and air navigation facilities for international civil aviation.
4. Meet the needs of the peoples of the world for safe, regular, efficient and economical air transport.
5. Prevent economic waste caused by unreasonable competition.
6. Ensure that the rights of the Contracting States are fully respected and that every Contracting State has a fair opportunity to operate international airlines.
7. Avoid discrimination between Contracting States.
8. Promote safety of flight in international air navigation.
9. Promote generally the development of all aspects of international civil aeronautics.

9.1.1.1 The International Standards

Since the ICAO was created, a main technical task of the organization has been the achievement of standardization in the operation of a safe, regular and efficient air service. This has resulted in high levels of reliability in the many areas that collectively shape international civil aviation, particularly in relation to the aircraft, their crews, and the ground-based facilities and services.

Standardization has been achieved through the creation, adoption, and amendments of 18 Annexes to the Convention, identified as **International Standards** and **Recommended Practices.**

Standards are directives which ICAO members agree to follow. If a member has a standard different from an ICAO Standard, that member must notify the ICAO of the difference.

Recommended practices are desirable but not essential practices. The basic principle for deciding whether a particular issue should be a Standard is an affirmative answer to the question: 'Is uniform application by all Contracting States essential?'

On the basis of the Convention, the Contracting States are engaged in achieving the highest practical degree of worldwide uniformity in regulations, organizing procedures in relation to aircraft, personnel, airways, and auxiliary services, whenever this will facilitate and improve air safety, effectiveness, and regularity.

The 18 Annexes are described as follows:

- **Annex 1. Personnel Licensing** – provides information on licensing of flight crews, air traffic controllers, and aircraft maintenance personnel, including medical standards for flight crews and air traffic controllers.
- **Annex 2. Rules of the Air** – contains rules relating to visual and instrument-aided flight.
- **Annex 3. Meteorological Service for International Air Navigation** – provides meteorological services for international air navigation and reporting of meteorological observations from aircraft.
- **Annex 4. Aeronautical Charts** – contains specifications for the aeronautical charts used in international aviation.
- **Annex 5. Units of Measurement to be used in Air and Ground Operations** – lists dimensional systems to be used in air and ground operations.
- **Annex 6. Operation of Aircraft** – enumerates specifications to ensure a level of safety above a prescribed minimum in similar operations throughout the world. The three parts of this Annex are as follows:
  - Part I. International Commercial Air Transport – Airplanes
  - Part II. International General Aviation – Airplanes
  - Part III. International Operations – Helicopters.
- **Annex 7. Aircraft Nationality and Registration Marks** – specifies requirements for registration and identification of aircraft.
- **Annex 8. Airworthiness of Aircraft** – specifies uniform procedures for certification and inspection of aircraft.
- **Annex 9. Facilitations** – provides for the standardization and simplification of border crossing formalities.
- **Annex 10. Aeronautical Telecommunications** – Volume 1 provides for standardizing communications equipment and systems, Volume 2 standardizes communications procedures.
- **Annex 11. Air Traffic Services** – includes information on establishing and operating ATC, flight information, and alerting services.
- **Annex 12. Search and Rescue** – provides information on organization and operation of facilities and services necessary for search and rescue (SAR).
- **Annex 13. Aircraft Accident Investigation** – provides for uniformity in notifying, investigating, and reporting on aircraft accidents.
- **Annex 14. Aerodromes** – contains specifications for the design and equipment of aerodromes.
- **Annex 15. Aeronautical Information Services** – includes methods for collecting and disseminating aeronautical information required for flight operations.
- **Annex 16. Environmental Protection** – Volume 1 contains specifications for aircraft noise certification, noise monitoring, and noise exposure units for land-use planning, Volume 2 contains specifications for aircraft engine emissions.
- **Annex 17. Security – Safeguarding International Civil Aviation against Acts of Unlawful Interference** – specifies methods for safeguarding international civil aviation against unlawful acts of interference.
- **Annex 18. The Safe Transport of Dangerous Goods by Air** – specifies requirements necessary to ensure hazardous materials are safely transported in aircraft while providing a level of safety that protects the aircraft and its occupants from undue risk.

Because aeronautical technology is continuously developing, the Annexes are constantly reviewed and updated when necessary. The **typical content** of an Annex is based upon:

1. Standards intended as specifications when their application is considered as *necessary* for the safety and regularity of international air navigation.
2. Recommended practices intended as specifications when their application is considered as a *recommendation* in the interest of safety, regularity, and efficiency of international air navigation.
3. Appendices dealing with the preceding points.
4. Definitions of the used terminology.

The Contracting States have issued norms not strictly copying the contents of the Annex, which essentially state some of the principles or objectives to attain. The norms contain the requirements used to reach the objectives. Furthermore, while the principles can remain the same, the requirements are often influenced by the state of the art (technical evolution, new technology, and acquired experience), and they are then likely to be improved and amended.

The applicable JAA/FAA/EASA airworthiness standards for the certification of aircraft to be internationally recognized are issued in accordance with the ICAO Annexes. Then, from a practical point of view, the certification process is based on these airworthiness standards rather than (directly) on the ICAO International Standards.

In order to remain within the scope and objectives of this book, we will consider the content of the three Annexes that are directly connected with airworthiness:

- **Annex 6. Operation of Aircraft.** This Annex contains the standards defining recommendations relating to the operation of aircraft for international commercial air transport, including the regulation for the certification of the operators. It also contains the technical and operational regulations for general international aviation activities, including maintenance.
- **Annex 8. Airworthiness of Aircraft.** This Annex contains the standards defining the minimum level of airworthiness for the development of the type certification requirements *as a basis for the international recognition of the certificates of airworthiness for aircraft* (according to Article 33 of the Convention) in order to fly into and land in the Contracting States. It also contains indications and provisions for the organization and functions of the civil aviation authorities.
- **Annex 16. Environmental Protection.** This Annex contains the standard applicable to the *aircraft noise* certification in relation to different noise levels proportionate to the type of aircraft (propeller-driven, jet-propelled, helicopters). It states with accuracy the test procedures for an effective and unequivocal measurement. The standard contained in this Annex is normally used as proposed because it is directly applicable to all of the technical requirements. The Annex also contains the standard relating to the *aircraft engine emission* certification with reference to the toxicity of some chemical components, such as azoth oxide. This is a particularly discussed Annex at present, because it deals with a sensitive social matter, noise, for people living near aerodromes, and the conflict between the sometimes opposing demands of economic development and the protection of citizens.

9.1.2 The Civil Aviation Authorities

9.1.2.1 Origins

The national states of developed countries have established institutions and authorities to guarantee flight safety. In many cases, these organizations evolved from pre-existing institutions for the safety of marine and river navigation. It is of interest to point out that, historically, the mainspring for the improvement of the safety of navigation is not a social principle, but an economical choice made by insurance companies.

The word 'register' was adopted by various navigational institutions and has a precise origin. In fact, it is derived from a register that a certain Edward Lloyd, owner of a tavern situated in the area of the river port of London at the end of the seventeenth century, filled with information on marine traffic gathered while talking to customers, such as ship owners and sailors. The collected information could be related to ships, traffic and, most importantly, to accidents resulting in the loss of men, goods and ships. This was the origin of the highly esteemed newsletters, 'Lloyd's News', that were first issued in 1696.

At the same time, marine insurance began to flourish and Lloyd's tavern rapidly became an important negotiation center. Lloyd was a practical man, well aware of the importance of the information he owned for the insurance business. Finally, Lloyd's, the incorporated society of underwriters in London, was born and was destined to become a world reference in the insurance field.

When Lloyd died in 1713, his heirs continued his work; 'Lloyd's List', filled with lists, data, and marine news, highly appreciated in the circle of marine traffic, was first published in 1734; the List, originally handwritten, first appeared in printed form in 1760.

Meanwhile, other lists with various ship classification criteria were published by different ship owners, until all the publications were unified into the 'Lloyd's Register' in 1833, the first register in the world, which acquired legal status in 1871. Other national registers were subsequently instituted in Europe.

Safety is obviously a matter of great importance for insurance companies: fewer accidents mean fewer indemnities to pay. It is also for this reason that the registers began to issue safety requirements for navigation.

Since the beginning of aviation, the operation of aircraft posed problems of an analogous nature to that of marine traffic, hence the necessity of the establishment of specific institutions, similar to the already existing institutions for marine traffic. In some cases, particular marine institutions took on the responsibilities of aviation regulations and control. Later, the growth of aviation led to the creation of autonomous registers and national authorities, dealing with aircraft and air navigation.

9.1.2.2 Tasks of Airworthiness Authorities[1]

From a general point of view, an airworthiness authority has the following tasks:

1. **To prescribe** airworthiness requirements and procedures.
2. **To inform** the interested parties regarding the above-mentioned prescriptions. This is performed in different ways. The authority publishes technical regulations, technical standards, circulars, etc., to be obtained on request or by other means. At present, much information can be found on the Internet.
3. **To control** aeronautical material, design, and manufacturing organizations, and aircraft operators. This is to ensure that all pertinent prescriptions are complied with. Control can be performed in different ways, with the appropriate involvement of the relevant authority.
4. **To certificate** aeronautical material and organizations. This is to declare in a legal form compliance with the applicable requirements of an aircraft or part of it, or a change to a type certificate, the capability of an organization, and so on.

9.1.3 The Joint Aviation Authorities (JAA)

The Joint Aviation Authorities (JAA) is an associated body of the European Civil Aviation Conference (ECAC)[2] representing the civil aviation regulatory authorities of a number of European states who have agreed to cooperate in developing and implementing common safety regulatory standards and procedures.

The JAA's work started in 1970 (when it was known as the Joint Airworthiness Authorities). The aim was to develop a harmonized airworthiness standard to meet the needs of the industry in Europe, particularly for products manufactured by international consortia. Since 1987, JAA activities have been extended to operations, maintenance, and licensing and certification design standards for all classes of aircraft.

JAA membership is based on the approval of the JAA Arrangements that the State Members signed in 1990 in Cyprus. Membership is open to the ECAC's members, which at present consist of 41 countries.

The JAA has a two-phase membership system. The current procedures start with a familiarization visit by a 'candidate' authority to the Central JAA (CJAA). After a satisfactory conclusion, the authority can formally apply for membership, expressing its willingness to commit itself to the terms and commitments in the Arrangements. At this stage, if the application is accepted, the authority becomes a 'candidate member' and will have access to meetings, documentation, etc., without voting rights.

Full membership is achieved after the JAA's satisfactory evaluations of the candidate member according to established procedures.

At present, the JAA consists of 33 full members and six candidate members.

Based on the Arrangements and related commitments, the JAA's objectives and functions may be summarized as follows.

9.1.3.1 Objectives

1. *Aviation safety*. To ensure, through cooperation among Member States, that JAA members achieve a high, consistent level of aviation safety.
2. *Transition from the JAA to the EASA*. To ensure the highest level of contribution to the European Union for establishing an Aviation Safety Agency that absorbs all functions and activities of the JAA in as short a time as possible and ensures the full participation of JAA non-EU Members.
3. *Business effectiveness*. To achieve a cost-effective safety system so as to contribute to efficient civil aviation.
4. *Consolidation of common standards*. To contribute, through the uniform application of common standards and through regular revision of existing regulations, to fair and equal competition within Member States.
5. *Worldwide aviation safety improvement*. To cooperate with other regional organizations or national authorities of states who are playing an important role in civil aviation, in order to reach at least the JAA safety level and to foster the worldwide implementation of harmonized safety standards and requirements through the conclusion of international arrangements.

9.1.3.2 Functions

1. To develop and adopt Joint Aviation Requirements (JARs) in the fields of aircraft design and manufacture, aircraft operations and maintenance, and the licensing of aviation personnel.
2. To develop administrative and technical procedures for the implementation of the JARs.
3. To implement JARs and the related administrative and technical procedures in a coordinated manner.
4. To adopt measures to ensure, whenever possible, that pursuance of the JAA safety objective does not unreasonably distort competition between the aviation industries of Member States or place companies of Member States at a competitive disadvantage with companies of non-Member States.
5. To provide the principal center of professional expertise in Europe on the harmonization of aviation safety regulations.
6. To cooperate on the harmonization of requirements and the procedures with other safety regulatory authorities, particularly the Federal Aviation Administration (FAA).

9.1.3.3 Organization of the JAA

The JAA system is run by the governing body in which a JAA Board (JAAB) and a JAA Committee (JAAC) work closely together.

The **JAAB** consists of the Director Generals of Civil Aviation and an appropriate European Community representation.

The functions of the JAAB are to determine the general policy and long-term strategy or the JAA, and to coverall the fieldsofcompetence of the JAA.

The **JAAC** is comprised of the representatives of the National Aviation Authorities and of the European Commission. It takes technical decisions concerning products, services, organizations, and people, to cover all the fields of competence of the JAA. It controls the Executive and reports to the JAA Board as specified in the Arrangements.

The Committee also has an **Executive Board** of seven members selected by the JAA Committee, one being the seat of the EASA, who meet more regularly to decide day-to-day matters, to prepare work for the full Committee, and to make provisional decisions.

The JAAC staff is headed by a Chief Executive.

The JAA deals with almost all aviation safety problems, except for air traffic control. In Europe, air traffic control is coordinated by EUROCONTROL,[3] an organization that maintains close relations with the JAA.

9.1.3.4 Transition from the JAA to the EASA

With the adoption of Regulation (EC) No. 1592/2002, the national regulation for EU Member States has been replaced by EU Regulation. Also, all certification tasks have been transferred from the national authorities (NA) to the EASA. Non-EU states retain their accountability in all fields.

While the EASA is building its own organization, the JAA continues to exist, ensuring the highest level of contribution to the EASA, which is subsequently to absorb all its functions and activities. Since 2002, the JAA has actively participated in the transition to the EASA by developing, in consultation with the Commission, a transition plan focused on regulatory aspects.

On 28 November 2003, the EASA signed the Cyprus arrangement, becoming a full member of the JAA. This arrangement allowed the EASA to simplify the relationship with the JAA, with the potential to act in the name of all EU Member States.

Within this new framework the JAA maintains all its functions and responsibilities in operation and licensing, while acting as service provider to the EASA in certification and maintenance.

9.1.3.5 The Future of the JAA

A set of options for the future of JAA/EASA relations was discussed during the ECAC's 51st Special Meeting of Director Generals of Civil Aviation (DGCA) in Yalta, from 30 August to 2 September 2002. It was decided that the so-called Option 3, by which the regulatory and certification activity would be conducted within the EASA system but with transparency to non-EASA Member States for the relevant decisions, offered the best prospects for the

transition period. It was also agreed that Option 4 should be the ultimate goal, under which JAA activities would be fully integrated with the EASA.

In light of the above decisions, the JAA Board took the initiative to develop a 'roadmap' for the establishment of clear milestones for its medium-term activities. Accordingly, a working group on the Future of the JAA (FUJA) was established to develop this roadmap under clearly defined objectives.

The main objective of the FUJA working group was to produce a detailed document regarding the future of the JAA (the so-called roadmap), inclusive of a precise indication as to when, where, and how each activity performed by the JAA would be transferred or disbanded (milestones), and to consider the most suitable means to continue the association with non-EU ECAC states in ongoing safety-related activities in Europe.

The final report was presented to the ECAC Director Generals at their Special Meeting on 26 August 2005 in Romania, at which time the report, together with the main decisions, were adopted.

9.1.3.6 General Remarks

The activity of this worthy organization, which is to be replaced by the EASA, has very often been limited by its own nature. It is worth mentioning that we are talking about **authorities,** not **authority.** This means that the JAA did not have the legal status of an authority and therefore a legally recognized power. They did not have the power, for example, to issue certificates.

Instead, they could only 'recommend' to the national authorities the release of such certificates under the relevant terms and conditions. For the same reasons, they could not impose rules and procedures – unless they became European directives – but only 'recommend' their implementation. The shortcomings of such a situation are clear, considering the variety of rules and laws in force in the Member States. This is why the institution of a true European authority was increasingly felt to be a necessity.

This is now a reality with the institution of the EASA, which has benefited from the substantial and complex work carried out by the JAA.

9.1.4 The European Aviation Safety Agency (EASA)

The EASA is an independent European Community body with a legal identity and autonomy in legal, administrative, and financial matters.

This single authority has been created by the adoption of a European Parliament and Council Regulation (EC) No. 1592/2002 of 15 July 2002 in order to put in place a Community system of air safety and environmental regulation.

The activity of the EASA started, as planned, on 28 September 2003 and, after a transitory period in Brussels, the Agency has now moved to Cologne (Germany).

9.1.4.1 Main Tasks

1. To assist the European Commission in preparing legislation, and support the Member States and industry in putting the legislation into effect.
2. To assist the European Commission in monitoring the application of European Community legislation.
3. To adopt its own certification specification and guidance material, conduct technical inspections, and issue certificates where centralized action is more efficient.

The Agency will develop its know-how in all the fields of aviation safety and environmental protection in order to assist Community legislators in the issuing of common rules for:

1. The certification of aeronautical products, parts, and appliances.
2. The approval of organizations and personnel engaged in the maintenance of these products.
3. The approval of air operations.
4. The licensing of aircrew.
5. The safety oversight of airports and air traffic services operators.

The EASA, as a body of the European Community, is subject to the provisions of the Financial Regulations applicable to the general budget of the European Communities, the EASA Financial Regulations, and the relevant directives relating to the coordination of procedures for the award of public contracts.

9.1.4.2 EASA Partnerships

The EASA works closely with representatives of other organizations to ensure that it takes their views into account:

1. Interested parties in industry, which are subject to rules drafted by the EASA, are pivotal in ensuring the success of civil aviation safety standards by assisting in the drafting and correct application of European Community and EASA rules.
2. European aviation authorities perform a critical role in assisting the EASA with the performance of its core rulemaking, certification, and standardization functions.
3. International aviation organizations such as the Joint Aviation Authorities, EUROCONTROL, and the International Civil Aviation Organization work together with the EASA to promote international civil aviation standards.
4. International aviation authorities such as the Federal Aviation Administration, Transport Canada, DAC/CTA (Brazil), and the Interstate Aviation Committee (Russia) work with the EASA to ensure compliance with international standards and to facilitate trade in aeronautical products.
5. Accident investigation bodies issue safety recommendations and analysis that guide the Agency's safety strategy.

9.1.4.3 Structure of the EASA

The EASA Headquarters includes:

1. Executive Directorate
2. Rulemaking Directorate
3. Certification Directorate
4. Quality and Standardization Directorate
5. Administrative Directorate.

The **Executive Director** is appointed by the Agency's **Management Board** (Figure 9.1-1). This Board, which brings together representatives of the Member States' authorities and the Commission, is responsible for the definition of the Agency's priorities, the establishment of the budget, and for monitoring the Agency's operation.[4]

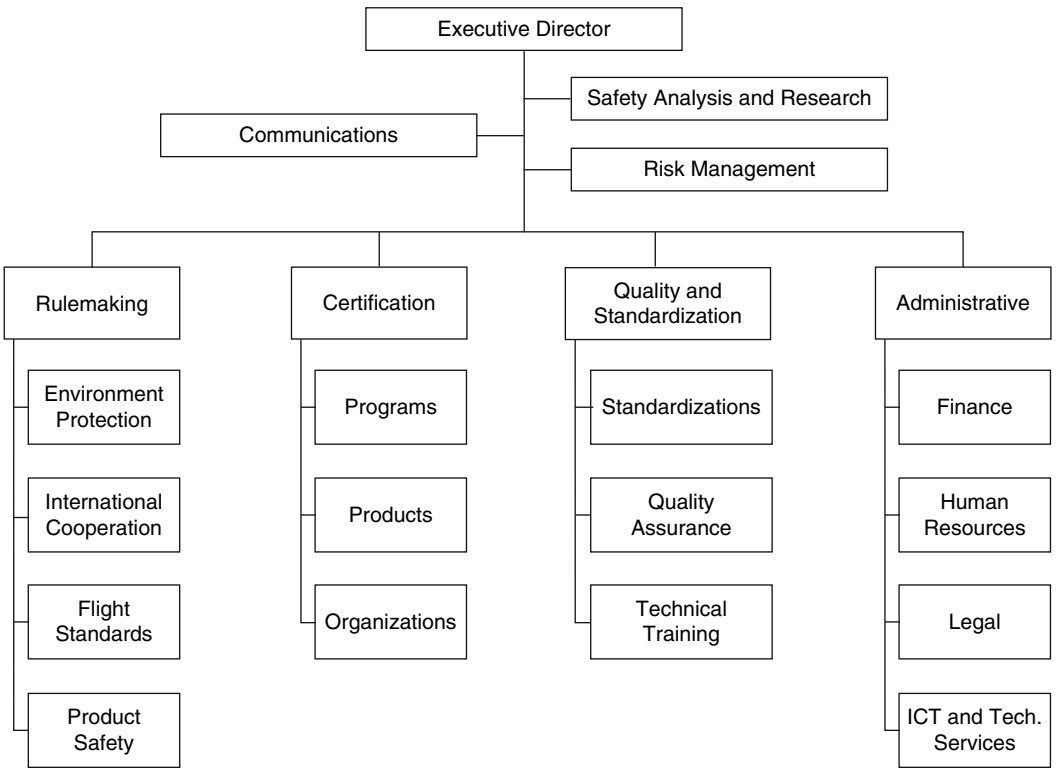


Figure 9.1-1 EASA organizational chart.

The **Rulemaking Directorate** contributes to the production of all EU legislation and implementation of material related to the regulation of civil aviation safety and environmental compatibility. It submits opinions to the European Commission and must be consulted by the Commission on any technical question in its field of competence. It is also in charge of the related international cooperation. Experts within the Rulemaking Directorate have direct contact with all relevant stakeholders, and make use of the knowledge available within the industry and national administrations across the European Union. The Agency's team of experts is comprised of people with a recognized background in aviation and Community regulations.

Currently, the Basic Regulation establishes Community competence only for the regulation of the airworthiness and environmental compatibility of aeronautical products, parts, and appliances. Work is underway to extend the scope of this regulation to embrace the regulation of pilot licensing, air operations, and third country aircraft. It is also envisaged to extend the scope of the Basic Regulation to the safety regulation of airport operations and air traffic control services.

The **Certification Directorate.** On 28 September 2003, the Agency took over responsibility for the airworthiness and environmental certification of all aeronautical products, parts, and appliances designed, manufactured, maintained, or used by persons under the regulatory oversight of EU Member States.

The Agency's certification work also includes all post-certification activities, such as the approval of changes to, and repairs of, aeronautical products and their components, as well as the issuing of Airworthiness Directives to correct any potentially unsafe situation. All type certificates are therefore now issued by the European Aviation Safety Agency and are valid throughout the European Union.

On the same date, the Agency became the competent authority to approve and oversee the organizations involved in the design of aeronautical products, parts, and appliances. It also carries out the same role for foreign organizations involved in the manufacture or maintenance of such products.

To execute its tasks within the present period of building up its resources, the Agency relies on *national aviation authorities* who have historically filled this role and concludes contractual arrangements to this effect.

The **Quality and Standardization Directorate.** Where Community law is implemented at Member State level, the Agency assists the Commission in overseeing its effective application and its uniform understanding.

The necessary standards are therefore being developed and maintained properly, uniformly, and consistently across the European Union.

Accordingly, the Agency conducts inspections of undertakings as well as national authorities throughout the EU, both to monitor the application of EU rules on aviation safety and to assess the effectiveness of these rules. The Agency also provides technical training, which is essential to achieve overall consistency.

The **Administrative Directorate** supports the operational activities of the Agency. Its role is to help the Agency to plan and manage its resources within the limits set out in the regulatory framework. The Directorate′s specialists deal with human resource issues, budgeting and finance, infrastructure, legal affairs, and procurement.

9.1.4.4 EASA Certification

9.1.4.4.1 Design approval

The Certification Directorate is responsible for the management of all applications for design approval and for the issue of related certificates or approvals, and may use either internal or external resources to deal with technical investigations.

According to Regulation (EC) No. 1592, the EASA takes responsibility for the design approval of products, parts, and appliances designed, manufactured, or used by persons under the regulatory oversight of EU Member States, **except for those excluded by its Annex II**.[5]

The European Commission then adopted *Regulation (EC) 1702/2003,* which specifies *inter alia* the requirements applicable to products, parts, and appliances, and also provides for the grandfathering of pre-existing certificates under conditions that aim at ensuring that they meet the level of safety required by the Basic Regulation (EC) No. 1592/2002 and its rules of implementation.

Consequently, only the products that comply with one of the provisions below are deemed to have been issued a type certificate compliant with Regulation (EC) No. 1592/2002, unless the Agency determines differently:

1. Products certificated in accordance with the JAA rules and procedures.
2. Products certificated by a Member State, acting as the State of Design, on the basis of a well-known airworthiness code.
3. Products certificated by a Member State in the framework of a bilateral agreement with the State of Design, on the basis of that State's airworthiness codes.

Products that do not benefit from these grandfathering provisions will remain under the national administrations' oversight until the time the Agency has determined their type certificates. This will be done by 28 March 2007.

Though the EASA is directly responsible for type certifications, at present some certification tasks can be executed by the national authorities to ease the necessary period of transition.

In relation to the products already type certificated, the Agency will expedite, in cooperation with the concerned Member States of design, the review of the type certification bases of these products with the view to determine their EASA type certificate and thus take over responsibility for their continued airworthiness.

Aircraft which were permitted to fly before 20 September 2003 and cannot be issued an EASA type certificate will remain under the responsibility of the Member State of Registry under applicable national regulations.

9.1.4.4.2 Organization approval

The Certification Directorate is responsible for:

1. The management of *all* applications for *Design Organization Approval*, the issue of related certificates and their continued surveillance.
2. The management of all applications from *non-EU countries* (or from EU countries on request of the competent authority), for *Production Organization Approval*, the issue of related certificates and their continued surveillance.
3. The management of all applications from *non-EU countries* for *maintenance and maintenance training organization approval*. All *EU production, maintenance, and maintenance training organizations* are approved by the *local competent authority*, in accordance with the rules of implementation.

9.1.4.4.3 General remarks

At the end of 2004, the EASA was still in the organization phase. According to *Flight International* (October 2004):

> *The EASA is currently engaged in extending its powers beyond its existing responsibility for airworthiness and maintenance into the operations arena. According to approved plans for centralizing all aviation safety rulemaking, the EASA is preparing to assume responsibility for operational issues, including air traffic management, airports and pilots, mirroring the US Federal Aviation Administration.*
>
> *Mr Goudou*[6] *used a speech to the European Parliament to address claims from several national aviation authorities, that supplementary national requirements licensing will continue to be enforced in the future. But, unlike the predecessor, the Joint Aviation Authorities, the EASA will not merely recommend regulations. As an agency of the EU's executive, the European Commission, it will have the power to enforce compliance.*

In spite of Mr Goudou's goodwill, in the same article *Flight International* mentioned that the EASA has had a setback in its recruitment of 95 certification staff by the end of 2004, mainly because of the Agency's move from Brussels to Cologne, which could not be considered attractive for experienced people living in other locations.

In any case, it is worth reading what Goudou wrote in an article (for a UVS International Publication):

> *During the set-up and transition phase, the keyword of the Agency's activities is 'continuity'. Indeed, it goes without saying that the Agency is not going to reinvent the wheel, as its initial tasks are based on the activities and existing procedures of the Joint Aviation Authorities (JAA), and on national know-how, which enables the Agency to provide continuity in terms of the certification work and the progressive resumption, without major upheaval, of the work carried out now by the JAA and national authorities. As such, no project has been delayed since the Agency has become operational.*

The future will tell if the choice of having all the certification staff in Cologne is the right one.

In a further development, in a letter to *Flight International* (December 2004), Gerd Muehlbauer[7] proposed the institution of regional EASA offices:

> *This would allow technical resources from national airworthiness authorities to be used and would suit those who are unwilling to move to the EASA's office in Cologne.*
>
> *The FAA approach works, financed by the ticket fee,*[8] *with its headquarters in Washington and regional offices around the nation, strategically located close to the industry.*

9.1.5 The Federal Aviation Administration (FAA)

9.1.5.1 Origins

The Air Commerce Act of 20 May 1926 was the cornerstone of the Federal government's regulation of civil aviation. This landmark legislation was passed at the behest of the aviation industry, whose leaders believed the

aircraft could not reach its full commercial potential without Federal action to improve and maintain safety standards. The Act charged the Secretary of Commerce with fostering air commerce, issuing and enforcing air traffic rules, licensing pilots, certificating aircraft, establishing airways, and operating and maintaining aids to air navigation. A new Aeronautics Branch of the Department of Commerce assumed primary responsibility for aviation oversight.

9.1.5.2 Early Responsibility

In fulfilling its civil aviation responsibilities, the Department of Commerce initially concentrated on functions such as safety rulemaking and the certification of pilots and aircraft.

In 1934, the Aeronautics Branch was renamed the Bureau of Air Commerce to reflect its enhanced status within the Department. As commercial flying increased, the Bureau encouraged a group of airlines to establish the first three centers for providing air traffic control (ATC) along the airways. In 1936, the Bureau itself took over the centers and began to expand the ATC system.

9.1.5.3 The Civil Aeronautics Act

In 1938, the Civil Aeronautics Act transferred the Federal civil aviation responsibilities from the Commerce Department to a new independent agency, the Civil Aeronautics Authority.

In 1940, President Franklin Roosevelt split the Authority into two agencies, the Civil Aeronautics Administration (CAA) and the Civil Aeronautics Board (CAB). The CAA was responsible for ATC, airman and aircraft certification, safety enforcement, and airway development. The CAB was entrusted with safety rulemaking, accident investigation, and economic regulation of the airlines. Both organizations were part of the Department of Commerce.

9.1.5.4 The Birth of the FAA

The approaching introduction of jet airliners and a series of mid-air collisions spurred passage of the Federal Aviation Act of 1958. This legislation transferred the CAA's functions to a new independent body, the Federal Aviation Agency (FAA), that had broader authority to combat aviation hazards. The act took safety rulemaking from the CAB and entrusted it to the new FAA. It also gave the FAA sole responsibility for developing and maintaining a common civil-military system of air navigation and air traffic control, a responsibility the CAA previously shared with others.

9.1.5.5 From Agency to Administration

In 1966, Congress authorized the creation of a cabinet department that would combine major Federal transportation responsibilities. This new Department of Transportation (DOT) began full operations on 1 April 1967. On that day, the FAA became one of several modal organizations within the DOT and was given a new name, the **Federal Aviation Administration.** At the same time, the CAB's accident investigation function was transferred to the new National Transportation Safety Board (NTSB).

9.1.5.6 Structural Changes

The FAA's organizational structure has continued to evolve since its creation. The agency's first Administrator favored a management system under which officials in Washington exercised direct control over programs in the field. In 1961, however, his successor began a decentralization process that transferred much authority to regional organizations. This pattern generally endured until a 1988 'straightlining' again charged managers at national headquarters with more direction of field activities.

9.1.6 FAA Activities

9.1.6.1 Safety Regulations

The FAA issues and enforces regulations and minimum standards covering manufacturing, operating, and maintaining aircraft. It also certifies airmen and airports that serve air carriers.

9.1.6.2 Airspace and Traffic Management

The safe and efficient use of navigable airspace is one of the FAA's primary objectives. The FAA operates a network of airport towers, air route traffic control centers, and flight service stations. It also develops air traffic rules, assigns the use of airspace, and controls air traffic.

9.1.6.3 Air Navigation Facilities

The FAA builds or installs visual and electronic aids to air navigation. It also maintains, operates, and assures the quality of these facilities, and sustains other systems to support air navigation and air traffic control, including voice and data communications equipment, radar facilities, computer systems, and visual display equipment at flight service stations.

9.1.6.4 Civil Aviation Abroad

The FAA promotes aviation safety, encourages civil aviation abroad, and takes part in international conferences. Aeronautical information is exchanged with foreign authorities. The FAA certifies foreign repair shops, airmen and mechanics, provides technical aid and training, negotiates 'Bilateral Safety Agreements' (BASA) with other authorities with the 'Implementation Procedures for Airworthiness' (IPA) to allow and facilitate the mutual certification of aeronautical products which are imported or exported between the USA and a signatory country, as well as promoting technical cooperation in matters of airworthiness, including maintenance, flight operations, and environmental certification.

The FAA deals with all the problems related to flight safety in the United States, but has representatives on five continents committed to ensuring and promoting the safety, security, and efficiency of international civil aviation. The FAA engages in dialog with its counterparts in 188 countries and works closely with the ICAO. This effort includes providing technical assistance and training, ensuring that countries with airlines flying to the USA meet international standards, and harmonizing global standards so that passengers can benefit from a seamless air transportation network.

It is clear that all these international activities have the final and institutional purpose of guaranteeing flight safety in the USA. However, we cannot ignore the considerable drive given by the FAA to the growth of safety on a global scale.

9.1.6.5 Commercial Space Transportation

The FAA regulates and encourages the US commercial space transportation industry. It licenses commercial space launch facilities and private launches of space payloads on expendable launch vehicles.

9.1.6.6 Research, Engineering, and Development

The FAA conducts research on and develops the systems and procedures needed for a safe and efficient system of air navigation and air traffic control. It helps develop better aircraft, engines and equipment, and tests or evaluates aviation systems, devices, materials, and procedures. The FAA also carries out aero-medical research.

9.1.6.7 Other Programs

The FAA registers aircraft and records documents reflecting title or interest in aircraft and their parts. It administers an aviation insurance program, develops specifications for aeronautical charts, and publishes information on airways, airport services and other technical subjects in aeronautics.

9.1.6.8 Summary of FAA Activities

The FAA is responsible for the safety of civil aviation. Its main roles include:

1. Regulating civil aviation to promote safety.
2. Encouraging and developing civil aeronautics, including new aviation technology.
3. Developing and operating a system of air traffic control and navigation for both civil and military aircraft.
4. Researching and developing the National Airspace System and civil aeronautics.
5. Developing and carrying out programs to control aircraft noise and other environmental effects of civil aviation.
6. Regulating US commercial space transportation.

9.1.7 FAA Certification

The organization of the FAA is very complex; this is understandable considering the plurality of tasks, the size of the USA, and its relationship with the rest of the world.

From an airworthiness point of view, we will try to describe which structure deals with each relevant issue.

In the vast FAA organizational chart (Figure 9.1-2) we can find the **Aviation Safety** headquarters located in Washington which, among its many offices (like the Office of Accident Investigation, Office of Aerospace Medicine, etc.), hosts the **Aircraft Certification Service,** structured as shown in Figure 9.1-3.

Figure 9.1-4 summarizes the main tasks of this Service.

9.1.7.1 The Aircraft Certification Service

The Aircraft Certification Service administers the type certification program to determine compliance with the prescribed regulations and to maintain certificate integrity (continued airworthiness).

The Service is composed of three headquarters divisions, four certifications directorates, an International Airworthiness Program Staff, and an Aircraft Certification Division located in Brussels, Belgium.

The Aircraft Certification Service's responsibility *for administering the Federal Aviation Regulations* is divided as follows:[9]

1. The **Aircraft Engineering Division** is responsible for overall policy and guidance for engineering portions of the Aircraft Certification Regulatory Program (ACRP).[10] Furthermore, it is responsible for FAR 21, 39, and all Special FARs[11] pertaining to type certification.
2. The **Production and Airworthiness Certification Division** is responsible for the overall policy and guidance for manufacturing and airworthiness certification portions of the Aircraft Certification Regulatory Program. It is also responsible for FAR 21, 43, 45, 183, and all Special FARs pertaining to certification conformity, airworthiness certification, and production.
3. The **Planning and Program Management Division** is responsible for the coordination of the Service's strategic and tactical planning initiatives and processes. It is also responsible for service technical, general and

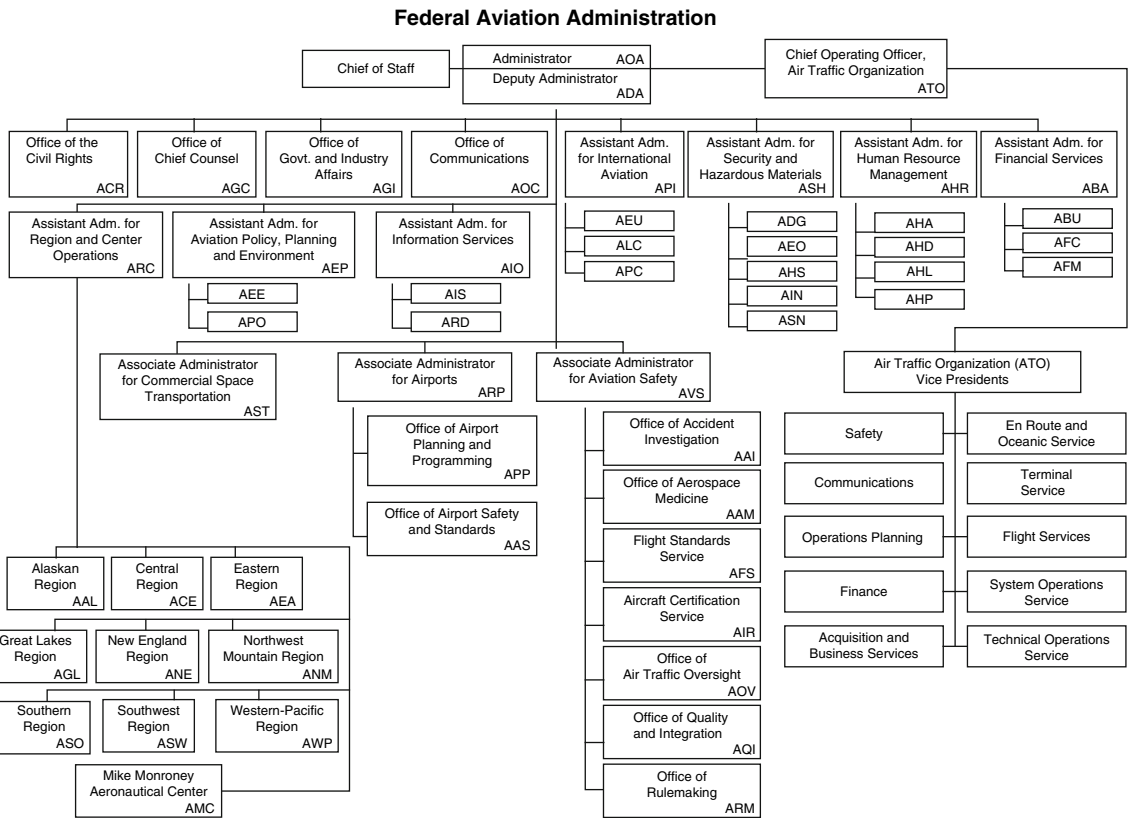


Figure 9.1-2 Organization of the Federal Aviation Administration (FAA).

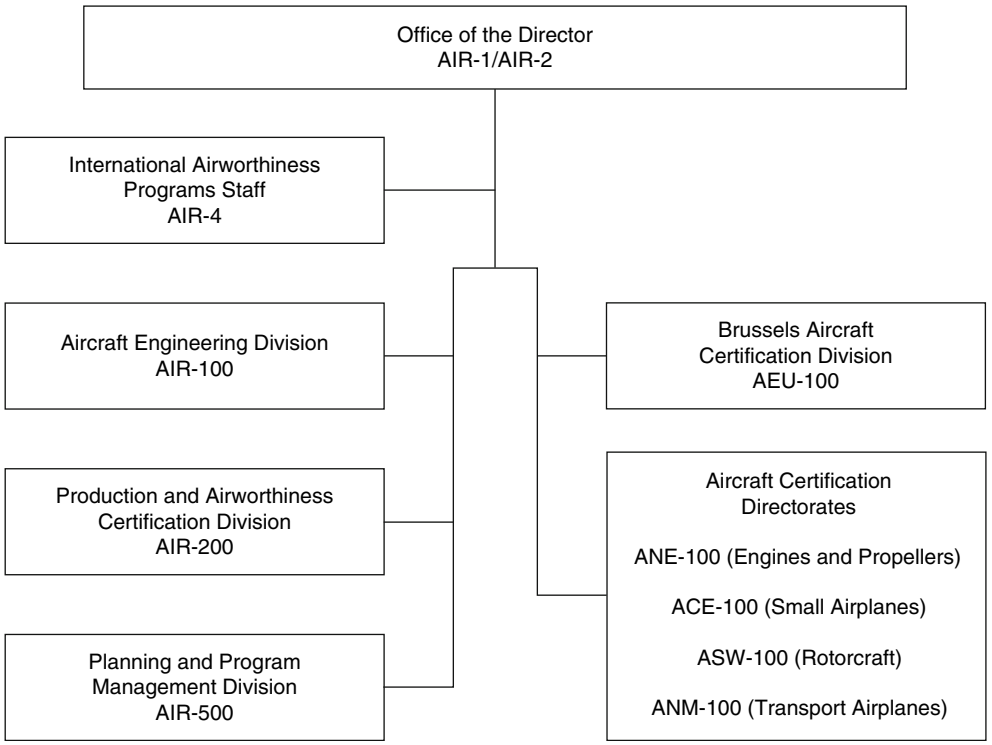


Figure 9.1-3 Structure of the Aircraft Certification Service.

managerial training requirements, administrative and program management guidance, coordination, and support for Service headquarters organizations.

4. The **Small Airplane Directorate** (Central Region) is responsible for FAR 23, 31, technical guidance for FAR 23 Restricted category airplanes; airworthiness standard for gliders and airships, and technical guidance for Primary category airplanes; FAR 23, glider and airship import TC projects from Europe, Africa, and Israel.
5. The **Transport Airplane Directorate** (Northwest Mountain Region) is responsible for FAR 25 and technical guidance for FAR 25, Restricted category airplanes, and FAR 25, import TC projects from Europe, Africa, and Israel.
6. The **Rotorcraft Directorate** (Southwest Region) is responsible for FAR 27 and 29, technical guidance for FAR 27 and 29 Restricted category rotorcraft, powered lift aircraft, and guidance for primary category aircraft; FAR 27 and 29 import TC projects from Europe, Africa, and Israel.
7. The **Engine and Propeller Directorate** (New England Region) is responsible for FAR 33 and 35, technical guidance on auxiliary power units (APUs).

The Service also has other functions:

1. The **International Airworthiness Program Staff (AIR-4)** is responsible for policy guidance on bilateral agreements, import and export of aeronautical products, and other international airworthiness issues, programs, and procedures.
2. **Aircraft Certification Offices (ACOs).** Each directorate incorporates three or more ACOs within their geographical areas issuing the actual certification of aircraft and products. They work directly with the applicant and provide the main interface between the public and the FAA.
3. **Aircraft Evaluation Group (AEG).** A Flight Standards group is co-located with each directorate and it is responsible for determining operational acceptability and continuing airworthiness requirements of newly certified or modified aircraft, engines, and propellers intended to be operated under the provisions of the FARs.

We will now describe in more detail the four Directorates mentioned above.

9.1.7.2 The Small Airplane Directorate

The Small Airplane Directorate (Central Region) consists of the **Directorate headquarters** located in Kansas City; four **Aircraft Certification Offices (ACO)** located in *Anchorage, Atlanta, Chicago* and *Wichita;* and seven **Manufacturing**

Aircraft Certification Service – Products and services		
Design Approvals	**Production Approvals for Aircraft, Engines, and Propellers**	**Airworthiness Certification**
• Transport airplanes • Small airplanes • Engines and propellers (*including APUs*) • Rotorcraft • Airships • Manned free balloons	• Production certificate • Production under a type certificate only • Approved Production Inspection System (APIS)	• Standard airworthiness certificate • Special airworthiness certificate (amateur-built) • Approved Production Inspection system (APIS) • Special flight authorization • Export approvals • Import approvals
Design Modifications for Aircraft, Engines, and Propellers	**Design and Production Approvals for Parts/ Articles/Appliances**	**Representatives of the Administrator (Designees)**
• Amended type certificate • Supplemental type certificate • Field approval	• Parts Manufacturer Approval • Technical Standard Order authorization	• Designee resources • Designee process overview • Designee/FAA selection and appointment process • Designee training
Continued Operational Safety	**International Aviation**	
• Airworthiness Directives (AD) process • Alternate Method of Compliance (AMOC) • Design approval holder reporting requirements	• Bilateral agreements	

Figure 9.1-4 Main tasks of the Aircraft Certification Service.

Inspection District Offices (MIDOs)[12] located in *Atlanta*, *Cleveland*, *Kansas City*, *Minneapolis*, *Orlando*, *Vandalia*, and *Wichita*.

The primary functions of the Directorate headquarters in Kansas City are to:

1. Provide administrative support and resource management for the Directorate field offices.
2. Develop type certification policies and regulations for *small airplanes*, *airships*, and *balloons*, and ensure standardized application of the policies and regulations.
3. Administer type certification of small airplanes, airships, and balloons in field offices outside the Directorate.
4. Monitor *service difficulty information* and process airworthiness actions for small airplanes, airships, and balloons.

The Small Airplane Directorate is responsible for several aspects of aviation, such as:

1. Continued airworthiness and general aviation safety.
2. Type certification
3. Technical Standard Orders (TSO)
4. Parts manufacturer approval (PMA)
5. Field approval.[13]

9.1.7.3 The Transport Airplane Directorate

The Transport Airplane Directorate (Northwest Mountain Region), functionally, has oversight responsibility for transport category airplane design approvals and modifications worldwide, as well as oversight responsibility for over 900 production approval holders. The Transport Airplane Directorate works closely with other FAA offices throughout the country and with foreign regulatory authorities to accomplish this goal.

Among the FAA offices working with the Directorate, it is worth mentioning:

1. The **Aircraft Certification Services (ACOs;** in Seattle, Los Angeles, and Denver).
2. The **Manufacturing Inspection District Offices (MIDOs;** in Renton, WA, Phoenix, AZ, Lakewood, CA, and Van Nuys, CA) – one Certificate Management Office in Renton, WA, with Certificate Management Units in Everett, WA, Renton, WA, Auburn, WA, and Long Beach, CA.

The Directorate relies on **Designated Representatives**[14] of the Administrator to act on behalf of the FAA. This Designee force includes Engineering Designees, Manufacturing Designees, and Organization Designees.

The Directorate's three most important responsibilities are:

1. Continued operational safety
2. Regulations and policy for all transport category airplanes
3. Design, production, and airworthiness certification.

9.1.7.3.1 Continued operational safety

1. Monitoring the transport category airplane fleet to ensure that airplanes continue to meet regulations and are safe throughout their operational life cycle.
2. Looking for conditions that affect the safety of airplanes. This is done by surveillance, inspection, review, investigation and analysis of service difficulties, incidents, and accidents.
3. If an unsafe condition is identified, this will trigger the following actions:
   (a) Work with the manufacturers to mandate corrective action through Airworthiness Directives (ADs), or
   (b) Revision of regulations/policy, or
   (c) Issuing of new regulations/policy.
4. Performing surveillance and oversight of production approval holders.

9.1.7.3.2 Regulations and policy for all transport airplanes

1. Developing and establishing FAA type design and airworthiness standards for all transport category airplanes.
2. The type design standards are codified in Title 14, Code of Federal Regulations (14 CFR), Part 25. This is commonly referred to as Part 25 of the Federal Aviation Regulations (FAR).
3. These FAR 25 standards are applied to aircraft worldwide, working with other civil aviation authorities to 'harmonize' these standards whenever possible.

9.1.7.3.3 Design, production, and airworthiness certification

1. The Directorate is responsible for the release of design, production, and airworthiness approvals of all aircraft and aircraft parts in Washington, Oregon, Idaho, Montana, Colorado, Wyoming, California, Arizona, Utah, Nevada, Hawaii, and the Pacific Rim countries.
2. Determining if and ensuring that each aircraft design meets the applicable regulations **(design certification)**.
3. When an applicant shows that its aircraft design meets the standards, a type certificate is issued.
4. Ensuring that each manufacturing facility is capable of producing aircraft to the approved **(production certification).**
5. Ensuring that each aircraft produced in the manufacturing facility is built to the approved design.
6. Ensuring that each aircraft produced is in a condition for safe operation **(airworthiness certification).**

9.1.7.4 The Rotorcraft Directorate

The Rotorcraft Directorate (Southwest Region) approves the design and production for all aircraft manufactured or modified in Texas, New Mexico, Oklahoma, Arkansas, and Louisiana for those seeking a type certificate, Supplemental type certificate, parts manufacturer approval, or a Technical Standard Order approval.

In addition to certifying all aircraft, the Directorate has the responsibility for writing rules and policy for rotorcraft and working with all the **Aircraft Certification Offices (ACOs)** – also outside the above-mentioned territory – to achieve standardized application of the rules for rotorcraft. Furthermore, it works with its counterparts in other countries to issue domestic approvals for foreign manufactured rotorcraft.

The Rotorcraft Directorate has one **Manufacturing Inspection District Office (MIDO)** in Oklahoma City and one in San Antonio. The rest of the Directorate, which includes three **Aircraft Certification Offices (ACOs),** a **Manufacturing Inspection District Office (MIDO),** the **Directorate Staff,** and the necessary administrative support and management, is located in the Southwest Regional Office, 2601 Meacham Blvd, Fort Worth, TX 76193.

1. Of the three Aircraft Certification Offices, two certify the engineering design of airplanes and one certifies rotorcraft.
2. The Manufacturing Inspection District Offices certify the production aspects of all products.
3. The Rotorcraft Standards Staff writes rules/policy and establishes standardized application of the rotorcraft certification rules, and
4. The necessary administrative support staff and management to operate the organization.

9.1.7.5 The Engine and Propeller Directorate

The Engine and Propeller Directorate (New England Region) is located in Burlington, MA. It is responsible for original type certification or changes to approved designs of aircraft engines and propellers in addition to Technical Standard Order (TSO) approvals of auxiliary power units (APUs).

The Engine and Propeller Directorate (E&PD) is responsible for developing rules, policy, and guidance for these products, and assures standardization across all FAA **Aircraft Certification Offices (ACOs)** that perform certification work on these products. The E&PD Standards Staff is the working element of the E&PD that directly carries out these functions.

The **Engine Certification Office (ECO)** and each of the **Aircraft Certification Offices (ACOs)** that perform E&PD-related certification work are accountable for planning, directing, and controlling engine and propeller type certification programs in addition to TSO approvals of auxiliary power units. Both the ECO's and ACOs' primary responsibilities are to find compliance to the applicable Airworthiness Standards (i.e. FAR 33 and 35 and TSO-C77B) and assure continued airworthiness of these products once in service.

The **Manufacturing Inspection Office (MIO)** is responsible for assuring aviation parts are manufactured to approved standards and issuing production certificates to manufacturers in accordance with the requirements of FAR 21.

9.1.8 'One World, One Goal: Aviation Safety'

In this chapter, in dealing with the JAA, we have emphasized the necessity of having in place a legally recognized European authority. In fact, in spite of a huge amount of work accomplished for unification of regulations and procedures in Europe, the JAA did not have the authority to impose these rules.

The EASA now has this power and can perform as a single authority. For instance, once an aircraft is type certificated by the EASA, this type certificate is valid for all the Member States, without being just a 'recommendation' for the issue of a national type certificate. Today, we have a single European Agency instead of 25 national authorities, and a single certificate for aeronautical products instead of 25.

Another shortcoming of the JAA was the complexity of bilateral agreements with authorities like the FAA or Transport Canada. For example, an Airbus certificated by the JAA could be accepted in the United States only when it was in possession of a type certificate issued by alityrequires European Member State.

The JAA has carried out long and complex work with the FAA and Transport Canada for the release of new bilateral agreements, also relating to single European Member States.

The new legal reality requires European Member States to comply with European Law; they cannot deviate from common European rules, nor impose additional requirements or conclude agreements with third countries. As a consequence, Member States are represented by the EASA. Furthermore, Member States are bound by and must reflect the Agency's decisions and positions when carrying out their representative roles in frameworks like the ICAO and ECAC.

The Agency is committed to establishing proper relations with non-EU members of the ECAC and to pursue relationships with other international partners through special arrangements, associations, partnerships, and mutual recognition agreements. It must also recognize that, legally, bilateral safety agreements are a competence of the European Commission.

At present, although the EASA has already agreed to some working arrangements with a certain number of non-EU states (Brazil, Canada, China, Israel, Russia, and the USA), no bilateral agreement has been formalized. Therefore, from a strictly legal point of view, the existing bilateral agreements of the EU Member States are still in force.

In this context, the EASA is carrying on the tradition of an annual US-Europe International Aviation Safety Conference. The Europe-US Aviation Safety Conference has been taking place for 50 years to promote cooperation and mutual recognition of safety standards.

The Europe-US International Aviation Safety Conference on 7–9 June 2005 was jointly organized by the European Aviation Safety Agency (EASA), the Joint Aviation Authorities (JAA), and the Federal Aviation Administration (FAA) of the USA. For many years, this event provided a forum for open discussion between the JAA and other civil aviation authorities and industry representatives on current initiatives and strategic directions. Today, this annual conference also provides a forum for interested parties to participate in harmonization and safety enhancement activities, and to present initiatives of their own to the global community.

More than 350 high-level aviation experts from all over the world came together in Cologne, Germany, to discuss future trends in aviation safety. Under the title 'Aviation Safety Regulation – Setting the Sights for the Future', this conference hosted by the EASA focused on bilateral agreements and future regulation in aviation safety.

In opening the conference, Patrick Goudou, Executive Director of the EASA, said:

> *Our mission is to set and achieve the highest common standards of safety and environmental protection in civil aviation. I am confident we can achieve our goals through international co-operation and a strong partnership with the United States in particular.*

Notes

1. This can be considered as the part of an aviation authority dealing with airworthiness.
2. The ECAC (European Civil Aviation Conference) was founded in 1955 as an intergovernmental organization. The ECAC's objective is to promote the continued development of a safe, efficient, and sustainable European air transport system. In so doing, the ECAC seeks to harmonize civil aviation policies and practices amongst its Member States, and promote understanding on policy matters between its Member States and other parts of the world. Close liaisons are maintained with the ICAO, EUROCONTROL (see note 3), and the JAA.
3. EUROCONTROL has the role of co-ordinating the development of a uniform system of Air Traffic Management (ATM) throughout Europe (41 states), working with its partners in the air transport industry to provide a range of servicesfrom air traffic controller training to managing air traffic flow, from regional control of airspace to development of innovative technologies and procedures.
4. The **Advisory Body of Interested Parties** assists the Management Board in this workIt comprises organizations representing aviation personnel, manufacturers, commercial and aviation operators, the maintenance industry, training organizations, and air sport.
5. Aircraft for which a type certificate or a certificate of airworthiness has not been issued on the basis of this EASA Regulation and its implementing rules(Generally, aircraft with a Special certificate of airworthiness and ultralights.)
6. Patrick Goudou, Executive Director of the EASA.
7. President of MT Propeller Entwicklung.
8. This is another proposal following the US system where the FAA is supported by ' $1 per airline ticket' This is why the FAA does not charge for certification. ' To charge fees for certification could kill small and medium enterprises in aviation, especially in general aviation.'
9. Details on the quoted FARs can be found in Chapter 9.2.
10. The Federal Aviation Act of 1958 directs the FAA to promote safety of flight of civil aircraft in air commerce prescribing and revising minimum standards for design, materials, construction, etc. The ACRP was developed to accomplish this goal.
11. Special FARs establish additional airworthiness standards for aircraft to cope (normally) with particular operation. For instance, Special FAR No. 23 is for aircraft to be certificated in the Normal category for a reciprocating or turbopropeller multi-engine-powered small airplane that is to be certificated to carry more than 10 occupants and that is intended for use in operations under FAR 135.
12. MIDOs assist with: production approval and certification (manufacturing); airworthiness certification; manufacturing facilities approval holder issues; manufacturing designee oversight; support to ACOs during design approvals.
13. Field approval is a maintenance performance approval for a major repair or major alteration that is performed by a Flight Standards Service, Aviation Safety Inspector.
14. A Designee is a representative of the FAA Administrator authorized by law to examine, test, and/or make inspections necessary to issue airman or aircraft certificates.

Aerospace Engineering Desk Reference; ISBN 9781856175753

9.2 Airworthiness Requirements

Filippo De Florio

9.2.1 Requirements, Regulations, and Standards

Before dealing with EASA regulations, it is worth considering the JAA requirements, which are the basis of these regulations, and their relationship with their FAA analogs. Even if the JAA requirements are to be superseded, it is necessary to start with them in order to establish continuity and gain an understanding of their origin.

Having already mentioned the **standards** as the technical documents issued to define design criteria, we will now consider the **'requirements'** (in the JAA terminology) or **'regulations'** (in the FAA terminology): the compulsory standards.

The OSTIV,[1] for example, publishes a standard for the design of sailplanes and powered sailplanes entitled 'OSTIV Airworthiness Standard'. This document defines the organization's vision on this subject. However, if anyone applies for the certification of a sailplane in Europe, they must make reference to JAR *22*,[2] 'Sailplanes and Powered Sailplanes', because this is the only set of sailplane airworthiness standards with legal value, adopted by all JAA national authorities. This means that the OSTIV Standard[3] can only be a guide as well as a valuable reference point (also for the JAR *22* Study Group).

9.2.2 JARs and FARs

When the JAR requirements were first issued in the 1970s, several different standards for aircraft certification were in force in different countries. If we consider the western world only, among the most renowned we can quote the Federal Aviation Regulations (FARs) issued by the FAA, adopted in the United States as well as in many other countries. In the UK, for example, the Civil Aviation Authority that in 1972 replaced the Air Registration Board (ARB) made use of the British Civil Air Regulations (BCARs). In France the *Direction Générale de l'Aviation Civile* (DGAC) had the *Régles AIR*. In Germany the *Luftfahrt Bundesamt* had its own regulations for sailplanes. This situation posed many difficulties in aircraft exportation.

Finally, on 1 January 1992, the JARs became part of the regulations of the European Community, assuming legal status in the Community Countries (all existing equivalent regulations had to be superseded). At present, only JARs (now replaced by the EASA regulations, as we will see) and FARs (or derivative regulations) are in practical use.

9.2.3 List of JARs and FARs[4] Directly or Indirectly Related to Airworthiness Certification

9.2.3.1 JAR 1/FAR 1. Definitions and Abbreviations

These codes contain definitions and abbreviations of terms used in other JAR/FAR codes. JAR 1 is based partly on those definitions contained in ICAO Annexes and partly on FAR 1. FAR 1 also contains *rules of constructions*, i.e. characterization of wording like the use of *'shall'*, *'may'*, *'a person may not'*, *'includes'*.

9.2.3.2 JAR 11. JAA Regulatory and Related Procedures

This code contains the requirements applicable to:

1. The retention by the Central JAA of documents related to the development and production of JARs.
2. The format and structure of JARs.
3. The development of JARs and amendments to JARs until their publication by the JAA.
4. The procedures for granting exemptions in the JARs.
5. The procedures for consultation on special conditions.
6. The development of ACJ until their publication by the JAA.

9.2.3.3 FAR 11. General Rulemaking Procedure

This part applies to the issuance, amendment, and repeal of any regulation for which the FAA follows public rule-making procedures under the Administrative Procedure Act ('APA'). In this context the code prescribes requirements applicable to:

1. Procedures for issuing a rule, from the 'advanced notice of proposed rulemaking' (ANPRM), through the 'notice of proposed rulemaking' (NPRM), to the 'final rule'.
2. Petitions for exemptions (from individual or entity).
3. Petitions for rule-making (from individual or entity).
4. Issuing of special conditions.

9.2.3.4 JAR 21. Certification Procedures for Aircraft and Related Products and Parts

See relevant paragraph in this chapter.

9.2.3.5 FAR 21. Certification Procedures for Products and Parts

See relevant paragraph in this chapter.

9.2.3.6 JAR 22. Sailplanes and Powered Sailplanes[5]

See relevant paragraph in this chapter.

9.2.3.7 JAR-VLA. Very Light Aeroplanes[6]

See relevant paragraph in this chapter.

9.2.3.8 JAR 23. Normal, Utility, Aerobatic and Commuter Category Aeroplanes

See relevant paragraph in this chapter.

9.2.3.9 FAR 23. Airworthiness Standards: Normal, Utility, Acrobatic and Commuter Category Airplanes

See relevant paragraph in this chapter.

9.2.3.10 JAR 25. Large Aeroplanes

See relevant paragraph in this chapter.

9.2.3.11 FAR 25. Airworthiness Standards: Transport Category Airplanes

See relevant paragraph in this chapter.

9.2.3.12 JAR 26. Additional Airworthiness Requirements for Operations

This code prescribes specific additional airworthiness requirements with which operators must ensure that compliance has been established if operating in accordance with the Part of JAR-OPS relevant to the particular type of operations.

1. Subpart B relates to Commercial Air Transportation (Aeroplanes).
2. Subpart C (*reserved*) relates to General Aviation (Airplanes).
3. Subpart D (*reserved*) relates to Commercial Air Transportation (Helicopters).
4. Subpart E (*reserved*) relates to General Aviation (Helicopters).

9.2.3.13 JAR 27. Small Rotorcraft

See relevant paragraph in this chapter.

9.2.3.14 FAR 27. Airworthiness Standards: Normal Category Rotorcraft

See relevant paragraph in this chapter.

9.2.3.15 JAR 29. Large Rotorcraft

See relevant paragraph in this chapter.

9.2.3.16 FAR 29. Airworthiness Standards: Transport Category Rotorcraft

See relevant paragraph in this chapter.

9.2.3.17 FAR 31. Airworthiness Standards: Manned free Balloons[7]

See relevant paragraph in this chapter.

9.2.3.18 JAR-E. Engines

This code is based on the English BCAR Section C and contains the airworthiness requirements for engines. Subsections B and C deal specifically with piston engines; subsections D and E deal specifically with turbine engines.

9.2.3.19 FAR 33. Airworthiness Standards: Aircraft Engines

This part prescribes airworthiness standards for the issue of type certificates for aircraft engines and changes to those certificates. Subparts C and D deal specifically with reciprocating aircraft engines; Subparts E and F deal specifically with turbine aircraft engines.

9.2.3.20 JAR-APU. Auxiliary Power Units[8]

This code is based on FAA Technical Standard Order TSO-C77a and provides airworthiness requirements for the release of JTSO authorizations for turbine powered auxiliary power units for use on aircraft.

9.2.3.21 FAR 34. Fuel Venting and Exhaust Emission Requirements for Turbine Engine Powered Airplanes[9]

The provisions of this subpart are applicable to all in-use aircraft gas turbine engines of the classes specified, certificated for operations within the United States.

As regards foreign airplanes, this FAR applies only to those foreign civil airplanes that, if registered in the United States, would be required by applicable Federal Aviation Regulations to have a US standard airworthiness certificate in order to conduct the operations intended for the airplane.

9.2.3.22 JAR-P. Propellers

The requirements of this code apply to propellers of conventional design.

9.2.3.23 FAR 35. Airworthiness Standards: Propellers

This part prescribes airworthiness standards for the issue of type certificates and changes to those certificates for propellers.

9.2.3.24 JAR 36. Aircraft Noise

The applicable noise requirements for the publication of a type certificate for an aircraft are prescribed according to the provisions of Chapter 1 of ICAO Annex 16.

9.2.3.25 FAR 36. Noise Standards: Aircraft Type and Airworthiness Certification

This part prescribes noise standards for the issue of the following certificates:

1. Type certificates, and changes to those certificates, and standard airworthiness certificates for subsonic transport category large airplanes and for subsonic jet airplanes regardless of category.

2. Type certificates, and changes to those certificates, standard airworthiness certificates, and restricted category airworthiness certificates for propeller-driven, small airplanes and for propeller-driven, commuter category airplanes, except those airplanes that are designed for 'agricultural aircraft operations' (as defined in FAR 137.3, as effective on 1 January 1966) or for dispersing fire-fighting materials to which FAR 36.1583 does not apply.
3. A type certificate, and changes to that certificate, and standard airworthiness certificates for Concorde airplanes.
4. Type certificates, and changes to those certificates, for helicopters, except those helicopters that are designated exclusively for 'agricultural aircraft operations', for dispensing fire-fighting materials or for carrying external loads.

9.2.3.26 FAR 39. Airworthiness Directives

The regulations in this part provide a legal framework for the FAA's system of Airworthiness Directives.[10]

9.2.3.27 FAR 43. Maintenance, Preventive Maintenance, Rebuilding, and Alterations

See relevant literature for details.

9.2.3.28 FAR 45. Identification and Registration Marking

This part prescribes the requirements for:

1. Identification of aircraft, and identification of aircraft engines and propellers, that are manufactured under the terms of a type or production certificate.
2. Identification of certain replacement and modified parts produced for installation on type certificated products.
3. Nationality and registration marking of US-registered aircraft.

9.2.3.29 JAR-TSO. Joint Technical Standard Orders

While the requirements for issue of JTSOs are found in JAR 21 Subparts O and N–O, the code provides the list of JTSOs as follows:

- **Index 1:** the JTSOs that are technically similar to FAA TSOs.
- **Index 2:** the JTSOs that are applicable only to JAR (different from FAA TSOs, or corresponding FAA TSOs not existing).

9.2.3.30 JAR-OPS 1. Commercial Air Transportation (Aeroplanes)

This code prescribes requirements applicable to operation of any civil aeroplane for the purpose of commercial air transportation by any operator whose principal place of business is in a JAA Member State, with exceptions indicated in the same code.

9.2.3.31 JAR-OPS 3. Commercial Air Transportation (Helicopters)

This code prescribes requirements applicable to any civil helicopter for the purpose of commercial air transportation by any operator whose principal place of business is in a JAA Member State, with exceptions indicated in the same code.

9.2.3.32 JAR-MMEL/MEL. Master Minimum Equipment List/Minimum Equipment List

See relevant literature for details.

9.2.3.33 FAR 91. General Operating and Flight Rules

Except as provided in cases indicated, this part prescribes rules governing the operation of aircraft (other than moored balloons, kites, unmanned rockets, and unmanned free balloons, which are governed by FAR 101, and ultralight vehicles operating in accordance with FAR 103) within the United States, including the waters within three nautical miles of the US coast.[11]

9.2.3.34 FAR 101. Moored Balloons, Kites, Unmanned Rockets, and Free Balloons

This part prescribes rules governing the operation in the United States, of moored balloons, kites, unmanned rockets, and free balloons, whose characteristics and limitations (as applicable: weight, gas capacity, quantity and quality of propellant, etc.) are defined.

9.2.3.35 FAR 103. Ultralight Vehicles

This part prescribes rules governing the operation of ultralight vehicles in the United States. For the purposes of this part, ultralights are defined in terms of maximum weight (powered and unpowered), maximum speed (powered), and maximum stalling speed; the operations are limited to a single occupant and their use to recreation or sport purposes only.

9.2.3.36 FAR 119. Certification: Air Carriers and Commercial Operators

This part applies to each person operating or intending to operate civil aircraft as an air carrier or commercial operator, or both, in air commerce or, when common carriage is not involved, in operations of US-registered civil airplanes with a seat configuration of 20 or more passengers, or a maximum payload capacity of 6000 lb or more. This part prescribes in particular the certification requirements an operator must meet in order to obtain and hold a certificate authorizing operations under FAR 121, 125, or 135.

9.2.3.37 FAR 121. Operating Requirements: Domestic, Flag, and Supplemental Operations

This part prescribes rules governing (in particular):

1. The domestic, flag, and operations of each person who holds an Air Carrier Certificate or Operating Certificate under FAR 119.
2. Each person employed or used by a certificate holder conducting operations under this part, including maintenance, preventive maintenance, and alteration of aircraft.

9.2.3.38 FAR 125. Certification and Operations: Airplanes Having a Seating Capacity of 20 or More Passengers or a Maximum Payload Capacity of 6000 Pounds or More; and Rules Governing Persons on Board Such Aircraft

This part prescribes rules governing the operations of the above-mentioned US-registered civil airplanes when common carriage is not involved, unless they are required to be operated under FAR 121, 129, 135, or 137, and unless other cases described in this part are applicable.

9.2.3.39 FAR 129. Operations: Foreign Air Carriers and Foreign Operators of US-Registered Aircraft Engaged in Common Carriage

This part prescribes rules governing the operations within the Unites States of each foreign air carrier holding (defined) permits issued by the Civil Aeronautic Board of the US Department of Transportation.

9.2.3.40 FAR 133. Rotorcraft External-Load Operations

This part prescribes airworthiness and operating certification rules for rotorcraft used in the above-mentioned operations in the United States by anyone, with the exceptions defined in the same document.

9.2.3.41 FAR 135. Operating Requirements: Commuter and on-Demand Operations and Rules Governing Persons on Board Such Aircraft

This part prescribes rules governing the commuter or on-demand operations of each person who holds or is required to hold an Air Carrier Certificate or Operating Certificate under FAR 119 and relevant items.

9.2.3.42 FAR 137. Agricultural Aircraft Operations

This part prescribes rules governing agricultural operations within the United States and the issue of commercial and private agricultural aircraft operator certificates for those operations.

9.2.3.43 FAR 145. Repair Stations

This part describes how to obtain a repair station certificate. This part also contains the rules a certificated repair station must follow relating to its performance of maintenance, preventive maintenance or alterations of an aircraft, airframe, aircraft engine, propeller, appliance, or component part to which FAR 43 applies. It also applies to any person who holds, or is required to hold, a repair station certificate issued under this part.

9.2.3.44 FAR 147. Aviation Maintenance Technician Schools

This part prescribes the requirements for issuing aviation maintenance technician school certificates and associated ratings and the general operating rules for the holders of those certificates and ratings.

9.2.3.45 JAR-AWO. All Weather Operations

This code prescribes requirements for:

1. Automatic landing systems.
2. Airworthiness certification of aeroplanes for operations with decision heights of 60 m (200 ft) down to 30 m (100 ft) – Category 2 operations.
3. Airworthiness certification of aeroplanes for operations with decision height below 30 m (100 ft) or no decision height – Category 3 operations.
4. Directional guidance for take-off in low visibility.

9.2.3.46 JAR/CS-VLR. Very Light Rotorcraft

See relevant paragraph in this chapter.

9.2.3.47 References for Certification of Parts of Aircraft

1. Joint Technical Standard Order authorization (JTSO) (JAR 21 Subpart O).
2. Technical Standard Order (TSO) (FAA AC 20–110).
3. Joint Part Approval authorization (JPA) (JAR 21 Subpart P).
4. Part Manufacturer Approval (PMA) (FAR 21.303).
5. Military and industrial specifications.
6. Specifications written in the aircraft certification process.

9.2.3.48 General Remarks

The airworthiness standards dealing with the same products[12] have been put in sequence in the above list. The list shows the existence of operational standards in addition to the product type certification standards. These operational standards contain airworthiness requirements that influence the aircraft configuration in relation to their particular operations.[13]

A JAR/FAR 23 aeroplane, for example, can obtain a type certificate with the installation (as flight and navigation instruments) of an airspeed indicator, an altimeter, and a magnetic direction indicator only. However, in order to obtain a certificate of airworthiness (the document which authorizes the flight), other instruments and equipment that depend on the particular type of operation (for example, tourism, aerial work) and on the flight conditions (VFR, IFR, night flight, etc.) must be installed as prescribed by the operational rules.

Furthermore, the environmental standards such as FAR 34 and JAR/FAR 36 must be considered. For the FAA and EASA, compliance with the environmental protection requirements is part of the type certification.

The JAA requirements are adopted by the JAA Member States, the EASA requirements are now used by the EU Member States, and the FAA regulations are used in the United States[14] (Canada has equivalent rules).

Nevertheless, the manufacturing companies wanting to sell their products on both sides of the Atlantic must perform a double certification, with a substantial increase in costs, especially when the standards are different. For many years the transport aeroplane industry has been penalized because the contents of JAR 25 and FAR 25 were not equivalent (and with the same paragraph numbering). JAR 25 originated under a strong English influence, with philosophies borrowed from BCAR Section D. It is also because of manufacturers' complaints that, for many years now, the JAA and FAA have carried out a harmonization process which is well advanced, but not yet accomplished. The situation is better for JAR/FAR 23 aeroplane standards and for the JAR/FAR 27 and 29 rotorcraft[15] standards,

because the JAR requirements were produced with the co-operation of the FAA, with the common will of avoiding the situation that has penalized transport aeroplanes: these standards are now almost harmonized.

9.2.4 Advisory Material

Some rules can be interpreted in different ways. This is the reason why the authorities issue advisory material for the explanation of the rule or, in certain cases, suggest suitable procedures to perform a demonstration of compliance to the same rule.

The FAA publishes 'Advisory Circulars' (ACs) as documents separate from the standards, while the JAA and EASA include similar documents at the end of the JAA/EASA standards.

If we look at the JARs, in Section 9.2.2, these standards contain the 'Advisory Circulars – Joint' (ACJs) that are 'Acceptable Means of Compliance, and Interpretations'.[16] The ACJs provide a means, *but not the only means*, by which a requirement can be met.[17] A numbering system is adopted in which the ACJ uses the same number as the paragraph of the JAR to which it is related.

By the same approach, the EASA Certification Standards (CS) contain the 'Acceptable Means of Compliance' (AMC), with the same meaning as the ACJs.

For the Implementing Rules of the EASA, like Part 21, Part M, Part 145, etc., documents containing the Acceptable Means of Compliance (AMC) and Guidance Material (GM) have been issued.

The AMC have the meaning already defined, while the GM helps to illustrate the meaning of a specification or requirements.

9.2.5 EASA Regulations

Figure 9.2-1 depicts the EASA's regulation organizational structure.

9.2.5.1 The Basic Regulations

The Basic Regulations (EC) of the EASA are contained in No. 1592/2002 already mentioned, which outlines the tasks of the Agency starting from the necessity that 'a high and uniform level of protection of the European citizen should at all

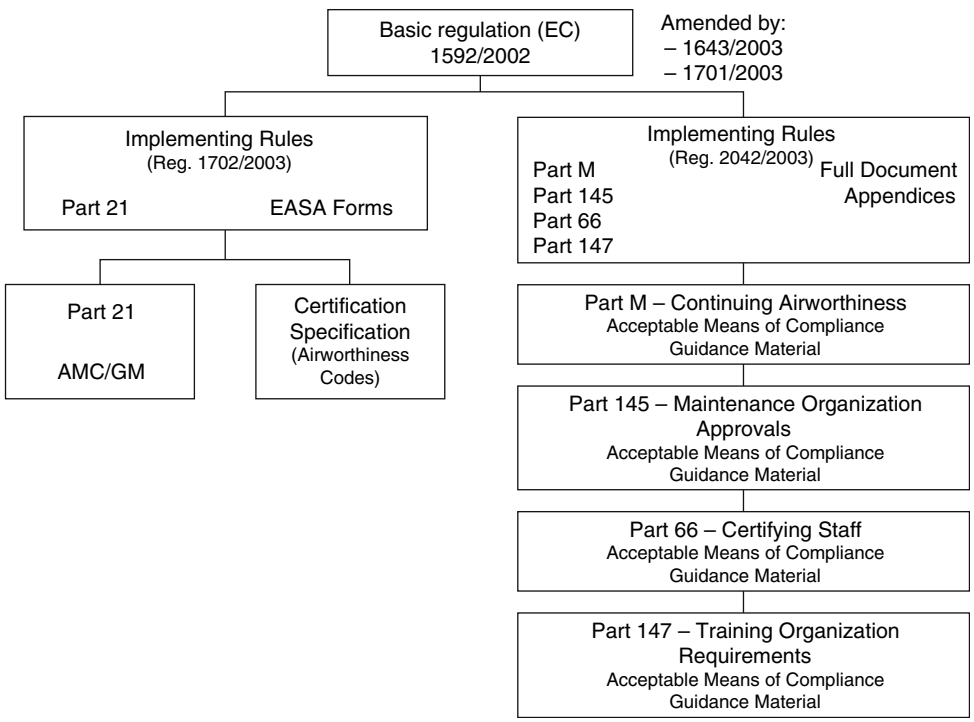


Figure 9.2-1 EASA regulations.

times be ensured in civil aviation, by the adoption of common safety rules and by measures ensuring that products, persons and organizations in the Community comply with such rules and those adopted to protect the environment.'

'As a consequence, aeronautical products should be subject to certification to verify that they meet essential airworthiness and environmental protection requirements relating to civil aviation. Appropriate essential requirements should be developed within one year after the entry into force of this Regulation to cover operation of aircraft and flight crew licensing and application of the Regulation to third-country aircraft and, thereafter, other areas in the field of civil aviation safety.'

On these bases, the EU Member States must adopt common rules in matters of civil aviation.

The EASA has produced, with the JAA's co-operation, a set of regulations that we will now summarize.

The Basic Regulations define the activities to be achieved through formalities detailed in second-level regulations adopted by the European Commission and defined as follows.

9.2.5.2 Implementing Rules (IR)

These Implementing Rules contain documents defined as **Parts**, which are divided into two sections: Section A, detailing the requirements to be satisfied by aeronautical subjects; and Section B, containing the procedures to be followed by the national authorities.

The Implementing Rules (EC) No. 1702/2003 of 24 September 2003 for the airworthiness and environmental certification of aircraft and related products, parts, and appliances, as well as for the certification of design and production organizations, specify:

1. The issuing of type certificates, Restricted type certificates, and changes to those certificates.
2. The issuing of certificates of airworthiness, Restricted certificate of airworthiness, permit to fly, and authorized release certificates.
3. The issuing of repair design approval.
4. The showing of compliance with environmental protection requirements.
5. The issuing of noise certificates.
6. The identification of products, parts, and appliances.
7. The certification of certain parts and appliances.
8. The certification of design and product organizations.
9. The issuing of Airworthiness Directives.

Annex to this document is **Part 21, 'Certification of aircraft and related products, parts and appliances, and design and production organizations'.**

This document replaces JAR 21, which remains the core of the same document. The changes to the JAR document reflect the new legal status of the EASA towards the national authorities, and a full revision of the document in the light of the JAA certification experience.

The Implementing Rules (EC) No. 2042/2003 of 20 November 2003 specify the continuing airworthiness of aircraft and aeronautical products, parts, and appliances, and the approval of organizations and personnel involved in these tasks.

Annexes to this document are:

1. **Annex I, Part M** establishes the measures to be taken to ensure that airworthiness is maintained, including maintenance. It also specifies the conditions to be met by persons and organizations involved in such continuing airworthiness management.
2. **Annex II, Part 145** establishes the requirements to be met by an organization to qualify for the issue or continuation of an approval for the maintenance of aircraft and components.
3. **Annex III, Part 66** establishes the requirements for the issue of an aircraft maintenance license and conditions of its validity and use, for aeroplanes and helicopters.
4. **Annex IV, Part 147** establishes the requirements to be met by organizations seeking approval to conduct training and examination as specified in Part 66.

9.2.5.3 Acceptable Means of Compliance (AMC) and Guidance Material (GM) for Part 21, Part M, Part 145, Part 66, and Part 147[18]

As already mentioned, Acceptable Means of Compliance (AMC) illustrate a means, but not the only means, by which a specification contained in an airworthiness code or a requirement in an implementing rule can be met. Guidance Material (GM) helps to illustrate the meaning of a specification or requirement.

9.2.5.4 Airworthiness Codes

All airworthiness codes are directly derived from the JARs. The JAR denomination has been changed in CS (Certification Specification).

Currently, the airworthiness codes are the following:

1. **CS-Definitions.** Derived from JAR **1.**
2. **CS-22 Sailplanes and Powered Sailplanes.** Derived from JAR 22.
3. **CS-23 Normal, Utility, Acrobatic and Commuter Aeroplanes.** Derived from JAR 23.
4. **CS-25 Large Aeroplanes.** Derived from JAR 25.
5. **CS-27 Small Rotorcraft.** Derived from JAR 27.
6. **CS-29 Large Rotorcraft.** Derived from JAR 29.
7. **CS-VLR Very Light Rotorcraft.** Derived from JAR-VLR.[19]
8. **CS-VLA Very Light Aeroplanes.** Derived from JAR-VLA.
9. **CS-E Engines.** Derived from JAR-E.
10. **CS-P Propellers.** Derived from JAR-P.
11. **CS-34 Aircraft Engine Emission and Fuel Venting.** Derived from JAR 34.
12. **CS-36 Aircraft Noise.** Derived from JAR 36.
13. **CS-APU Auxiliary Power Units.** Derived from JAR-APU.
14. **CS-ETSO European Technical Standard Orders.** Derived from JAR-TSO.
15. **CS-AWO All Weather Operations.** Derived from JAR-AWO.

For the certification of parts of aircraft, the references are the following:

1. European Technical Standard Order authorization (ETSO) (Part 21 Subpart O).
2. Specifications written in the aircraft certification process.
3. Standard parts in accordance with officially recognized standards.

9.2.6 General Considerations on Airworthiness Standards

Before considering the single standards (at least the more representative ones relating to the information this book is aimed to provide), it is worth considering the 'philosophies' that are the basis of their compilation.

9.2.6.1 Publication

The standards are made by Working Groups that are responsible for their compilation and amendments. Before publication, the Authorities concerned (the JAA, FAA, or EASA) submit the standards to public evaluation, allowing interested people and organizations to express their opinions on the matter. The relevant rules and the procedures for these phases are contained in JAR 11, 'Regulatory and Related Procedures', and in FAR 11, 'General Rulemaking Procedures'.

The EASA does not yet have a similar standard.

9.2.6.2 Special Conditions

As mentioned earlier, the standards do not anticipate aeronautical progress. Therefore, in several cases a 'non-conventional aircraft' is the object of the certification, or one with some peculiarities for which the 'applicable' airworthiness requirements of the relevant JAR/FAR/CS do not contain adequate or appropriate safety standards (Figures 9.2-2). As we have also considered that a 'blocked' standard might prevent aeronautical progress, what should be done in such situations? JAR/FAR 21, paragraph 16, and EASA Part 21, paragraph 21A.16B, provide an answer mentioning **'special conditions'**. It is a matter of adding such safety standards as the authority finds necessary to establish a level of safety equivalent to that established in the applicable JAR/FAR/CS. The special conditions are issued in accordance with JAR/FAR 11.

We will return to the 'level of safety' concept. However, to mention just one of the numerous possible examples, special conditions were issued for turbine engine installations on FAR 23 aircraft when FAR 23 did not yet contain safety standards for this kind of installation. It is not difficult to imagine the number of special conditions issued for the certification of 'Concorde' in the 1960s.

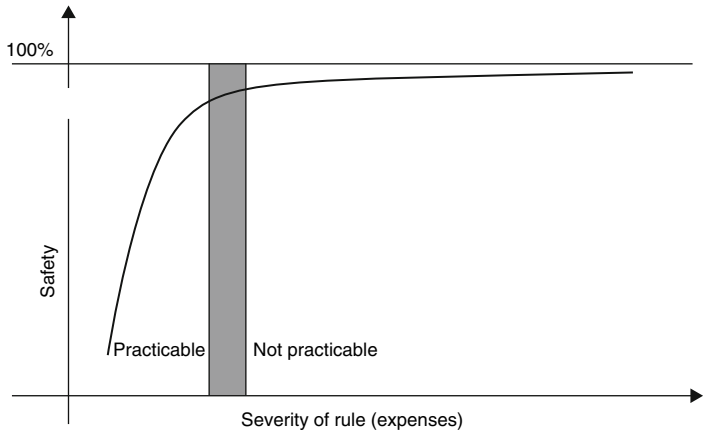


Figure 9.2-2 Airworthiness rules

In many cases, if design peculiarities that require special conditions become commonplace in the aeronautical field – 'winglets', for example – such special conditions are included (after discussions and evaluations according to JAR/FAR 11) in the JAR/FAR/CS standards via amendments.

9.2.6.3 Severity of the Airworthiness Standards

The 'level of safety' concept is a matter of serious concern regarding the compilation of the standards. The authorities could be tempted to play safe by issuing very restrictive standards. The immediate result would be to make it impossible for an aircraft to be certified for technical or simply for economical reasons.[20] Within airworthiness standards it is therefore necessary to balance criteria of **acceptability** (from the safety point of view) and the **practicability** of the same criteria.

The application of a rule involves expense. Increase of safety is not always proportional to the severity of the rule, even before considering the expense: at and beyond a certain point, negligible safety increases incur great expenditure. At this point the rule is no longer 'practicable'.

Various airworthiness standards have been produced for different types of aircraft (aeroplanes, rotorcraft, etc.), but also for different categories of the same type of aircraft (for weight, passengers number, etc.). An attempt has been made to arrange the aircraft in groups that are as 'homogeneous' as possible. Neglecting the obvious necessity to distinguish free balloons from transport aeroplanes, a distinction was made, for example, among the categories of aeroplanes named Normal, Utility, Aerobatic, and Commuter in JAR/FAR 23 and the Large Aeroplanes/Transport category airplanes in JAR/FAR 25. We must not infer that the airworthiness standards are different because the transport aeroplanes should be safer than other types of aircraft. Safety must be maximized for all aircraft, taking into account the criteria of 'practicability' mentioned above. As a fundamental concept, simple aircraft should have simple airworthiness standards to comply with.

It is certainly less easy to understand why a 19 000-lb commuter certified according to JAR/ FAR/CS-23, if it 'adds on weight' by some pounds, then comes under the JAR/FAR/CS-25 regulations. However, it is clear that the creation of classes implies that significant parameters are established, consequently involving precise numbers.[21] The designer should be capable of choosing the right airworthiness standard in light of the possible development of the project.

In any case, it is worthwhile considering that airworthiness, like medicine, is not an exact science!

9.2.6.4 Stalling Speed for Single-Engine Aeroplanes

In cases of 'acceptability', 'practicability', and examples of 'philosophies', we will now see how single-engine aeroplanes are regulated from a particular point of view: the stalling speed. A single-engine aeroplane, in case of engine failure, by definition can only glide. If this condition is not manageable in safety terms, it should 'never' happen. In fact, despite the great progress made in engineering techniques, the engine that 'never'[22] fails does not exist. We can also add that the present engine failure rate should not be compatible – and therefore not acceptable – with safety if any engine failure were to cause an accident. It is then necessary that the gliding and (especially) the power-off landing of a single-engine aeroplane be managed by a pilot of average skill.[23] It is evident that the result of an out-of-field landing is mainly influenced by the approach speed. However, the minimum gliding approach speed in

the landing configuration is a function of the power-off stalling speed in the same configuration; hence, a limitation of this speed is required. As a result, the stalling speed of single-engine aeroplanes in landing configuration (V_{so}) is limited to 61 knots. The same limitation exists for twin-engine aeroplanes that cannot meet a certain minimum rate of climb with an inoperative engine.

For all other twin-engine aeroplanes (even with an engine failure probability that is double that of a single-engine aeroplane), the probability of a twofold engine failure in the same flight is considered close to 'never', and therefore acceptable, so that no stalling speed limit is prescribed.[24]

It is interesting to note that, on the basis of the above-mentioned principles, JAR-VLA[25] contains a speed limitation in landing configuration of 45 knots, because it allows the installation of JAR 22 powered sailplane engines that, at least in principle, are considered less reliable than the engines installed on JAR/FAR 23 aeroplanes.

A stalling speed limitation in landing configuration was also introduced in JAR 22, because the trend towards the increase of water ballast quantity for speed contests was producing such an increase in wing loading, and therefore in the stalling speed, to jeopardize the possibility of a safe landing in the case, for example, of an aborted take-off or breaking of the tow rope; normally, in these cases, there is insufficient time to dump the water ballast.

9.2.6.5 Crashworthiness

We have mentioned a stalling speed limitation of 61 knots. However, is it really true that such a limitation could produce a safe power-off landing?

When limitations of this type are introduced, generally they are the result of experience and analysis of accidents that occurred in relevant situations. They are certainly not chosen at random. Nevertheless, the limitation cannot take into consideration all the conditions of the area where the aeroplane is likely to land (or crash if the ground is particularly uneven). Then the possibility of a crash must be considered, for whatever reasons and not only for single-engine aeroplanes. The airworthiness standards have become more and more stringent from this point of view. This is what we call **crashworthiness.**

JAR/FAR/CS-23 contains appropriate safety standards for emergency landing conditions. It deals with structural rules for the occupants' protection, also requiring expensive static and dynamic tests for the seat/restraint system, the seats, and the fuselage structure supporting the same.

FAR 23 prescribed something more. In order to allow the certification of speedy single-engine aeroplanes (turbine engine-powered), the design of which is severely penalized by the 61-knot limitation, these regulations enable an increase of stalling speed to be 'exchanged' with the additional severity of the crashworthiness regulations; we again encounter the acceptability/practicability balance. At the time of writing, JAR 23/CS-23 has not (yet) adopted this amendment.

Crashworthiness concerns all types of aircraft.

The JAR 22 Study Group tried to avoid dynamic tests for aircraft such as sailplanes and powered sailplanes. These aircraft are normally produced in such small numbers that it becomes economically difficult for the manufacturers to sustain the cost of dynamic crash tests. Nevertheless, the problem does exist and it is a serious one because these machines, for which an out-of-field landing is not even an emergency, tend to crash land quite frequently. We therefore have to consider the classical configuration of these aircraft that, in principle, does not offer suitable protection for the occupants.

We could think of a 'survival cage' able to bear some tens of grams, but this is not the solution because, even if the cage does not break, the occupants could still sustain extended or maybe fatal injuries. The OSTIV Sailplane Development Panel (SDP) studied this problem for some time, appointing a Crashworthiness Panel and reaching solutions recalling (to a certain extent) those adopted in Formula 1. The criteria they adopted could be summarized as 'stiff cage and soft nose', i.e. a sufficiently strong structure to protect the occupants but with a yielding front part, able to absorb impact energy. The OSTIV also provided advice on the seat design; these should be devised as 'energy absorbing'.[26] Standards for headrests were introduced, very effective items in the rebound phase after impact. The seat profile and the safety harness configuration were studied; the accident analysis pointed out the possibility of spine damage due to sliding under the safety harness in the impact phase, a movement defined as 'submarining'. Furthermore, the accident analysis showed that the landing gear standards did not offer sufficient energy absorption, with consequences for the occupants' spine. Therefore, these standards were improved.

The criteria coming from the OSTIV SDP are very often introduced in JAR 22, after evaluation of the relevant Study Group.

JAR-VLA contains a paragraph dealing with crash landing, which has not been updated since the original publication in 1990 and could be considered in need of modernization.

Dynamic crash tests should also be avoided for these aeroplanes, but an update of the crash-worthiness criteria is reasonable – for instance, taking into consideration the studies performed for sailplanes.

The airworthiness standards for transport aeroplanes (JAR/FAR/CS-25) and for rotorcraft (JAR/FAR/CS-27 and -29) contain paragraphs on crash landing inclusive of dynamic crash tests.

9.2.6.6 Fire Protection

An aircraft has engines, electrical installations, and other components, making it subject to fire hazard. Firstly, the 'fire zones' of the aircraft, i.e. those where a fire can develop, must be located – an engine compartment, for example. There are essentially three methods of protecting the occupants from fire: (a) abandoning the aircraft;[27] (b) passive protection in order to contain the fire for the time necessary for landing; (c) active protection by means of extinguishers. Of course, the combination of these last two means is possible. For military aircraft, normally carrying explosive material, abandoning the aircraft is favored (unless the fire is so limited that it can be put out by means of extinguishers), the active or passive protection being limited to the time necessary for the acknowledgment of the situation by the crew and their bailing out.

In the case of civil aircraft, passive protection is prescribed in order to allow a safe emergency landing whenever possible. This is achieved by suitable isolation of the fire zones so that essential structures and installations can be protected for the time necessary for landing. The use of extinguishers is not excluded, but they are not considered as primary protection.

Active protection, by means of portable or fixed extinguishers, is prescribed in some categories of aircraft (e.g. transport and commuter aeroplanes), for accidental fires in the cockpit, the cabin, and the baggage or cargo compartments.

The airworthiness standards also provide rules for materials used for the cabin interiors, from the points of view of flammability and noxious smoke emissions.

Because the requirements must normally be substantiated by tests, the certification standards provide acceptable procedures for such tests. To give an idea of the content of these documents, an example can be found in Appendix F to FAR 23, 'Test Procedure', of which an extract is reported here.[28]

9.2.6.6.1 Acceptable test procedure for self-extinguishing materials for showing compliance with paragraphs 23.853, 23.855, and 23.1359

1. **Conditioning.** Specimens must be conditioned to 70°F, ±5°F, and at 50 ±5 per cent relative humidity until moisture equilibrium is reached, or for 24 hours.
2. **Specimen configuration.** Except as provided for materials used in electrical wire and cable insulation and in small parts, materials must be tested either as a section cut from a fabricated part as installed in the airplane or as a specimen simulating a cut section, such as a specimen cut from a flat sheet of the material or a model of the fabricated part. The specimen may be cut from any location in a fabricated part; however, fabricated units, such as sandwich panels, may not be separated for a test. The specimen thickness must be no thicker than the minimum thickness to be qualified for use in the airplane, except that: (1) thick foam parts, such as seat cushions, must be tested at 1/2-inch thickness; (2) when showing compliance with paragraph 23.853(d)(3)(v) for materials used in small parts that must be tested, the materials must be tested at no more than 1/8-inch thickness; (3) when showing compliance with paragraph 23.1359(c) for materials used in electrical wire and cable insulation, the wire and cable specimens must be the same size as used in the airplane. In the case of fabrics, both the warp and fill directions of the weave must be tested to determine the most critical flammability conditions. When performing the tests prescribed in paragraphs (d) and (e) of this appendix, the specimen must be mounted in a metal frame so that: (1) in the vertical tests of paragraph (d) of this appendix, the two long edges and the upper edge are held securely; (2) in the horizontal test of paragraph (e) of this appendix, the two long edges and the edge away from the flame are held securely; (3) the exposed area of the specimen is at least two inches wide and 12 inches long, unless the actual size used in the airplane is smaller; and (4) the edge to which the burner flame is applied must not consist of the finished or protected edge of the specimen, but must be representative of the actual cross-section of the material or part installed in the airplane. When performing the test prescribed in paragraph (f) of this appendix, the specimen must be mounted in a metal frame so that all four edges are held securely and the exposed area of the specimen is at least eight inches by eight inches.
3. **Vertical test.** A minimum of three specimens must be tested and the results averaged. For fabrics, the direction of weave corresponding to the most critical flammability conditions must be parallel to the longest dimension. Each specimen must be supported vertically. The specimen must be exposed to a Bunsen or Tirrill burner with

a nominal 3/8-inch I.D. tube adjusted to give a flame of 1.5 inches height. The minimum flame temperature measured by a calibrated thermocouple pyrometer in the center of the flame must be 1550°F.

4. **Horizontal test.** A minimum of three specimens must be tested and the results averaged. Each specimen must be supported horizontally. The exposed surface when installed in the airplane must be face down for the test. The specimen must be exposed to a Bunsen or Tirrill burner with a nominal 3/8-inch I.D. tube adjusted to give a flame of 1.5 inches height. The minimum flame temperature is measured.
5. **Forty-five-degree test.** A minimum of three specimens must be tested and the results averaged. The specimens must be supported at an angle of 45° to a horizontal surface. The exposed surface when installed in the aircraft must be face down for the test. The specimens must be exposed to a Bunsen or Tirrill burner with a nominal 3/8-inch I.D. tube.
6. **Sixty-degree test.** A minimum of three specimens of each wire specification (make and size) must be tested. The specimen of wire or cable (including insulation) must be placed at an angle of 60°.
7. **Burn length.** Burn length is the distance from the original edge to the furthest evidence of damage to the test specimen due to flame impingement, including areas of partial or complete consumption, charring or embrittlement, but not including areas sooted, stained, warped or discolored, nor areas where material has shrunk or melted away from the heat source.

9.2.6.7 Safety Assessment

Let us consider the control system of a light aeroplane: cables, pulleys, perhaps some rods. These items are very often in view and easy to inspect. For such systems, if designed according to good design practice and applicable airworthiness standards and maintained following the maintenance manual instructions (providing the replacement of worn parts), no particular studies will be needed to assure the system's safety during the entire operating life of the aeroplane. We can therefore talk of a system that 'never' fails. It is quite different if, considering a more sophisticated aircraft, the control system depends on the electrical and hydraulic systems, or even the mechanical transmissions are eliminated, as for fly-by-wire systems, with computers playing an important part.

The above example on control systems can obviously be extended to all aircraft systems and equipment.

In this case, the safety assessment would require more refined rules and instruments. The essentially informative nature of this book cannot provide a thorough discussion on this very specific topic. Nevertheless, it is worth outlining some basic concepts.

The rules for safety assessment are contained in different aircraft airworthiness standards at paragraph XX.1309,[29] and advisory material in the respective ACJs/ACs/AMC&GM. As specified by the title of paragraph 1309, they are related to 'Equipment, Systems and Installations'.

As a consequence, these rules do not apply to performance, flight qualities, and structural strength of Subparts B and C,[30] but they can influence every system installed in order to comply with these subparts.

As a typical example (contained in FAA AC 23-1309-1A), paragraph 23.1309 does not apply to the stall characteristics of paragraph 23.201, but nevertheless it applies to a stick pusher (stall barrier) installed to satisfy the latter paragraph.

That being said, if we were to ask a layman (better still, a passenger) what kind of reliability a vital aircraft system should have, the answer would be immediate: 100 per cent. It has nevertheless been demonstrated that such reliability is an impossibility. As an example, setting in parallel 'n' items (redundancy), 100 per cent reliability can be obtained for 'n' tending to infinity!

A system with a high degree of redundancy would be heavy, expensive, and complex: so subject to drawbacks that it would make such redundancy questionable. It is then more convenient to design such systems with a minimum degree of redundancy (the reliability of the single components can be increased), in order that its reliability, even if not amounting to 100 per cent, is such as to ensure **an acceptable safety level.**

The definition of an acceptable safety level implies the definition of an **acceptable accident rate;** this cannot be defined as abstract wishful thinking, but on the basis of what is **practicable.**

What is practicable for the future can be forecast by the analysis of past accident rates. Therefore, after taking into consideration the accident rate in commercial (occidental) aviation in the 10-year period from 1970 to 1980, a rate of catastrophic accidents[31] a little less than 1×10^{-6} flight hours was detected. (An extrapolation forecasts a probable accident rate of 0.3×10^{-6} accidents in the 1990s). From this accident analysis it was also found that about 10 per cent of the catastrophic accidents could be attributed to system failures. Hence, the portion of catastrophic accidents attributed to systems was of the order of 1×10^{-7} flight hours.

Starting from the arbitrary hypothesis that a commercial large aircraft could present some 100 hazards (potential failure conditions) leading to a catastrophic effect, it follows that, for each system, the acceptable probability of a catastrophic failure is less then 10^{-9} flight hours.

This is the basic concept for 'the maximum probability of a catastrophic effect for a single system'[32] of a transport aeroplane.

The general intention is that effects of a catastrophic nature should virtually never occur in the fleet life of a particular type of aircraft. This would mean (for example) that in the case of a fleet of 100 aircraft of a particular type, each flying 3000 hours per annum, one or more of the various catastrophic effects might be expected to occur once in 30 years, which is close to the concept of 'virtually never',[33] a situation near to that 'never' we have already considered.

We have to bear in mind that there are some systems operating constantly and others operating in a certain flight phase only (the latter could make up as much as 80 per cent of the total: a landing gear system, for example). Hence, a probability failure per flight hours of such systems can be established by dividing the probability by the average flight duration estimated for the particular type of aircraft.

9.2.6.7.1 Failure conditions

Failure conditions are defined as effects on the aircraft and its occupants, both direct and consequential, caused or contributed to by one or more failures, considering relevant adverse operational or environmental conditions. Failure conditions may be classified according to their severity as follows (AMJ 25.1309):

1. **Minor.** Failure conditions which would not significantly reduce aeroplane safety, and which involve crew actions that are well within their capability.
2. **Major.** Failure conditions which would reduce the capability of the aeroplane or the ability of the crew to cope with adverse operating conditions to the extent that there would be, for example, a significant reduction in safety margins or functional capabilities, a significant increase in crew workload or in conditions impairing crew efficiency, or discomfort to occupants, possibly including injuries.
3. **Hazardous.** Failure conditions which would reduce the capability of the aeroplane or the ability of the crew to cope with adverse operating conditions to the extent that there would be:
   (a) A large reduction in safety margins or functional capabilities
   (b) Physical distress or higher workload such that the flight crew cannot be relied upon to perform their tasks accurately or completely, or
   (c) Serious or fatal injury to a relatively small number of the occupants.
4. **Catastrophic.** Failure conditions that would prevent continued safe flight and landing.

An **Inverted relationship** between the **severity** of the failure conditions and the **probability** of occurrence is established.[34] Hence:

1	Minor failures	become	probable
2	Major failures	become	remote
3	Hazardous failures	become	extremely remote
4	Catastrophic failures	become	extremely improbable

Each of the above probabilities has a maximum value assigned, which depends on the type of aircraft considered – for example, for large aircraft, extremely improbable is 10^{-9}, as we have already seen; extremely remote is 10^{-7}; remote is 10^{-5}, etc.

Figures 9.2-3 and 9.2-4, extracted from Book 2 of CS-25, show the above criteria.

We can gain a better indication of the safety levels relating to the above figures through another example. A single aircraft might fly a total of 5×10^4 hours and a large fleet of 200 aircraft (same type) might then accumulate a fleet total of 10^7 hours. Thus:

1. A catastrophic failure condition (at worst 10^{-9}) would be unlikely to arise in the whole fleet's life.
2. A hazardous failure condition (at worst 10^{-7}) might arise once in the whole fleet's life.
3. A major failure condition (at worst 10^{-5}) might arise once in an aircraft's life and would arise several times in the whole fleet's life.
4. A minor failure could arise several times in the aircraft's life.

The safety assessment of equipment, systems, and installation is a very important (and fascinating) part of aircraft design. It is of paramount importance to start the assessment from the very beginning of the design. A late assessment could bring unpleasant surprises, leading to expensive design changes.

As mentioned before, the techniques of safety assessment are a specialist matter.

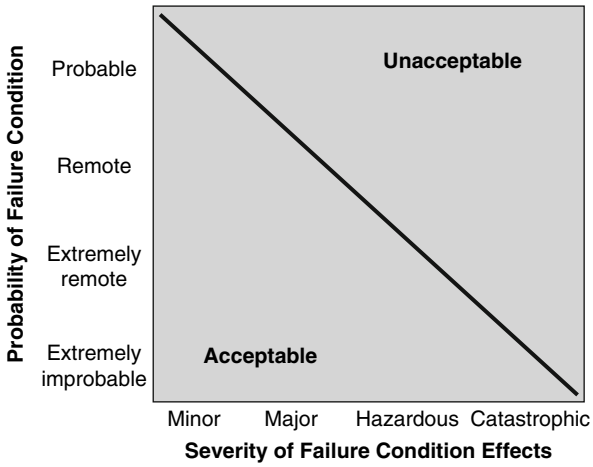


Figure 9.2-3 Classification of failure conditions.

9.2.6.8 Fatigue Strength

In order to remain within the limits of general information and guidance, leaving the rules, the advisory material, and the numerous treatises dealing with the subject as direct reference for further studies, we can see in summary how the airworthiness standards confront the structural fatigue that has caused so many air crashes, especially in the past.[35]

The airworthiness standards essentially consider two types of structure:

1. **Single load path** structures, where the applied loads are eventually distributed through a single member, the failure of which would result in the loss of the structural capability to carry the applied loads.[36]
2. **Multiple load path** structures, identified with redundant structures in which (with the failure of an individual element) the applied loads would be safely distributed to other load-carrying members.[37]

In the first case, the structure must result in **safe-life,** i.e. be able to sustain a certain number of events such as flights, landings or flight hours, during which there is a low probability that the strength will degrade below its design ultimate value due to fatigue cracking.

In the second case, the structure must be of **damage-tolerance** design, i.e. be able to retain its required residual strength for a period of unrepaired use after the failure or partial failure of a principal structural element due to fatigue, corrosion, accidental damage, and bird strikes.[38] Such a structure is defined as **fail-safe.**

For large aeroplanes and large rotorcraft, the relevant airworthiness standards require fail-safe structures,[39] unless this entails such complications that an effective damage-tolerant structure cannot be reached within the limitations of geometry, inspection, or good design practice. Under these circumstances, a design that complies with the safe-life fatigue evaluation requirements is used. A typical example of a structure that might not be conducive to damage-tolerance design is the landing gear and its attachments.

The divided opinions within the National Transportation Safety Board (NTSB)[40] at the end of 2002, as described by *Flight International,* give an idea of the importance of the choice between fail-safe and safe-life. The question was whether a component with exceptional reliability needs to be fail-safe.

A report on a January 2000 MD-83 crash led to the conclusion that the stabilizer was jammed fully, leading edge up, pitching the aircraft nose-down so strongly that elevator forces could not counteract it. The failure of the screw-jack assembly controlling the horizontal stabilizer pitch angle, due to inadequate lubrication, was also ascertained. Because the screw-jack mechanism is the sole component that sets and limits stabilizer pitch, its failure becomes catastrophic. Therefore, in some NTSB staff's opinion, the mechanism should have been redesigned according to a 'more rational' fail-safe criterion. Other NTSB staff argued that the component was reliable (the retrofit would have involved more than 800 civil aircraft), because:

1. The same basic mechanism has been used in all MD DC-9, MD-80, and -90 aircraft since 1965.
2. This kind of accident was the only one in more than 100 million flying hours.
3. The accident was caused by inadequate maintenance (the carrier was fined by the FAA).

Effect on aeroplane	No effect on operational capabilities or safety	Slight reduction in functional capabilities or safety margins	Significant reduction in functional capabilities or safety margins	Large reduction in functional capabilities or safety margins	Normally with hull loss
Effect on occupants excluding flight crew	Inconvenience	Physical discomfort	Physical distress, possibly including injuries	Serious or fatal injury to a small number of passengers or cabin crew	Multiple fatalities
Effect on flight crew	No effect on flight crew	Slight increase in workload	Physical discomfort or a significant increase in workload	Physical distress or excessive workload impairs ability to perform tasks	Fatalities or incapacitation
Allowable qualitative probability	No probability requirement	<...Probable...>	<...Remote...>	Extremely <.....> Remote	Extremely improbable
Allowable quantitative probability: Average probability per flight hour on the order of:	No probability requirement	$<10^{-3}$ <.....> Note 1	$<10^{-5}$ <.....>	$<10^{-7}$ <.....>	$<10^{-9}$
Classification of failure conditions	No safety effect	<...Minor...>	<...Major...>	<...Hazardous...>	Catastrophic

Note 1: A numerical probability range is provided here as a reference only. The applicant is not required to perform a quantitative analysis, nor substantiate by such an analysis, that this numerical criteria has been met for minor failure conditions. Current Transport category aeroplane products are regarded as meeting this standard simply by using current commonly-accepted industry practice.

Figure 9.2-4 Relationship between probability and severity of failure conditions.

In the end, common sense prevailed and the issue was resolved without modifications to the mechanism. The FAA issued an Airworthiness Directive[41] (AD 2000-15-15) providing inspection, check, and test 'to prevent loss of pitch trim capability due to excessive wear of the jackscrew assembly of the horizontal stabilizer, which could result in reduced controllability of the airplane…'.

The AD prescribes in particular the replacement of the jackscrew assembly with a new or serviceable assembly in case of metallic parts deterioration, corrosion, pitting, or distress.

The section 'Airworthiness Limitations' of the 'Instructions for Continued Airworthiness' must contain the inspections, replacement of parts, and other procedures necessary to prevent accidents caused by structural failures.

For JAR/FAR/CS-23 aeroplanes, it is possible to choose between the two philosophies safe-flight/fail-safe. Exceptions are made for composite airframe structures that must be designed according to fail-safe criteria, excluding the already mentioned cases of impracticability.

The previous remarks about continued airworthiness are also applicable to these aircraft.

For JAR/FAR/CS-27 rotorcraft that are prone, like all rotorcraft, to particular fatigue problems, mixed criteria are generally followed, based mainly on time change items (parts to be replaced according to prearranged and approved schedules). Everything must obviously be clear in the 'Instructions for Continued Airworthiness'.

For JAR/CS-VLA aeroplanes and JAR/CS-22 sailplanes, the airworthiness standards contain very little information on fatigue,[42] similar to the old airworthiness standards for light aeroplanes. As matter of fact, the low average utilization of these aircraft (100–200 flight hours/year) was not worrisome. However, after several years, and the intensive use of the machines in certain kinds of operation (e.g. school, aerial work, air taxi), fatigue problems became one of the causes of accidents, so much so that step by step the fatigue airworthiness standards have been amended in JAR/FAR 23. Similar problems also occurred for sailplanes (perhaps less for VLAs as they are 'younger'), so that no sailplane or VLA is now certified without fatigue assessment made by the manufacturers and authorities. For instance, LBA, in Germany, a long time ago issued standards for fatigue assessment of sailplane composite airframe structures.

In the case of loads and loading spectra, the assumptions made for fatigue assessment are:

1. **For rotorcraft,** it is explicitly required that for each portion of the flight structure of which failure could be catastrophic, loads or stresses must be verified (or determined) by in-flight measurements; the same must be done for the loading spectra being considered. Then the identification of these 'critical' structural items is of paramount importance.
2. **For large aeroplanes,** the principal loads that should be considered in establishing a loading spectrum are flight loads (gust and maneuver), ground load, and pressurization loads. The loading spectra are based on measured statistical data derived from government and industry load history studies and, where no sufficient data are available, on a conservative estimate of the anticipated use of the aeroplane. In assessing the possibility of serious fatigue failures, the design is examined to determine probable points of failure in service. In this examination, consideration is given, as necessary, to the results of stress analysis, static and fatigue tests, strain gauge surveys, tests of similar structural configurations, and service experience.
3. **For JAR/FAR/CS-23 aeroplanes,** criteria similar to the above are adopted.
4. **For sailplanes and VLAs,** apart from the general design recommendation to avoid stress concentration areas as far as possible, fatigue tests are performed, but only if they are essential, for economic reasons. If possible, reference is made to data resulting from fatigue tests performed on similar structures and service experience. Another way to avoid fatigue tests is the design of critical structures with stress levels under the fatigue limit of the material involved. Obviously, this must be properly demonstrated by static tests and strain gauge surveys.

If fatigue tests are necessary, the technical literature provides typical load spectra and programs for the repeated application of loads.

Apart from the consideration made for sailplanes and VLAs, the fatigue life assessment is performed through analysis, and fatigue tests on structures or single parts, according to criteria that are detailed in airworthiness standards and ACJs/ACs/AMC&GM. All analysis and test schedules are normally agreed with the authorities.

Fatigue test programs for large aeroplanes can last some years; hence, it is not generally possible to complete them before the aeroplanes' type certification. It is therefore required that at least one year of safe operations be demonstrated when the type certificate is issued. Subsequently, in order to maintain the validity of the type certificate, the fatigue life substantiation must always exceed the number of cycles/flight hours reached by the 'oldest' aeroplane (lead aeroplane).

9.2.7 JAR/FAR 21

JAR/FAR 21 contains the 'Certification Procedures for Aircraft and Related Products and Parts' for JAA and FAA certification respectively. JAR 21 deals with:

1. Procedural requirements for the issue of type certificates and changes to those certificates, the issue of standard certificates of airworthiness, and the issue of Export Airworthiness Approvals.
2. Procedural requirements for the approval of certain parts and appliances.[43]
3. Procedural requirements for the approval of organizations related to the subject of the previous points.
4. Rules governing the holders of any certificate or approval specified in the previous points.

In a similar way, FAR 21 deals with:

1. Procedural requirements for the issue of type certificates and changes to those certificates, the issue of production certificates, the issue of airworthiness certificates, and the issue of Export Airworthiness Approvals.
2. Rules governing the holders of any certificate specified in paragraph (a)(1) of paragraph 21.1 (Applicability).
3. Procedural requirements for the approval of certain materials, parts, processes, and appliances.

JAR/FAR 21 are therefore the rules upstream of the airworthiness standards, dictating, so to speak, the 'rules of the game'. The relationship between authorities and enterprises for certification of design and production of aeronautical materials are established.

9.2.8 EASA Part 21

As we previously mentioned, this document replaces JAR 21, which remains the core of the same document. The changes to the JAR document reflect the new legal status of the EASA towards the national authorities and a full revision of the document in light of the JAA certification experience.

Without attempting a full comparison between the two documents, it is worth making the following observations.

9.2.8.1 Type Certificates

Subpart H of Part 21 (Article 21A.184) includes the 'Restricted type certificates' missing in JAR 21.

9.2.8.2 Airworthiness Certificates

Subpart H of Part 21 (Article 21A.173) classifies the airworthiness certificates as follows:

1. A certificate of airworthiness shall be issued to aircraft which conform to a type certificate that has been issued in accordance with this Part.
2. Restricted certificates of airworthiness shall be issued to aircraft:
   (a) that conform to a Restricted type certificate that has been issued in accordance with this Part, or
   (b) that have been shown to the Agency to comply with specific certification specifications ensuring adequate safety
3. Permits to fly shall be issued to aircraft that do not meet, or have not been shown to meet, applicable certification specifications but are capable of safe flight under defined conditions.

The certificates in 1 are equivalent to the Standard certificates of airworthiness of JAR 2.

The certificates in 2 are consequent to the Restricted type certificates and do not exist in JAR 21.

The certificates in 3 have the characteristics of the Special certificates of airworthiness currently issued by national authorities and are not included in JAR 21.

The Special certificates of airworthiness (present, for example, in FAR 21) constitute a very complex and much discussed matter for which there is no harmonization amongst EU Member States. It is possible that in the future the EASA will deal with this issue more meticulously.

9.2.8.3 Environmental Protection

Part 21 of the type certification includes the designation of applicable environmental protection requirements and certification specifications, missing in JAR 21.

9.2.9 Structure of Aircraft Airworthiness Standards

If we look at the airworthiness standards for aircraft certification (JAR/CS-22, JAR/CS-VLA, JAR/CS-VLR, JAR/FAR/CS-23, -25, -27, and -29), we note a common structure that entails a certain unity and uniformity. Apart from the forewords, the lists of pages, etc., we find 'sub-parts' and 'appendices'. As mentioned previously, the JARs/CS also contain advisory material. Each subpart contains paragraphs under a title (e.g. 'Ground Loads', 'Control Systems', etc.), and it is of interest to see that, in all the above standards, the same topics are generally dealt with in paragraphs bearing the same number (e.g. 'Weight limits' paragraph XX.25; 'Materials and workmanship' paragraph XX.603; etc.). This makes it easier to pass from one standard to another, and to define comparisons when that is needed.

Some details of this structure are as follows:

1. **Subpart A: General.** This Subpart provides information about the types and categories of aircraft to which the standard is applicable.
2. **Subpart B: Flight.** This Subpart deals with the flight tests to be carried out to show compliance with the requirements for performance, controllability and maneuverability, stability, etc. It is worth stating that this Subpart does not exclusively cover certification flight tests; other Subparts contain some requirements that must be complied with through flight tests.

3. **Subpart C: Structure.** This Subpart contains the requirements for flight and ground load assessment, and for structural design of airframes, control systems, landing gears, and other components. Crashworthiness and fatigue requirement parameters are also provided.
4. **Subpart D: Design and Construction.** This Subpart deals with the design technique, materials, safety factors, control system and landing gear design, structural tests to be carried out, cockpit and passenger cabin design, fire protection and flutter requirements, etc.
5. **Subpart E: Power Plant.** This Subpart contains the requirements for power plant installations and related systems (like fuel, oil, exhaust systems, etc.). Power plant controls, accessories, and fire protection are also considered.
6. **Subpart G: Operating Limitations and Information.** This Subpart provides requirements for all the information that must be available to the pilot and other personnel for correct aircraft operations: from marking and placards, to the flight manual content.
7. **Appendices.** These documents are of various natures; they can provide simplified design load criteria, test procedures for assessment of material flammability, instructions for continued airworthiness, and other information.

9.2.10 Aircraft Airworthiness Standard Applicability

As mentioned above, Subpart A of aircraft airworthiness standards defines types and categories of specific aircraft. We will consider this in more detail.

9.2.10.1 JAR/CS-22. Sailplanes and Powered Sailplanes

1. Sailplanes with a maximum weight not exceeding 750 kg.
2. Single-engine (spark or compression ignition) powered sailplanes with a design value W/b^2 (weight to span^2) not greater than 3 (W in kg, b in m), and maximum weight not exceeding 850 kg.

The maximum number of occupants for both sailplanes and powered sailplanes must not exceed two.

The term 'powered sailplane' includes those powered sailplanes which may be incapable of complying with the minimum rate of climb required by paragraph 22.65 and maximum takeoff distance required by paragraph 22.51, and which must consequently be prohibited from taking off solely by means of their own power (so they are launched like sailplanes). These powered sailplanes are referred to as 'self-sustaining powered sailplanes' and additional requirements of Appendix I are applicable to them.

JAR/CS-22 contains a Subpart H and a Subpart J with standards for engines and propellers to be installed on powered sailplanes (engines and propellers certificated according to JAR-E and JAR-P as relevant can, of course, be installed).

Note. In JAR/CS-22 the term 'sailplane' is related both to sailplanes and powered sailplanes. In JAR 22 the requirements applying only to powered sailplanes are annotated with the letter P in the margin.

Sailplane categories are as follows:

1. **Utility.** Sailplanes intended for normal soaring flight and some aerobatic maneuvers (listed in the requirements) if demonstrated during type certification.
2. **Aerobatic.** Sailplanes intended for aerobatic maneuvers additional to those permitted in the Utility category. The permitted aerobatic maneuvers must be established during type certification.

Notes. Powered sailplanes are considered as 'sailplanes with an auxiliary engine'. That is why the installation of power plants complying with less severe requirements is allowed. Furthermore, unlike aeroplanes, powered sailplanes are allowed to fly with the engine off (and retracted if that is possible), like sailplanes. Flight tests must be carried out both with power on and power off – and power plant retracted when this is possible.

In order to explain that strange-looking parameter W/b^2, we can say that, when the first standards for powered sailplanes began to appear, some manufacturers envisaged the possibility of producing aeroplanes 'disguised' as powered sailplanes. The intent was to comply with requirements that were less severe - at that time. It is clear that for an 850-kg powered sailplane, the formula W/b^2 not greater than 3 leads to a minimum wing span of 16.8 m, which gives the aircraft an appearance more like a sailplane than an aeroplane.

9.2.10.2 JAR/CS-VLA. Very Light Aeroplanes

This includes aeroplanes with a single engine (spark or compression ignition) having no more than two seats, with a maximum certificated take-off weight of not more than 750 kg and a stalling speed in the landing configuration of not more than 45 knots (CAS). The approval to be given must be for day-VFR[44] only.

Notes. JAR-VLA was issued as a simplification of FAR 23 (JAR 23 did not yet exist). Unfortunately, while other requirements have been continuously updated by the relevant study groups, JAR-VLA has remained practically unchanged since publication; this is why some inaccuracies are still present as well as some old-fashioned concepts (e.g. in crashworthiness matters). An increase in maximum take-off weight would also be advisable, because 750 kg is somewhat insufficient, especially in the case of composite airframes (avoiding expensive carbon fibers).

In the United States, where the requirements have been adopted (see note 6 of this chapter), the VLA certification for night flight and IFR[45] is possible, in compliance with the additional requirements of AC 23–11. In Europe, this possibility has not yet been approved.

9.2.10.3 JAR/FAR/CS-23. Normal, Utility, Aerobatic, and Commuter Category Aeroplanes

1. Aeroplanes in the Normal, Utility, and Aerobatic categories that have a seating configuration, excluding the pilot seat(s), of nine or fewer and a maximum certificated take-off weight of 5670 kg (12 500 lb) or less.
2. Propeller-driven, twin-engine aeroplanes in the Commuter category that have a seating configuration, excluding the pilot seat(s), of 19 or fewer and a maximum certificated take-off weight of 8618 kg (19 000 lb) or less.

Aeroplane categories are as follows:

1. **Normal.** The Normal category is limited to non-aerobatic operations. Non-aerobatic operations include stalls (except whip stalls) and some simple maneuvers (listed in the requirements) in which the angle of bank is not more than 60°.
2. **Utility.** The Utility category is limited to the operations of the Normal category, spins (if approved for the particular type of aeroplane), and some aerobatic maneuvers (listed in the requirements) in which the angle of bank is between 60° and 90°.
3. **Acrobatic.** The Acrobatic category has no restrictions other than those shown to be necessary as a result of required flight tests.
4. **Commuter.** The Commuter category is limited to any maneuver incident to normal flying, stalls (except whip stalls), and steep turns in which the angle of bank is 60° or less.

9.2.10.4 JAR/CS-25. Large Aeroplanes/FAR 25. Transport Category Airplanes

These comprise:

- Large turbine-powered aeroplanes (JAA/EASA)
- Transport category airplanes (FAA).

Note. There are no limitations as regards weight, number of engines, and number of occupants.

Maximum weights corresponding to the airplane's operating conditions (such as ramp, ground or water taxi, take-off, en route, and landing), environmental conditions (such as altitude and temperature), and loading conditions (such as zero fuel weight, center of gravity position, and weight distribution) are established so that they are not more than:

1. The highest weight selected by the applicant for the particular conditions.
2. The highest weight at which compliance with each applicable structural loading and flight requirement is shown.
3. The highest weight at which compliance is shown with the certification requirements of FAR 36.

JAR 25 takes into account turbine-powered aeroplanes only. Actually, large aeroplanes powered by reciprocating engines have not been designed for many years. FAR 25, issued before JAR 25 as a derivation of older regulations, does not have this limitation. In effect, transport aeroplanes powered by reciprocating engines are still flying in some parts of the world, with valid type certificates.

9.2.10.5 JAR/CS-27. Small Rotorcraft/FAR 27. Normal Category Rotorcraft

Rotorcraft with a maximum weight of 3175 kg (7000 lb) or less and nine or less passenger seats.

Multi-engine rotorcraft may be type certificated as **Category A** provided the requirements referenced in Appendix C are met.

Note. For Category A definition, see the Notes on JAR/FAR/CS-29.

9.2.10.6 JAR/CS-29. Large Rotorcraft/FAR 29. Transport Category Rotorcraft

Rotorcraft categories are as follows:

1. Rotorcraft must be certificated in accordance with either the **Category A** or **Category B** requirements of JAR/FAR/CS-29. A multi-engine rotorcraft may be certificated as both Category A and Category B, with appropriate and different operating limitations for each category.
2. Rotorcraft with a maximum weight greater than 9072 kg (20 000 lb) and 10 or more passenger seats must be type certificated as Category A rotorcraft.
3. Rotorcraft with a maximum weight greater than 9072 kg (20 000 lb) and nine or less passenger seats may be type certificated as Category B rotorcraft provided the Category A requirements of Subparts C, D, E, and F of JAR/FAR/CS-29 are met.
4. Rotorcraft with a maximum weight of 9072 kg (20 000 lb) or less but with 10 or more passenger seats may be certificated as Category B rotorcraft provided the Category A requirements of JAR/FAR/CS-29.67(a)(2), 29.87, 29.1517, and of Subparts C, D, E, and F of JAR/FAR/CS-29 are met.
5. Rotorcraft with a maximum weight of 9072 kg (20 000 lb) or less and nine or less passenger seats may be type certificated as Category B rotorcraft.

Notes. Category A means a multi-engine rotorcraft designed with engine and system isolation features specified in JAR/CS-27/JAR/CS-29 and capable of operations using take-off and landing data scheduled under a critical engine failure concept, which assures adequate designated surface area and adequate performance capabilities for continued safe flight or safe rejected take-off in the event of an engine failure.

Category B means a single-engine or multi-engine rotorcraft that does not meet Category A standards. Category B rotorcraft have no guaranteed capability to continue safe flight in the event of an engine failure, and unscheduled landing is assumed.

9.2.10.7 FAR 31. Manned Free Balloons

1. Captive gas balloons deriving lift from a captive lighter-than-air gas.
2. Hot-air balloons deriving lift from heated air.

Notes. There are no limitations on weight and number of occupants. The certification maximum weight is the highest weight at which compliance with each applicable requirement of this part is shown.

9.2.10.8 JAR/CS-VLR. Very Light Rotorcraft

Very light rotorcraft with a single engine (spark or compression ignition) having no more than two seats, with a maximum certificated take-off weight of not more than 600 kg. The approval has to be for day-VFR only.

Note. The Italian RAI-ENAC issued a standard for this type of rotorcraft in the 1990s (it was approved on 22 April 1996), in order to allow the certification of ultralight helicopters which were limited, by the relevant low in Italy, to a maximum weight of 450 kg. Such a maximum weight was considered inadequate and, furthermore, these ultralights had no certification standards.

RAI-VLR (this was the title of the standard) was issued as a simplification of JAR 27, following a path resembling that adopted for the issue of JAR-VLA (simplification of FAR 23). RAI-VLR is now a national standard that could lead to a type certification and the issue of Standard certificates of airworthiness. Nevertheless, because the RAI was bound to the Cyprus agreement, an Elementary Aircraft category was created in Italy, for which a special certification is required.

Then, the standard was 'offered' to the JAA, which set up a study group for the evaluation and eventual revision of this document, in order to finally issue a JAR-VLR.

The JAR-VLR was issued in September 2003 and than adopted by the EASA as CS-VLR.

9.2.11 Airworthiness Standards for Unmanned Aircraft

These aircraft are internationally known as **Unmanned Aerial Vehicles (UAVs)** or **Uninhabited Aerial Vehicles (UAVs),** and also as **Remotely Piloted Vehicles (RPVs)** and **Remotely Operated Aircraft (ROAs).**

The most common denomination for this type of aircraft is **UAVs,** and we will use this abbreviation.[46]

UAVs have been mainly used by the world's armed forces for wartime operations for more than 60 years, for battlefield observations in the past, and more recently as a wartime means of attack. We can therefore argue that UAVs have already reached a technical maturity and this is going to evolve as for any other kind of aircraft. However,

up to the present, UAV missions have been limited on battlefields, to restricted flying areas, outside the zones open to civil aircraft operations.

Today, when the great potential of this type of machine has been recognized, the global industry has requested the opportunity of also using them commercially in civil airspace. This possibility is also of interest to the defense industry, because they can achieve better operational flexibility – for example, in the case of transfer flights.

We have mentioned the potential of UAVs for civil applications. We now consider what kind of applications these might be.

As a first example, thousands of rotary wing UAVs are used for agricultural purposes in Japan (crop spraying – pesticides and fertilizers). These machines, all built in Japan, carry a useful load of 25-150 kg.

Some classifications have been drafted and the list below includes just a few examples taken from the multitude of possible UAV uses:

- Forestry services – fire control and other kinds of surveillance.
- National weather services – atmospheric sampling, meteorology, etc.
- Agriculture and wildlife – agricultural monitoring, river and estuary surveys, illegal waste disposal surveys, crop dusting, mapping, fishing law enforcement, etc.
- Electricity authorities – monitoring nuclear facilities, power line verification, etc.
- Postal services – urgent package delivery in remote areas.
- Coastguards – surveillance for counter narcotics, illegal alien intrusion detection, illegal fishing control, search and rescue missions, etc.
- Civil aviation – noise measurement for aircraft certification purposes.
- Telecommunications – as telecom relays (replacing satellites), local TV coverage, etc.

It is clear from these examples that, in many cases, the scope of UAVs is to carry out the 'dirty jobs', i.e. dangerous tasks, or tasks too long or too tedious for a crew.

Can UAVs be legally defined as 'aircraft'? An answer has been given by the ICAO EURNAT Office: UAVs are aircraft as defined by Annex 2 of the ICAO. Furthermore, the Chicago Convention in Article 8 declares that:

> *No aircraft capable of being flown without a pilot shall be flown without a pilot over the territory of a contracting State without the special authorization by that State and in accordance with the terms of that authorization. Each contracting State undertakes to ensure that the flights of such aircraft without pilot in regions open to civil aircraft shall be so controlled as to obviate danger to civil aircraft.*

Put simply, the flight of UAVs is allowed in regions open to civil aircraft only if authorized by the concerned states (something existing also for other aircraft) and if they do not cause any danger to other aircraft.

Therefore, the real problem is now to develop concepts for the safe integration of UAVs in general air traffic. It is then necessary to develop rules harmonized with the existing rules for air traffic control.

The issues concerning the above rules can be easily classified, as for 'manned' aircraft, into three basic segments:

1. Personnel licenses
2. Air traffic management
3. Airworthiness.

Hence we return to the main safety factors: man, the environment, and the machine.

Studies and conferences on the above subjects have been taking place for many years. In Europe there are some institutes and associations dealing with these issues. One is the European UVS International (formerly EURO UVS), equivalent to the AUVSI (Association for Unmanned Vehicle System International) in the United States. A great contribution to this discussion has also been made by EUROCONTROL, particularly concerning air traffic management matters. Other initiatives have been taken all over the world. In this book, we will limit our discussion to the airworthiness of UAVs.

9.2.11.1 Airworthiness Standards

We should not be misled by the title: at the time of writing (July 2005), no official airworthiness standards for UAVs exist.

In the 1990s, at the request of the national industry, the Italian RAI-ENAC issued a draft of a UAV airworthiness standard. This document was presented at the annual EURO UVS conference in June 1999, triggering great debate on the subject. It was indeed probably the first attempt to define some sort of airworthiness standards for civil UAVs.

Instead of attempting to invent everything from scratch, the JAR-VLA standard was chosen as a basic standard to be adapted to fixed-wing UAVs up to 750 kg.

It could be argued that, in order to transform a standard for 'manned' aircraft into a UAV standard, it would be sufficient to delete all requirements inherent to the occupants, such as the cockpit and the passenger cabin requirements. But it is not so simple because the airworthiness 'philosophies' considered for 'manned' aircraft would not be fully utilized in doing so.

It is therefore necessary to set up new philosophies specific for UAVs before trying to convert them into new standards.

The definition of 'airworthiness' prevalent in current literature is perfectly suitable to UAVs ('requirements' and 'allowable limits' should also exist for these machines), provided that we clarify the meaning of 'safe condition'. In other words, what constitutes 'safety' for UAVs? This is a topic that requires debate and validation.

If we consider the various airworthiness standards, we clearly understand that they are written having in mind the **occupant's protection.** The protection of people and properties on the ground is an added purpose of the safety obtained through compliance with the standards. For some categories of aircraft (aerobatic aeroplanes, sailplanes, and powered sailplanes) the abandonment of the aircraft in emergency cases is even considered. (The presence of a pilot on board could, in certain cases, avoid or limit the damage on the ground, but this is not guaranteed, as demonstrated by various accident reports.)

On the other hand, it is always very difficult to establish exactly what to do to avoid damage on the ground when considering air accidents: the most reasonable way to approach this problem is by trying to prevent the accident from occurring.

From the above considerations we can logically assume that we need to approach UAV standards, which by definition have no occupants, from a totally different perspective.

If we start from a general, but fundamental, safety principle of the protection of human beings, we can state that the UAV standards should aim to avoid any damage to people (and properties) in the UAVs range of action. This can only mean one thing: **to avoid in-flight collisions and uncontrolled ground impact.**

This can be achieved by applying the system safety assessment concepts that we have already mentioned in this chapter, and the standards for flight, structural strength, etc., which can be acquired from the current standards for 'manned' aircraft. This also leads to an additional parameter, **mission effectiveness,** which applies equally to UAVs and manned aircraft. This can be achieved by taking from the current standards, based on a century of experience, everything that might be applicable to UAVs.

In the case of UAV safety assessment, it is clear that the **severity of failure conditions** for UAVs will be very different compared to 'manned' aircraft.

For example, for 'manned' aircraft, a **catastrophic failure condition** is one that **would prevent continued safe flight and landing.** For a UAV this situation would be not at all catastrophic if the aircraft has a 'Flight Termination System' (FTS) capable (using parachutes, for instance) of bringing the machine back to the ground. An FTS failure could instead become catastrophic and there are other numerous examples supporting this argument.

We can infer that a new setting of standards taken from paragraph XX. 1309 has to be arranged for the determination of **severity failure conditions** and **probability of occurrence.** But in the case of JAR-VLA, for example, we have seen that paragraph 1309 has a small number of applicable requirements.[55] Therefore, we have to provide something different for a UAV standard based on JAR-VLA.

Furthermore, the installation of anti-collision systems, certainly not installed on VLAs, should be compulsory.

Another peculiarity of UAV standards should be the incorporation of requirements for the 'Air Vehicle Station' (AVS) – the ground guidance station – that have to be considered as an integral part of the flying material and should be consistent with it.

In conclusion, we can argue from the analysis made so far that, in order to produce UAV airworthiness standards, many difficulties must be overcome; these difficulties are not so much related to UAV technology, which already exists and is evolving, but specifically associated with the creation of the related airworthiness philosophies and their correct transfer into the standards to be issued for the different UAV categories (also to be defined).

9.2.11.2 The State of the Art

At the beginning of this review of airworthiness standards for UAVs, it was stated that: 'at the time of writing (July 2005), no official airworthiness standards for UAVs exist'.

When *Aeronavigabilità* was written at the end of 2002, it was emphasized that:

> *In any case we must be able to discuss such requirements only when these problems will be faced with determination by authorities like the FAA and JAA (and subsequently the EASA).*

So, why bring up this issue now?

The answer is that, even without having achieved common and approved rules, much work has been carried out in the last two years.

Peter van Blyenburgh, President of UVS International, produced a summary of the major achievements in this field in the publication 'UAVs – A Global Perspective 2004'. In particular, the creation in October 2002 of a Joint JAA/EUROCONTROL UAV Task Force, under the auspices of the JAA and EUROCONTROL, has led to progress in this area.

This Task Force had several unique characteristics:

1. It was a multi-national effort (France, Germany, Greece, Italy, the Netherlands, Sweden, UK, USA).
2. It consisted of all major stakeholders - the JAA, EUROCONTROL, NATO, European National Civil Aviation Authorities, the FAA, AECMA, and UVS International - representing the UAV industry.
3. It was a European initiative, in which the FAA and US industry representatives participated.
4. It was a one-year non-funded study, which had as its objective the concurrent formulation of concepts (outline and guiding principles) to drive the establishment of UAV-related airworthiness and certification rules.
5. This was the first international effort of its kind and the UAV Task Force's final report, which came out in May 2004, is certain to have an important international outcome.

Furthermore, a number of studies and initiatives have now been initiated by civil aviation authorities (in Australia, France, Italy, Japan, South Africa, Sweden, Switzerland, the UK, and the USA), military flight safety authorities (Belgium, Finland, France, Germany, the Netherlands, Singapore, South Africa, Sweden, UK, and USA), and in some cases UAVs can already, with certain restrictions and special waivers, operate in civil managed airspace (Australia, Belgium, Finland, France, Germany, Switzerland, UK, and USA).

From the EASA standpoint, in the same aforementioned publication, the Executive Director M. Goudou pointed out that an action to consider appropriate certification specifications for the UAV system was earmarked for action in the Agency's rulemaking planning in the future, with contributions from UVS International and other stakeholders.

As a result, substantial international movement is under way and it will able to solve the issues summarized here, also taking into consideration the vast experience available in the field of manned aircraft, which is certainly the starting point.

Notes

1. The OSTIV (Organisation Scientifique et Technique International du Vol à Voile) is an independent organization linked to the FAI (Fédération Aéronautique Internationale). The organization's aim is to encourage and internationally co-ordinate the science and techniques of sailplane flight and design.
2. Now CS-22.
3. Before the issue of JAR 22, this was adopted as a national requirement by some states.
4. The correct denomination of the FAA regulations should be FAR Part XX (Ex. Part 11). For the sake of practicality, and to clearly see the difference from JAA and EASA requirements, we use the denomination 'FAR XX (Ex. FAR 11)'.
5. The FAA adopted JAR 22 as an acceptable standard for the certification of sailplanes and powered sailplanes in the United States.
6. An equivalent FAA standard does not exist. The FAA adopted these requirements for the certification of very light aeroplanes in the USA. The acceptable criteria for the adoption of JAR-VLA are included in AC 21.17–2A. The FAA also issued adjunctive rules in AC 23–11 to authorize IFR and night flight of such aeroplanes.
7. The JAA has not issued requirements for free balloons.
8. The FAA rules for APU certification are contained in the Technical Standard Order (TSO) C 77 B.
9. Exhaust emissions refer to substances emitted into the atmosphere from the exhaust nozzle of an aircraft engine. Fuel venting emissions refer to raw fuel, exclusive of hydrocarbons in the exhaust emissions, discharged from aircraft gas turbine engines during all normal ground and flight operations.
10. The FAA's Airworthiness Directives are legally enforceable rules that apply to the following products: aircraft, aircraft engines, propellers, and appliances.
11. While this part is essentially operative, airworthiness is recalled for equipment, instrument, and certification requirements. The same applies to other operative parts like FAR 121, 125, 129, 133, 135, JAR-OPS, and JAR-AWO.
12. According to JAR and FAR, products are aircraft, aircraft engines, and propellers.
13. Obviously, the JAA and FAA operational standards are related to the aircraft registered in the country having those standards as legal operational rules.
14. Many states adopt the FAA regulations as a basis for their national regulations.
15. The term 'rotorcraft' is not limited to helicopters, but also includes gyroplanes (even if they are less common).
16. JAR 11 defines ACJ as 'an accompanying text, containing explanations, interpretations or acceptable means of compliance, in order to clarify and to provide guidance for the application of requirements'.
17. This means that the designer (or better, the 'applicant', as it is normally defined) can choose other means of compliance, but in this case has to convince the authority about the validity of the choice.
18. See also the 'Advisory material' section in this chapter.
19. See the 'JAR-VLR' section in this chapter. At the end of 2002 it was still in a status of NPA (Notice of Proposed Amendment).
20. It used to be said that the limit trend of the airworthiness standards was to make aircraft certification impossible!
21. The same also applies for boxers' categories or for tax brackets.
22. We will see later the meaning of 'never' in flight safety assessment.
23. As a basic concept, an aircraft must be manageable in all foreseen conditions, by a crew of average skill (for that class of aircraft), and not necessarily by over-skilled people.

24. We can understand the lack of single-jet engine aeroplanes in the civil market, because they have performance normally incompatible with the above stalling speed limitation. We will mention this issue again with regard to 'crashworthiness'.
25. As we will see, this airworthiness standard concerns aeroplanes up to 750 kg maximum weight.
26. It may seem trivial, but how many people know that foam rubber cushions can be dangerous? They could indeed return most of the absorbed impact energy.
27. Independently of the fire emergency, the abandonment of the aircraft is considered in limited categories of civil aircraft like sailplanes/powered sailplanes and aerobatic aeroplanes. This is necessary for the hazard of flight collisions, especially during thermal flights of sailplanes, and for the hazard of the structure overloading and the critical situation that can occur during aerobatic operations. The applicable airworthiness standards provide suitable rules for this type of emergency.
28. JAR and EASA standards contain equivalent procedures.
29. JAR 22 does not contain this paragraph; JAR-VLA provides general indications only in order to minimize hazards in case of failure. This is consistent with the (generally) simple systems of the relevant aircraft.
30. See the 'Structure of aircraft airworthiness standards' section in this chapter.
31. A 'multi-fatality' accident, normally leading to the loss of the aircraft.
32. Accident analysis for other types of aircraft leads to different values. For example, for JAR 23 single engines it becomes 10^{-6}.
33. Total hours per year 3×10^5. In 30 years, 9×10^6, near to 10^7, which could imply a catastrophic accident (considering all aircraft systems).
34. Where the effects are less hazardous, they are permitted to occur more frequently.
35. Example: the 'Comet' crashes in the 1950s, due to fatigue caused by fuselage pressurization.
36. Example: a wing-fuselage attachment made by a single structural element. Such an arrangement is common in light aircraft.
37. Example: a multiplex wing-fuselage attachment, made by several structural elements. Such an arrangement is classical in large aeroplanes.
38. The bird strike is considered for large aeroplanes only, but it could be the object of special conditions for other types of aircraft.
39. Rotorcraft structures include rotors, rotor drive systems between the engines, and rotor hubs, controls, fuselage, fixed and movable control surfaces, engine and transmission mountings, landing gear, and their related primary attachment.
40. NTSB: the US organization dealing with aircraft accidents and issuing recommendations based on the investigation performed.
41. Airworthiness Directives: documents issued by the authority making particular mandatory actions (changes, inspections, etc.).
42. JAR-VLA offers some simplified criteria, but these must be considered carefully.
43. Appliance means any instrument, mechanism, equipment, part, apparatus, appurtenance or accessory, including communications equipment, that is used or intended to be used in operating or controlling an aircraft in flight, is installed in or attached to aircraft, and is not part of an airframe, engine, or propeller.
    JAR 21 normally uses 'parts and appliances' together, to include also the 'parts' of airframes, engines, and propellers.
44. VFR: Visual Flight Rules.
45. IFR: Instrumental Flight Rules.
46. ROA (Remotely Operated Aircraft) is preferred by the FAA and NASA. From a technical point of view it seems to be the most appropriate.
47. JAR-VLA deals with a simple aeroplane and the safety of two occupants. A UAV of the same weight is a sophisticated machine if we look at the systems, and we have also to consider the lives of the three or four hundred people aboard a large aircraft that could be brought down.

Index

B

C

D

G

H

I

M

Q

R

S

T

This page intentionally left blank

Information on Source Books

Mike Tooley and David Wyatt
Aircraft Communications and Navigation Systems
9780750681377
08/2007
£29.99/$59.95
The complete text for aerospace engineering and maintenance students and professionals

- Delivers the essential principles and knowledge base required by Airframe and Propulsion (A&P) Mechanics for Modules 11 and 13 of the EASA Part-66 syllabus and BTEC National awards in aerospace engineering
- Supports Mechanics, Technicians and Engineers studying for a Part-66 qualification
- Complete with Self-test questions, exercises and multiple choice questions to enhance learning for both independent and tutor-assisted study

Butterworth Heinemann

Lloyd Jenkinson and Jim Marchman
Aircraft Design Projects
9780750657723
2003
£38.99/$58.95
Draws together the various theoretical elements of aircraft design – structures, aerodynamics, propulsion, control and others

- Demonstrates how basic aircraft design processes can be successfully applied in reality
- Case studies allow student and instructor to examine particular design challenges
- Covers commercial and successful student design projects, and includes over 200 high quality illustrations

Butterworth Heinemann

Mike Tooley
Aircraft Digital Electronic and Computer Systems
9780750681384
11/2006
£29.99/$57.95
The comprehensive guide to aerospace engineering and maintenance. For both students and professionals

- Essential principles and knowledge base for Airframe and Propulsion (A&P) Mechanics
- With questions and exercises to enhance learning for both independent and tutor-assisted study
- Accompanied by online image bank for use in lecture presentations and teaching materials

Butterworth Heinemann

T.H.G. Megson
Aircraft Structures for Engineering Students
9780750667395
03/2007
£39.99/$79.95
The complete aircraft structures text: from mechanical fundamentals to analysis

- The leading Aircraft Structures text covering a complete course from the basics to analyses
- Features enhanced pedagogy with additional case studies, worked examples and exercises
- Online Solutions Manual available for adopting teachers

Butterworth Heinemann

Filippo De Florio
Airworthiness
9780750669481
2006
£31.99/$65.95
The only airworthiness guide available: a real contribution to understanding flight safety

- The only airworthiness guide available: a real contribution to understanding flight safety
- Helps anyone involved in the manufacture, flying and maintenance of aircraft to understand this complex yet essential topic
- Covers European and US requirements

Butterworth Heinemann

Michael V. Cook
Flight Dynamics Principles
9780750669276
08/2007
£39.99/$79.95
The study of flight dynamics is required of all aerospace and aeronautical engineers. This book is an all-encompassing reference for anyone in the field

- Extensive coverage of this broad topic – ideal for study and professional use
- Covers key contemporary topics including all aspects of optimisation, emissions, regulation
- End of chapter exercises and Instructor's Manual. Includes accompanying modelling code for performance model generation and optimisation solution

Butterworth Heinemann

Antonio Filippone
Flight Performance of Fixed Rotary Aircraft
9780750668170
2006
£41.99/$92.95
The most comprehensive coverage of fixed and rotary wing aircraft on the market

- Unique coverage of flight performance, including optimisation, emissions control and regulation
- Ideal for students, aeronautical engineering capstone projects and for professional reference in the aero industry, civil and military sectors
- With end of chapter exercises and Instructor's Manual. Includes accompanying website with downloadable flight performance modelling code

Butterworth Heinemann

Howard Curtis
Orbital Mechanics for Engineering Students
9780750661690
2004
£43.99/$91.95
From classical mechanics to spaceflight: this is the only OM book students need

- A complete, stand-alone text for this core aerospace engineering subject
- Richly-detailed, clearly and logically developed to meet the needs of students
- Highly illustrated and fully supported with downloadable MATLAB algorithms; with fully worked examples, Q&A material, homework exercises and an Instructor's Manual

Butterworth Heinemann