



Brigham Young University
BYU ScholarsArchive

Books

2018-12

Flight Vehicle Design

Andrew Ning

Brigham Young University, aning@byu.edu

Follow this and additional works at: <https://scholarsarchive.byu.edu/books>

 Part of the Aerospace Engineering Commons

Recommended Citation

Ning, Andrew, "Flight Vehicle Design" (2018). *Books*. 26.
<https://scholarsarchive.byu.edu/books/26>

This Book is brought to you for free and open access by BYU ScholarsArchive. It has been accepted for inclusion in Books by an authorized administrator of BYU ScholarsArchive. For more information, please contact scholarsarchive@byu.edu, ellen_amatangelo@byu.edu.

Flight Vehicle Design

Andrew Ning
Brigham Young University



Figure: [NASA](#), public domain

Copyright

© 2018 Andrew Ning. All rights reserved.

Colophon

This document was typeset with [L^AT_EX](#).

Publication

First published electronically Dec 2018.

Version

Last compiled: December 19, 2019

Contents

Nomenclature	8
1 Atmosphere	16
1.1 Layers	16
1.2 Altitude	18
1.3 Hydrostatic Equation and Geopotential Altitude	20
1.4 International Standard Atmosphere	22
1.5 Curve Fit to Standard Atmosphere	25
1.6 Other Atmospheric Models	25
2 Fundamentals	27
2.1 Airfoil Geometry	27
2.2 Wing Geometry	29
2.3 Aerodynamic Forces and Moments	31
2.4 Nondimensional Parameters	33
2.5 Pressure Distributions	35
2.6 Lift Curves and Drag Polars	40
3 Drag	44
3.1 Breakdown	44
3.2 Parasitic Drag	45
3.2.1 Skin Friction Drag	45
3.2.2 Pressure Drag	49

3.2.3	Putting it Together	51
3.2.4	Strip Theory	53
3.3	Induced Drag	54
3.4	Tradeoffs in Parasitic and Induced Drag	62
3.5	Compressibility Drag	62
3.5.1	Transonic Wave Drag	65
3.5.2	Supersonic Wave Drag	69
3.5.3	High Mach or Blunt Bodies	72
4	Wing Design	74
4.1	Fundamental Parameters	74
4.2	Lift Distributions and Lift Coefficient Distributions	77
4.3	Finite Wing Lift Curve Slopes	81
5	Stability	84
5.1	Definitions and Fundamentals	84
5.2	Longitudinal Static Stability	90
5.3	Lateral Static Stability	96
5.4	Coordinated Turns and Adverse Yaw	99
5.5	Tail Types and Statistical Tail Sizing	100
5.6	Dynamic Stability	102
6	Propulsion	108
6.1	DC Electric Motor	108
6.2	Propellers	110
6.3	Motor-Prop Matching	113
6.4	Gas Turbine Engines	117
6.5	Propulsion System Placement	123
7	Performance	126
7.1	Range	126
7.1.1	Jet Powered	127
7.1.2	Electric Powered	130
7.2	Endurance	132
7.3	Optimal Lift Coefficients	132
7.4	Rate of Climb	137
7.5	Glide Ratio	139
7.6	Turning Radius	140
7.7	High Lift Devices	141

7.8 Takeoff	144
7.9 Landing	149
8 Structures	152
8.1 Typical Structural Elements	152
8.2 Wind Bending	154
8.3 Flight Envelope Preliminaries	160
8.4 Speed-Altitude Flight Envelope	163
8.5 V-n Diagrams	164
9 Rocket Fundamentals	169
9.1 Types	169
9.2 Thrust and Efficiency (Specific Impulse)	173
9.3 Tsiolkovsky Rocket Equation	175
9.4 Staging	178
9.5 Stability	180
10 Rocket Propulsion	182
10.1 Thermodynamics Review	182
10.1.1 Properties and Equations of State	182
10.1.2 Laws of Thermodynamics	184
10.1.3 Specific Heats	185
10.1.4 Isentropic Relationships	187
10.2 Compressible Flow Primer	188
10.2.1 Energy Equation	189
10.2.2 Area Velocity Relationship	189
10.3 Nozzle Sizing	194
10.3.1 Exit Velocity	194
10.3.2 Throat Size	195
10.3.3 Combustion Chamber Sizing	195
10.3.4 Exit Area Sizing	196
10.4 Lengths	198
11 Orbital Mechanics	203
11.1 Gravity	203
11.2 Orbits	208
11.3 Hohmann Transfer	214

A Supplementary Material	216
A.1 Skin Friction Coefficient for a Flat Plate with Transition	216
A.2 Supersonic Wave Drag Equations	220
A.3 Rocket Altitude	224
A.4 Rocket Nozzle Performance	225
A.4.1 Exit Velocity	227
A.4.2 Throat Size	228
A.4.3 Combustion Chamber Sizing	228
A.4.4 Exit Area Sizing	229

Preface

Aircraft design is a multidisciplinary topic with many applications and can't possibly be fully covered in one text. Aircraft configurations are changing rapidly, particularly with recent advances in electric power systems and automated controls. Thus, this text primarily focuses on fundamental physics, as opposed to handbook or statistical methods. The last three chapters provide an introduction to rocket design.

A few sections are more advanced, or involve lengthier derivations, and these are put in the end in an Appendix. These chapters are not necessary to understand the main text, but are provided as a reference for the interested reader. Words in italics denote important terminology.

Nomenclature

Aircraft Nomenclature

\emptyset_∞	freestream
\emptyset_r	root
\emptyset_t	tip
\emptyset_{ac}	aerodynamic center
\emptyset_{cg}	center of gravity
\emptyset_{ref}	reference quantity
\emptyset_{sl}	sea level
α	angle of attack
α_0	zero lift angle of attack
β	sideslip angle
\dot{m}	mass flow rate
η	efficiency
η	spanwise lift centroid location

Γ	circulation
γ	flight path angle
Λ	sweep
λ	taper ratio
μ	dynamic viscosity
μ	rolling coefficient of friction
Ω	rotation rate (rad/s)
Φ	bank angle
ϕ	dihedral
ρ	density
σ	stress
τ	shear stress
θ	twist
A	axial force
a	speed of sound
AR	aspect ratio
b	span
c	chord
C_D	drag coefficient
c_d	2D drag coefficient
C_{Dc}	compressibility drag coefficient
C_{Di}	induced drag coefficient
C_{Dp}	parasitic drag coefficient
C_f	skin friction coefficient

C_L	lift coefficient
c_l	2D lift coefficient
C_{Lmax}	maximum lift coefficient
c_{lmax}	2D maximum lift coefficient
C_m	pitching moment coefficient
c_m	2D pitching moment coefficient
C_P	power coefficient
C_p	pressure coefficient
C_Q	torque coefficient
C_T	thrust coefficient
C_Y	side force coefficient
c_{mac}	mean aerodynamic chord
C_n	yawing moment coefficient
C_{roll}	rolling moment coefficient
D	drag
D	propeller diameter
D'	drag per unit span
D_c	compressibility drag
D_i	induced drag
D_p	parasitic drag
e	Oswald efficiency factor
e_b	battery specific energy
e_{inv}	inviscid span efficiency
EAS	equivalent air speed

fr	fineness ratio
g	acceleration of gravity
h	altitude
h_f	fuel enthalpy
i	current
J	advance ratio
k	form factor
K_v	motor velocity constant
L	lift
L'	lift per unit span
M	Mach number
M	pitching moment
M'	pitching moment per unit span
M_b	bending moment
m_b	battery mass
m_f	fuel mass
M_{br}	root bending moment
M_{cc}	crest critical Mach number
M_{dd}	drag divergence Mach number
m_{TO}	total takeoff mass
N	normal force
n	load factor
n	rotation rate (rev/s)
P	power

p	roll rate
p	pressure
Q	torque
q	pitch rate
R	range
r	yaw rate
R/C	rate of climb
R_t	turning radius
Re	Reynolds number
S	wing area
S_{exp}	exposed area
S_{gross}	gross area
S_{ref}	reference area
S_{wet}	wetted area
sfc	specific fuel consumption
T	thrust
t	thickness
T	temperature
t/c	thickness-to-chord ratio
TAS	true air speed
V	velocity
v	voltage
V_C	design cruise speed
V_D	design dive speed

V_s stall speed W weight w downwash Y side force

Rocket Nomenclature

ρ_∞	freestream
ρ_b	burnout
ρ_c	after combustion
ρ_e	earth
ρ_e	exit
ρ_t	throat
ρ_{sl}	sea level
\dot{m}_f	propellant mass flow rate
ϵ	nozzle expansion ratio
γ	ratio of specific heats
μ	GM , gravitational constant
ρ	density
θ	flight angle
θ_0	orbit phase shift
A	area
a	semimajor axis of ellipse
C	effective exit velocity
C_p	specific heat at constant pressure

C_v	specific heat at constant volume
D	drag
e	orbit eccentricity
e	specific internal energy
G	universal gravitation constant
g	acceleration of gravity
h	altitude
h	angular momentum per unit mass
h	specific enthalpy
h_T	total enthalpy
I_{sp}	specific impulse
L^*	characteristic mixing length
M	Mach number
m	mass
m_p	propellant mass
m_s	structural mass
M_w	molecular weight
$m_{payload}$	payload mass
MR	mass ratio (initial/final)
p	pressure
q	heat added
r	radius
R_e	radius of the earth
R_u	universal gas constant

s specific entropy

T kinetic energy

T temperature

T thrust

t_b burn time

U potential energy

v specific volume

V_e exit velocity

CHAPTER 1

Atmosphere

Aircraft, rockets, and spacecraft operate in or pass through the earth's atmosphere. Thus, we need to understand how atmospheric properties vary, at least with altitude. Primarily we are interested in pressure, density, and temperature, as well as what is meant by altitude. For the lower portion of the atmosphere, where aircraft operate, we will examine a "standard atmosphere" that is widely used as a first model. We will also mention, though not discuss, a couple more detailed models for other use cases.

1.1 Layers

The earth's atmosphere is not really composed of layers, but it is convenient to characterize it that way. Figure 1.1 provides an overview of the lower layers of the atmosphere and a typical temperature profile. The troposphere extends from the surface of the earth to about 10–12 km in altitude (varies with weather). About 80% of the mass of the atmosphere is contained in this layer and almost all of the water vapor. It is the layer where almost all weather occurs.

The next layer is the stratosphere, which extends roughly to 50 km. Most of the ozone in the atmosphere exists in this layer. In this layer temperature actually increases with altitude because the ozone absorbs

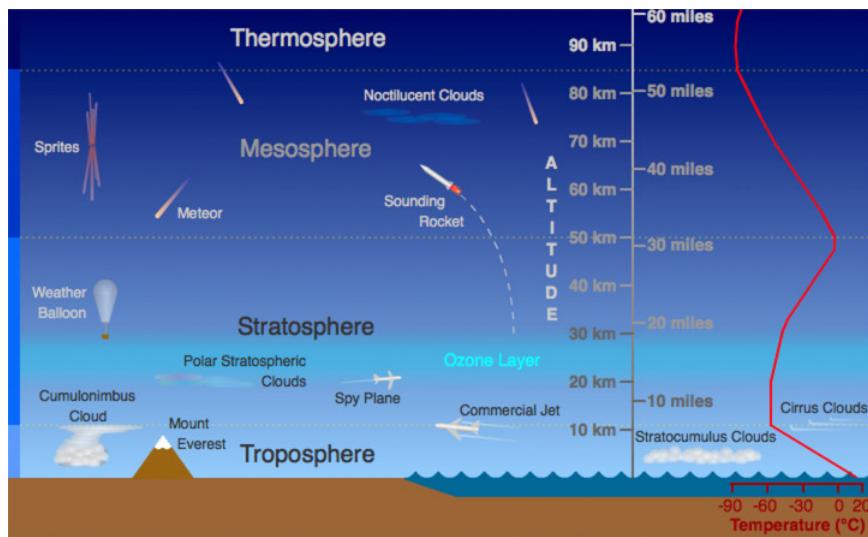


Figure 1.1: Depiction of layers in our atmosphere, and a typical temperature profile. Figure from Randy Russell, UCAR, public domain.

ultraviolet (UV) radiation from the sun (though it is still very cold!). This temperature inversion makes this layer very stable and thus there is little turbulence, clouds, or other weather effects. Most commercial aircraft fly near the border between the troposphere and the stratosphere. A typically cruising altitude for a commercial transport is somewhere between 10–13 km (approximately 33,000–42,000 ft).

Temperatures decrease again in the mesosphere, which extends to about 85 km. This is the coldest part of our atmosphere. There is not much activity by aircraft or rockets in this area, other than sounding rockets (a rocket designed to take measurements for research purposes).

The thermosphere again experiences a rise in temperature. In fact the temperature can be as high as 2000°C ! Why then would you still feel freezing cold? Recall that temperature is a measure of the average kinetic energy of the air molecules. But heat transfer depends also on the number of particles. In this layer the air is so thin, or in other words the molecules are spaced so far apart, that even though individual molecules have high energy there are so few that very little heat is transferred. The height of this layer varies considerably with solar energy and can range from 500 to 1,000 km in altitude. Some satellites exist in this layer.

The exosphere is the last layer and varying definitions have it extending to

somewhere between 10,000 km to 190,000 km (halfway to the moon). The air is so thin here that molecular collisions are rare. Most satellites operate in this layer.

1.2 Altitude

We have used the word altitude several times already, but there are actually multiple definitions of altitude that we need to be aware of. The first two are straightforward while the next two will need further explanation.

- *Geometric Altitude*: the geometric height from sea level.
- *Absolute Altitude*: the geometric height from the center of the earth.
- *Pressure Altitude*: corresponding altitude in the standard atmosphere that has the same pressure as the current pressure.
- *Geopotential Altitude*: an equivalent altitude assuming constant acceleration of gravity.

A device that measures altitude is called an *altimeter*. One way to measure height above ground is to use a radar altimeter. These are common in aircraft. They bounce radio waves off the ground and measure the return time in order to calculate height above ground level (AGL). Note that this is not the same as geometric altitude as the ground generally is not at sea level.

The most widely used type of altimeter is a pressure altimeter, also known as a barometric altimeter. Pressure altimeters are used in aircraft, and they are included in many smart phones and fitness tracking devices. A pressure altimeter takes a local reading of the pressure, then calculates the altitude based on the pressure (we will see how this works shortly). This calculation requires a model of the atmosphere, as it is not directly measuring height. Estimating altitude using pressure and an atmospheric model is what is meant by *pressure altitude* as it is not necessarily a true altitude.

Because of the reliance on pressure, this approach is subject to inaccuracies due to changes in air pressure that have nothing to do with altitude change. For highly accurate altitude measurements, periodic calibration is necessary. An aircraft must recalibrate before takeoff and make

adjustments in flight based on measurements reported by air traffic control. For casual-use with personal devices, calibration is usually less necessary as time scales are smaller and often only relative changes are needed. More sophisticated altimeters blend pressure readings with other sensors (e.g., accelerometer, GPS) to provide a more accurate altitude estimate.

Another way to measure altitude is by using the Global Positioning System (GPS). Four or more satellite readings are necessary to provide an altitude estimate, though generally more satellites than four are available overhead at a given time. Figure 1.2 shows 9 visible satellites for a particular location at a particular point in time (see [here](#) for an animated version). The GPS device's position is triangulated based on the known positions of the satellites and the passage time for the radio waves (like a radar altimeter). The computed altitude is referenced relative to sea level, and is thus a *geometric altitude*. GPS-derived altitude estimates are subject to inaccuracies caused by things like signal blockage and reflections (e.g., buildings, trees) and interaction with particles in the ionosphere¹ that delay the radio signals. GPS altitude measurements are less accurate than their horizontal measurements (with a rule of thumb of 3X difference in accuracy). Typically, a GPS altitude measurement is accurate to within about 10–20 meters (35–70 ft), if the signal is good.

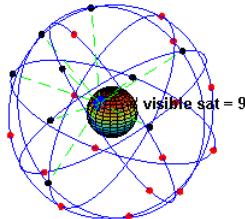


Figure 1.2: Visualization of GPS satellite locations around the earth for a given instant in time. Wikimedia Commons, [El pak](#), public domain.

¹part of the atmosphere that is ionized by solar radiation and can include parts of the mesosphere, thermosphere, and exosphere

1.3 Hydrostatic Equation and Geopotential Altitude

You probably already know that as you go up in altitude the air pressure and density generally decrease, but we need to be able to quantify this relationship. We can relate pressure changes with altitude using a hydrostatic balance (hydrostatic means a fluid at rest). Consider an arbitrary cubic control volume (Fig. 1.3) and let's assume the atmosphere is at static equilibrium. Gravity exerts a force downward from the center of the control volume with a magnitude of: $\rho dV g$ where dV is the volume of the control volume. There are also pressure forces on the top and bottom. The pressures can't be identical, otherwise the total vertical force (including gravity) would not sum to zero. Because this is an infinitesimally small control volume, the pressure on top differs by some infinitesimal amount $p + dp$. Thus the force balance in the vertical direction is:

$$pdA - (p + dp)dA - \rho g dV = 0 \quad (1.1)$$

We can simplify this, and use the fact that $dV = dAdh$:

$$\begin{aligned} pdA - (p + dp)dA - \rho g dV &= 0 \\ -dpdA - \rho g dV &= 0 \\ -dp - \rho g dh &= 0 \\ dp &= -\rho g dh \end{aligned} \quad (1.2)$$

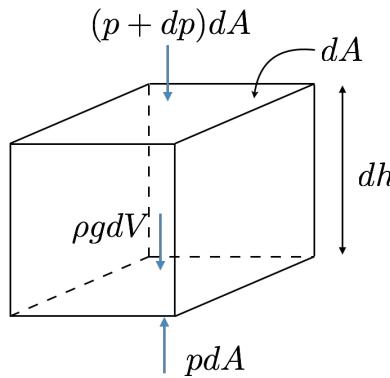


Figure 1.3: An arbitrary control volume in the atmosphere with pressure and gravitational forces.

This is called the hydrostatic equation, which we can integrate to find pressure as a function of altitude. One complication in the integration is

that g actually varies with altitude. However, it is conventional to use a constant value of g in the integration, specifically the acceleration of gravity at sea-level g_{sl} . In making this choice, h no longer corresponds to the geometric altitude, but rather is a fictitious altitude called the geopotential altitude. In other words:

$$gdh = g_{sl}dh_{GP} \quad (1.3)$$

where the subscript GP corresponds to the geopotential altitude, and dh without a subscript is the geometric altitude. We can use Newton's universal law of gravitation to relate the acceleration between sea-level and some other altitude:²

$$g = g_{sl} \left(\frac{R}{R + h} \right)^2 \quad (1.4)$$

where R is the mean radius of the earth (mean sea-level), and h is the geometric altitude from sea level corresponding to the location where we compute g . If we substitute this relationship into the differential equation above we have a relationship between geometric and geopotential altitudes:

$$dh_{GP} = \left(\frac{R}{R + h} \right)^2 dh \quad (1.5)$$

This equation can be integrated (skipping those steps) yielding:

$$h_{GP} = \frac{Rh}{R + h} \quad (1.6)$$

or inverted:

$$h = \frac{Rh_{GP}}{R - h_{GP}} \quad (1.7)$$

Thus, we can easily convert between geometric and geopotential altitudes. However, most of the time the distinction between geometric and geopotential altitude is not particularly relevant. Only at altitudes larger than 65 km do the two altitudes differ by more than 1%.

²This is still just an approximation because it assumes that the earth is a perfect sphere and that its mass distribution is radially symmetric.

1.4 International Standard Atmosphere

Generally pressure and density both decrease with altitude, but, as noted in the description of the atmosphere layers, the temperature profile is more complex. We need to quantify this relationship so we can estimate thermodynamic properties as a function of altitude. A common way this is done is by using the international standard atmosphere. The international standard atmosphere is an idealized, steady-state model. It does not represent actual atmospheric conditions, which of course vary both spatially and temporally across the earth. Instead, it is a empirical model meant to represent typical conditions.

The core of the standard atmosphere is the definition for the temperature distribution with altitude. The temperature profile is tabulated in Table 1.1, which contains some defined altitudes, and the temperature gradient (also known as the lapse rate) between these altitudes. For example, from -0.61 km to 11 km the temperature decreases with altitude at a rate of -6.5 K/km. From 11 km to 20 km the temperature is constant (isothermal), and so on. Note that the table uses geopotential altitude, not geometric altitude, but as discussed that's generally not an important distinction for most of the range of this table.

The temperature profile is depicted pictorially in Fig. 1.4. For reference, the figure contains an aircraft around the altitude band where commercial aircraft fly, and a rocket near the Karman line or the “edge of space”.

There is no precise transition from earth’s atmosphere to outer space; the Karman line is a convention defined as 100 km altitude.

To compute temperature as a function of altitude, we still need a starting point, or reference point. The reference conditions are defined at sea level, and are tabulated in Table 1.2.

Table 1.2: Sea-level reference conditions.

T_{sl}	temperature	288.15 K
p_{sl}	pressure	1.01325×10^5 Pa
g_{sl}	acceleration of gravity	9.80665 m/s ²

With a reference temperature, and the defined temperature gradients, we can estimate temperature at any altitude within the defined range.

Pressure and density require additional relationships. We already derived a hydrostatic balance that relates pressure and density as a function of

Table 1.1: Defined temperature gradients (lapse rates) in the standard atmosphere.

Geopotential altitude (km)	Temperature gradient (K/km)
-0.61	-6.5
11	0
20	+1
32	+2.8
47	0
51	-2.8
71	-2.0
84.852	

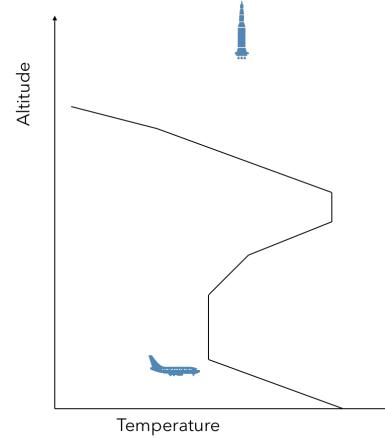


Figure 1.4: Graphical depiction of the temperature profile for the standard atmosphere.

altitude Eq. (1.2). We can then relate pressure, density, and temperature through the ideal gas law:

$$p = \rho RT \quad (1.8)$$

where R is the specific gas constant for air, and as defined in the standard atmosphere is 287.053 J/(kg-K).

There are two types of regions in the standard atmosphere: constant temperature gradients (isothermal), and linear temperature gradients. For both cases, we can derive analytic expressions for these thermodynamic quantities. In the following derivations we will use the acceleration of gravity at sea level, and thus all altitudes are geopotential altitude. However, we will drop the GP subscript for convenience.

Let's start with an isothermal layer. Using the hydrostatic equation, but replacing density with pressure and temperature using the ideal gas law:

$$dp = -\frac{p}{RT} g_{sl} dh \quad (1.9)$$

Separating terms to prepare for integration:

$$\frac{dp}{p} = -\frac{g_{sl}}{RT} dh \quad (1.10)$$

Because this is an isothermal region, everything in the right-hand

integrand is constant. Integrating between two points yields:

$$\ln\left(\frac{p_2}{p_1}\right) = -\frac{g_{sl}}{RT}(h_2 - h_1) \quad (1.11)$$

or

$$p_2 = p_1 \exp\left(-\frac{g_{sl}(h_2 - h_1)}{RT}\right) \quad (1.12)$$

For a linear temperature gradient we can express the temperature distribution as:

$$T_2 = T_1 + s(h_2 - h_1) \quad (1.13)$$

where s is the lapse rate. In differential form: $dT = sdh$. We start with the same differential expression from before:

$$\frac{dp}{p} = -\frac{g_{sl}}{RT} dh \quad (1.14)$$

This time temperature is not constant so the integral is not as straightforward. While we could express T as a function of altitude and integrate with respect to altitude, it will be easier to integrate with respect to temperature. We make the substitution $dh = dT/s$:

$$\frac{dp}{p} = -\frac{g_{sl}}{sRT} dT \quad (1.15)$$

Pulling out the constants for integration

$$\int_{p_1}^{p_2} \frac{dp}{p} = -\frac{g_{sl}}{sR} \int_{T_1}^{T_2} \frac{dT}{T} \quad (1.16)$$

After integrating, using the rules of logarithms, and simplifying we have:

$$p_2 = p_1 \left(\frac{T_2}{T_1}\right)^{\frac{-g_{sl}}{sR}} \quad (1.17)$$

Thus, with a starting point for pressure, and a known temperature distribution we can now compute the pressure as a function of altitude. From pressure and temperature, we can compute density from the ideal gas law:

$$\rho = \frac{p}{RT} \quad (1.18)$$

Viscosity and the speed of sound are both functions of temperature. The absolute (dynamic) viscosity can be estimated using Sutherland's Law:

$$\mu = \frac{\beta T^{3/2}}{T + S} \quad (1.19)$$

where $\beta = 1.458 \times 10^{-6} \frac{kg}{smK^{1/2}}$ and $S = 110.4$ K. The speed of sound (a) is given by the following expression for an ideal gas:

$$a = \sqrt{\gamma RT} \quad (1.20)$$

where γ is the ratio of specific heats, which is $\gamma = 1.4$ for air.

[Online calculators](#) can be handy for looking up atmospheric properties and/or comparing your results. Further details on this model can be found in a [NASA report](#).

1.5 Curve Fit to Standard Atmosphere

The piecewise nature of the standard atmosphere can be a bit cumbersome to work with.³ Drela devised a curve fit that requires only one equation for temperature and one for pressure [1]. The fit is only applicable for altitudes below 47 km, but that spans the range of interest for most aerodynamic aerospace applications. The equations are:⁴

$$T(h) = T_{sl} - 71.5 + 2 \ln [1 + \exp(35.75 - 3.25h) + \exp(-3 + 0.0003h^3)] \quad (1.21)$$

$$p(h) = p_{sl} \exp \left(-0.118h - \frac{0.0015h^2}{1 - 0.018h + 0.0011h^2} \right) \quad (1.22)$$

where h must be in km and T in K. Figure 1.5 compares these fits against the standard atmosphere.

1.6 Other Atmospheric Models

Many other atmospheric models exist, and the appropriate one depends on the type of data and level of precision required. We will not provide a

³Also, it is not ideal for use in gradient-based optimization because it is not continuously differentiable.

⁴The temperature equation is slightly modified from Drela's to explicitly include sea level temperature.

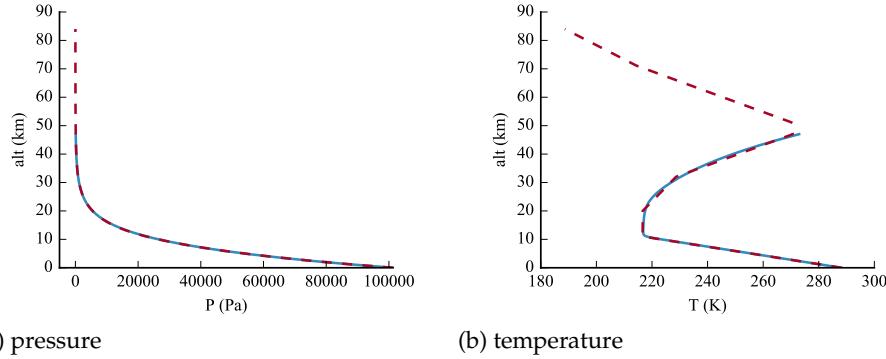


Figure 1.5: Comparison between standard atmosphere (red dashed line) and the curve fit (blue solid line).

comprehensive list, but merely mention two examples. The [NRLMSISE-00](#) is a model of the upper atmosphere (up to 1000 km). It is an empirical model and is particularly useful for higher altitudes (mesosphere and thermosphere). This atmospheric model can be used for things like estimating drag on satellites, which would be far too inaccurate using the standard atmospheric model.

[Cameron Beccario](#) developed a nice visualization based on data from multiple sources. Looking at just temperature, for example, it is clear that the assumption that temperature only varies with altitude is a crude one. Latitude, longitude, and time may be critical inputs depending on the accuracy required for the application.

Bibliography

- [1] Drela, M., *Flight Vehicle Aerodynamics*, MIT Press, Feb 2014.

CHAPTER 2

Fundamentals

This chapter discusses a range of topics that are foundational in our study of flight vehicle design. These topics include: nomenclature for airfoils and wings, common nondimensional parameters, and lift curves and drag polars.

2.1 Airfoil Geometry

A starting point for creating a lifting surface (e.g., a wing), is to define the airfoils that it is composed of. An airfoil is a 2D cross-section of a lifting surface (Fig. 2.1). Airfoils use streamlined shapes to produce lift with low drag.

Figure 2.2 highlights some of the main nomenclature associated with an airfoil. The *leading edge* of an airfoil is the forward most point, whereas the *trailing edge* is the aft most point. The *chord line* is a straight line connecting the leading and trailing edge. Its length is called its *chord length* or just *chord*.

The shape of the airfoil between the leading and trailing edge is divided into *upper* and *lower* surfaces. For some applications, like a propeller, the concept of upper and lower is less natural (because upper is not necessarily up) and so the two sides are called the *suction* and *pressure* side respectively.

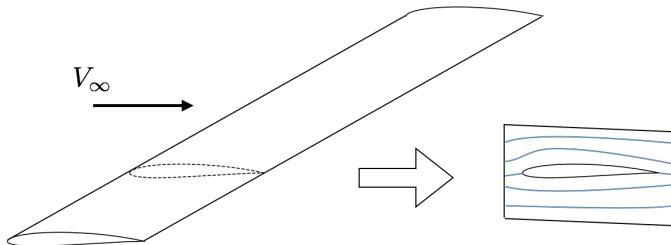


Figure 2.1: An airfoil is a 2D cross-section of a wing or other lifting surface. The streamlined shape is designed to produce lift efficiently with low drag.

The curve that is exactly halfway between the upper and lower surface is called the *camber line*. The local camber is the distance between the chord line, and the camber line. The maximum camber is simply referred to as the airfoil's *camber*. The local thickness is the distance between the upper and lower surface. The maximum thickness is simply referred to as the airfoil's *thickness*. An important parameter for an airfoil is its thickness-to-chord ratio, written as t/c . A typical t/c for aircraft airfoils (sometimes referred to as thickness) is between 8-14%.

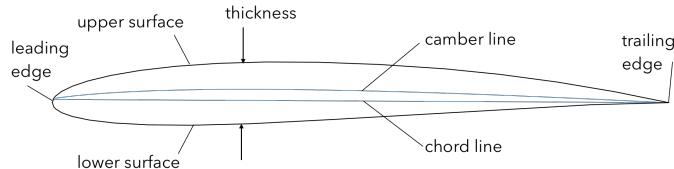


Figure 2.2: Some nomenclature for an airfoil.

The shape of an airfoil can be described in many ways. One well-known designation is the NACA 4-series. A NACA 4-series airfoil is defined by a set of equations, one for thickness and one for camber, that are parameterized by four digits. To describe what these parameters are we will use the NACA 2412 as an example (see Fig. 2.3). The first digit defines the maximum camber as a percentage of the chord (2% chord in this example). The second digit defines the location of the maximum camber in tenths of chord (0.4 chord in this example). The last two digits define the thickness of the airfoil in percent chord ($t/c = 0.12$ in this example). A

NACA 00XX (where XX means any number), is a *symmetric section*. In other words, there is no camber so the airfoil is symmetric. Symmetric sections are commonly used in tails and winglets.

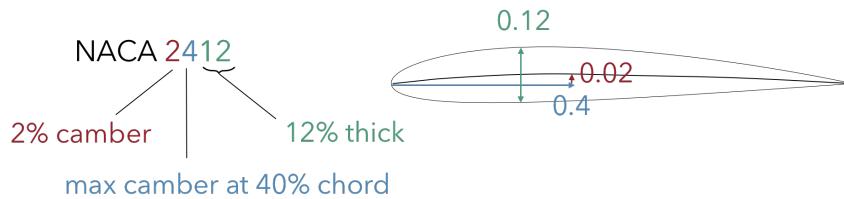


Figure 2.3: An illustration of the key dimensions for a NACA 2412 airfoil with unit chord.

While the NACA 4-series are well known, and are still sometimes used in simpler geometries like tails, they are rather limited and thus not generally useful for design. Several more general parameterization methods exist, like B-splines, but we will not elaborate on these methods.

Airfoils for supersonic sections, like a low-sweep wing of a supersonic aircraft or fins on a rocket, will have a sharp rather than a blunt leading edge (Fig. 2.4). This is because shock waves become the dominant source of drag and a sharp leading edge decreases the strength of the shock waves.

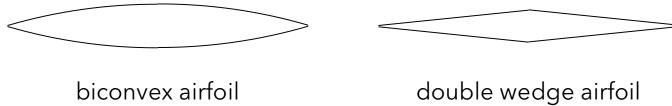


Figure 2.4: Two examples of airfoils used for locally supersonic flow.

2.2 Wing Geometry

Figure 2.5 highlights some of the main nomenclature associated with a wing. The *wingspan* or just *span* (b) is the distance from tip-to-tip, projected on the plane of the wing. *Wing area* is generally denoted with an S and different types of wing areas are used depending on the application. In general, the chord distribution can vary continuously along the wing, but for a simple linearly tapered wing it can be parameterized in terms of the *root chord* (c_r) and *tip chord* (c_t). Similarly, the *twist* distribution can vary continuously, but for a linearly twisted wing is parameterized by the

root twist (θ_r) and *tip twist* (θ_t). The *dihedral angle* (ϕ) is a rotation angle from horizontal. It can also be a distribution, rather than a constant value. *Sweep* is usually measured from the quarter-chord (a line/curve connecting 1/4 the distance along the chord), but is sometimes measured from the leading edge instead. Sweep need not be constant, but can vary continuously along the span.

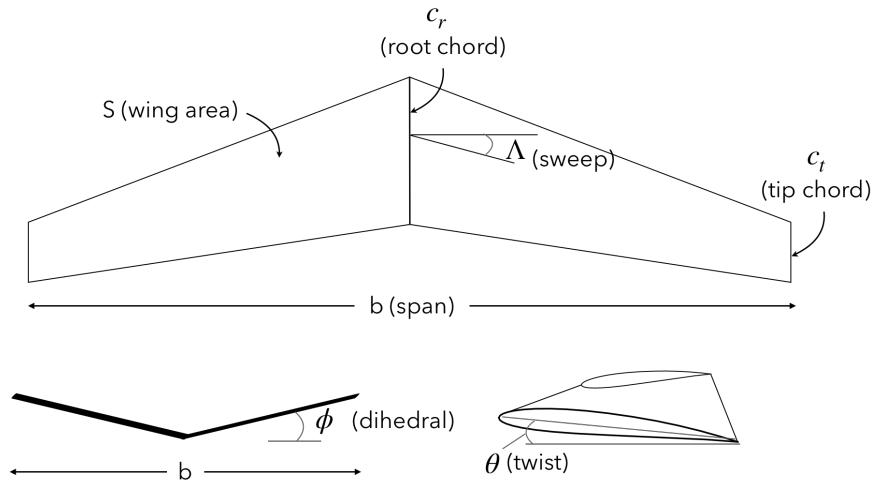


Figure 2.5: Some nomenclature for a wing or other lifting surface.

For a simple wing we define the *taper ratio*, which is the ratio of the tip chord to the root chord:

$$\lambda = \frac{c_t}{c_r} \quad (2.1)$$

Another fundamental parameter is aspect ratio, which applies to any wing, and is defined as:

$$AR = \frac{b^2}{S} \quad (2.2)$$

Figure 2.6 shows wings with a high and low taper ratio and a high and low aspect ratio.

Often, we need a representative chord length for a wing. This is needed for things like Reynolds number calculations and for normalizing moments (both of these will be reviewed momentarily). A simple option is the *mean geometric chord*:

$$\bar{c} = \frac{S}{b} \quad (2.3)$$

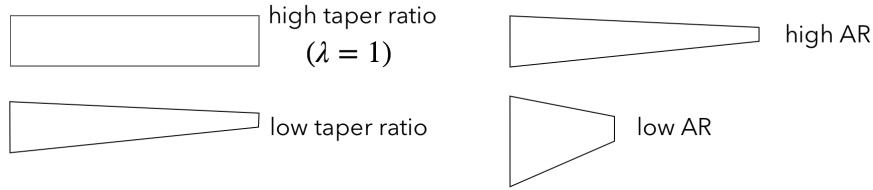


Figure 2.6: Examples of a low and high taper ratio and a low and high aspect ratio.

However, this is not often used in practice. A better representative length is the *mean aerodynamic chord*. The mean aerodynamic chord is a chord-weighted average chord:

$$c_{mac} = \frac{2}{S} \int_0^{b/2} c^2 dy \quad (2.4)$$

The mean aerodynamic chord is particularly important for stability analysis. For a linearly tapered wing the mean aerodynamic chord is:

$$c_{mac} = \frac{2}{3} \left(c_r + c_t - \frac{c_r c_t}{c_r + c_t} \right) \quad (2.5)$$

2.3 Aerodynamic Forces and Moments

A body in a fluid, is subject to both pressure and shear forces¹ as seen in Fig. 2.7. The differences in pressure drag and shear drag will have important design implications, as we will see later. For now the important point is that all aerodynamic forces come from pressure or shear. From statics you should recall that we can reduce any distributed load into an equivalent representation with point loads/moments. In this case, the result of the pressure and shear loads can be equivalently described by two forces and one moment as shown in Fig. 2.8 (for a 2D airfoil, in general there is three forces and three moments).

The location where the forces/moment are specified could be any point on the body, but typically it is defined at the *quarter-chord*. The quarter-chord is the point on the chord line that is 1/4 of the chord length from the leading edge. The reason why the quarter-chord is typically used

¹To be accurate, pressure and shear are not actually forces, but rather are tractions. Their application over some area results in a force.

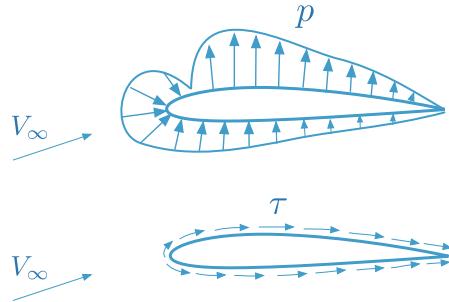


Figure 2.7: Pressure and shear are the two types of forces for a body immersed in a fluid.

is that the quarter-chord is the theoretical location for a thin airfoil's aerodynamic center. We will define the aerodynamic center in the chapter on stability (Chapter 5). For now just think of it as the most convenient point to resolve aerodynamic forces/moment at. The aerodynamic center for a wing approximately occurs at the quarter chord point of the spanwise location on the wing where the local chord is equal to the mean aerodynamic chord. Hence, the significance of the mean aerodynamic chord.

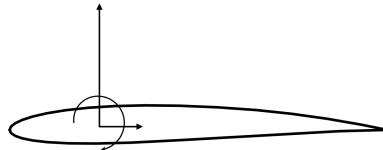


Figure 2.8: The pressures and shear over any body can be resulted into an equivalent set of forces and moments, two forces and one moment in 2D, and three forces and three moments in 3D.

The moment for an airfoil is called the pitching moment. Two different axes are typically used in defining the forces as shown in Fig. 2.9. The first is the body-axes, which are aligned with the airfoil. In this coordinate system to the two forces are called the *normal force* (N) and the *axial force* (A). For aerodynamicist, a more frequently used coordinate system is the wind-axes, which are aligned with the freestream direction, which is opposite the direction of flight. In this coordinate system the forces are called *lift* and *drag*. Drag is always defined parallel to the freestream direction, and lift is always defined perpendicular to the freestream

direction. The angle between the freestream and the chord line is called the angle of attack (α). In practice, it is usually the aircraft that is rotated relative to a horizontal flight direction as shown in the insert. Mathematically it is equivalent to draw the chord line as horizontal, with the freestream rotated as shown in the main part of the figure. The latter is merely a convenience.

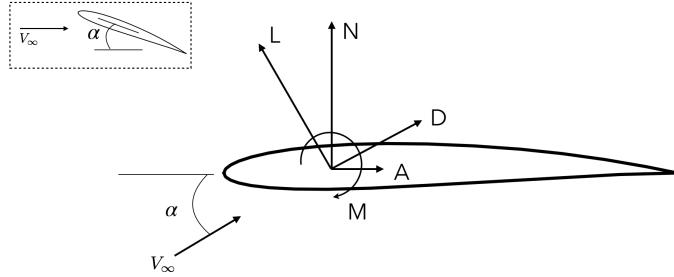


Figure 2.9: Definitions for normal and axial force in the body coordinate system, or lift and drag in the wind coordinate system. Angle of attack is the angle between the freestream direction and the chord line.

2.4 Nondimensional Parameters

In fluid dynamics you were exposed to the power of nondimensional numbers, and these are used heavily in aerodynamics. Before discussing further, let's review a few nondimensional numbers from fluid mechanics. The *Reynolds number*:

$$Re = \frac{\rho V c}{\mu} \quad (2.6)$$

where ρ and μ are the fluid's density and dynamic viscosity respectively, V is the freestream (or local) speed, and c is some relevant length scale used in normalization, which for an airfoil would be the chord length and for a wing would be the mean aerodynamic chord. The Reynolds number is a ratio of inertial to viscous forces. For a commercial transport, the Reynolds number is typically in the millions or tens of millions. In other words, outside of the boundary layer, the air speed is so high that inertial forces are much, much more important than viscous forces. For a small UAV, the Reynolds number might be of order 10,000.

The *Mach number*:

$$M = \frac{V}{a} \quad (2.7)$$

is a ratio between the freestream (or local) speed and the speed of sound. In other words, a Mach number of one means that the vehicle is moving at the speed of sound. Mach numbers less than 1 are called *subsonic*, at 1 is called *sonic*, and above 1 is called *supersonic*. Mach numbers larger than about 5 are called *hypersonic*, although the number 5 is somewhat arbitrary as there is no fundamental change in flow physics right at that point like there is at the sonic point. A flow field with Mach numbers less than about 0.3 is essentially incompressible, and thus the Mach number is not a relevant parameter for that flow. There are other parameters you were exposed to like the Froude number (ratio of inertial to gravitational forces), but, excepting lighter-than-air vehicle, the buoyancy force of the air is negligible. For aerodynamics generally Reynolds number and Mach number are the most important.

The power of nondimensionalization is that it can express a complicated relationship succinctly and universally, and it is the basis for all wind tunnel testing. For example, the drag of a given airfoil is a function of many things, the angle of attack, air density, viscosity, speed of sound, freestream speed, its size, etc.:

$$D = f(\alpha, \rho, \mu, V_\infty, a, c) \quad (2.8)$$

However, we can express the same relationship in terms of a non-dimensional drag coefficient as:

$$C_D = f(\alpha, Re, M) \quad (2.9)$$

The power of this form is that the drag coefficient doesn't depend on density directly (for example), but rather on Reynolds number. Thus, if we create the same Re and M with a scaled geometry, we can determine the drag coefficient that would be experienced on the full airplane. Similarly, for a given shape, we can tabulate or calculate lift and drag coefficients as a function of α , Re and if high speeds are important M and re-use them for various applications.

Because an airfoil is 2D, the lift/drag is actually lift/drag per unit span, which we will represent as L' and D' . We will also frequently use the *dynamic pressure*, which is defined as:

$$q_\infty = \frac{1}{2} \rho_\infty V_\infty^2 \quad (2.10)$$

The lift/drag/moment coefficients are defined as follows:

$$c_l = \frac{L'}{q_\infty c} \quad (2.11)$$

$$c_d = \frac{D'}{q_\infty c} \quad (2.12)$$

$$c_m = \frac{M'}{q_\infty c^2} \quad (2.13)$$

where c is the chord. Note that lower case letters are used to denote 2D (airfoil) coefficients.

A three-dimensional body, like a wing, is defined similarly but with an extra length dimension in the denominator:

$$C_L = \frac{L}{q_\infty S_{ref}} \quad (2.14)$$

$$C_D = \frac{D}{q_\infty S_{ref}} \quad (2.15)$$

$$C_m = \frac{M}{q_\infty S_{ref} c} \quad (2.16)$$

Note the capital letters are used for denoting 3D coefficients. The type of reference area S_{ref} used must be specified when providing coefficients.

For a wing it is typically the planform area, whereas for a fuselage the cross-sectional area is typical. For a full aircraft the *trapezoidal reference area* is typically used (Fig. 2.10). This area is a trapezoid that matches the main portion of the wing and extends all the way to the centerline. Its not the exact same shape as the wing, but is easier to use. Any area can be used as long as it is consistent and is defined, it merely provides the length scales used in the normalization.

One last important nondimensional parameter is the pressure coefficient:

$$C_p = \frac{p - p_\infty}{q_\infty} \quad (2.17)$$

We will make use of this parameter in the next section.

2.5 Pressure Distributions

In fluid dynamics you likely studied flow around a cylinder. Figure 2.11a shows streamlines, and the pressure distribution around a cylinder,



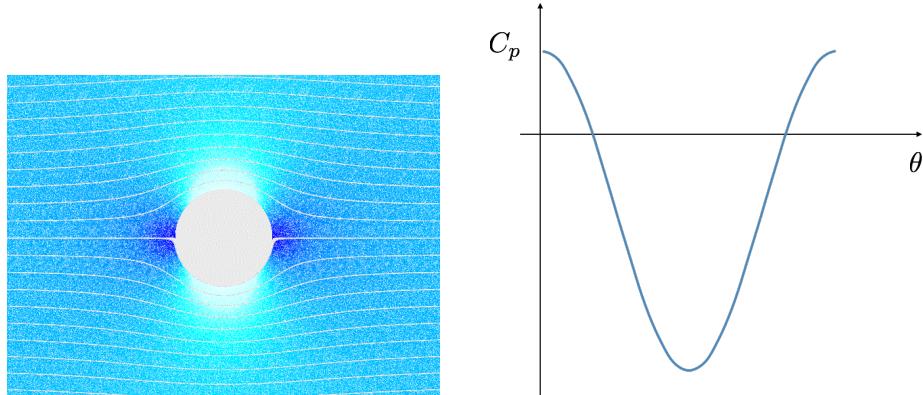
Figure 2.10: Half of the trapezoidal reference area shown overlaid on the aircraft. Attribution: [Airplane Drawing Top View](#), CC BY-NC 4.0

assuming inviscid flow. The dark blue indicates high pressure, while white indicates low pressure.

We can visualize the pressure distribution around the cylinder in a 2D plot as shown in Fig. 2.11b. The angle θ starts at 0 at the leading edge of the cylinder and wraps around to the back (in either direction as the flow is symmetric). The pressure coefficient starts at 1 at the leading edge (which indicates a stagnation point in incompressible flow, see Eq. (2.17)), and reaches a low of -3 and the top/bottom of the cylinder.

Recall that this is an idealized, inviscid case and does not represent flow around a real cylinder. One way to think about this is to use an analogy of a ball rolling down a hill. It starts out at the top with no kinetic energy, all the potential energy is converted to kinetic energy at the bottom, and if there is no friction it will return to the exact same height. The pressure distribution follows the same kind of pattern. However, a real flow is viscous, and like the analogy with the ball, energy is lost and so it cannot return to the initial energy state.

As the flow traverses the front part of the cylinder it is moving from high pressure to low pressure, this is called a *favorable pressure gradient* (“downhill” in our analogy). Conversely, as it moves from the top of the cylinder to the back it is moving from low pressure to high pressure, which is called an *adverse pressure gradient* (“uphill” in our analogy). With the presence of friction the flow cannot stay attached to the body all the way through the adverse pressure gradient. The flow will separate from the body creating a wake as depicted in Fig. 2.12. The corresponding



(a) Streamlines and pressure contours for inviscid flow around a cylinder. Image from [Thierry Dugnolle](#), Wikimedia, public domain.

(b) Pressure distribution around the cylinder starting from $\theta = 0$ corresponding to the leading edge and traversing to the back. The pressure coefficient starts at 1 at the stagnation points, drops to -3 at the top/bottom of the cylinder, and recovers back to 1 and the back end.

Figure 2.11: Pressure distribution around a cylinder

pressure distribution is shown in Fig. 2.13a. Note that the pressure distribution does not recover to the same height. Now the pressure distribution is no longer symmetric front to back (higher pressure in front), hence there is now a net drag force.

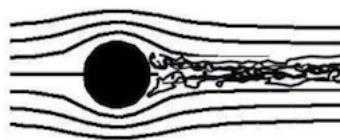
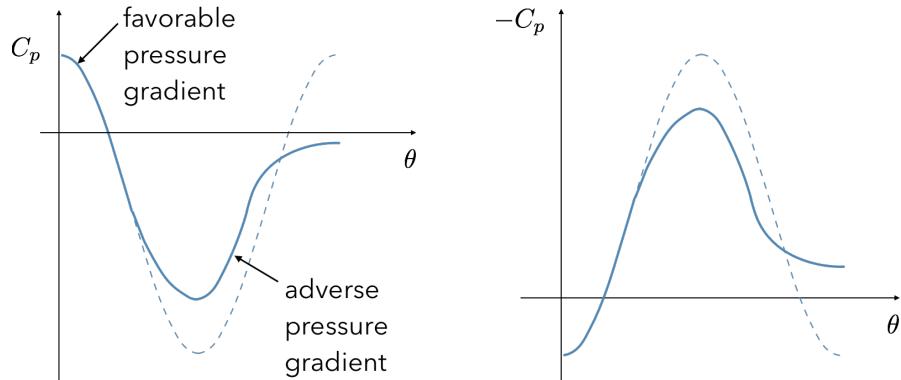


Figure 2.12: A real, viscous, fluid cannot navigate the adverse pressure gradient without separating and creating a wake. Image from [NASA](#), public domain.

By convention we usually plot the negative of the pressure distribution on the y-axis, for reasons we will see shortly. For the cylinder this is shown in Fig. 2.13b. All the information is the same, but the rolling ball analogy is less natural.

Now, let's look at an airfoil. For an airfoil the pressure distribution and



(a) The adverse gradient is too strong to navigate without separating causing the real pressure distribution to deviate from the idealized inviscid case.

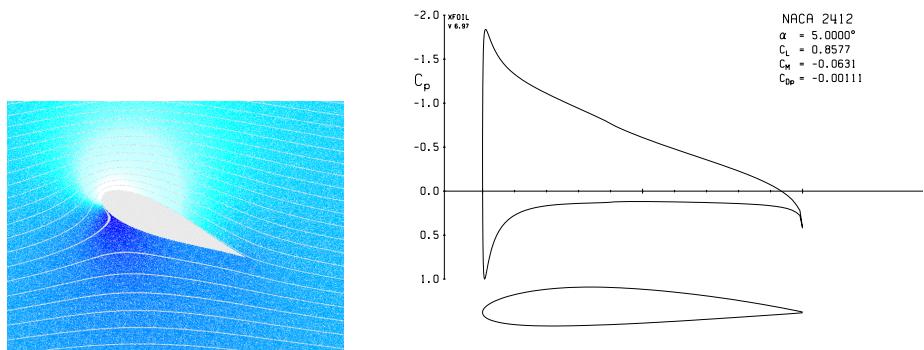
(b) Same figure but with $-C_p$ plotted on the y-axis.

Figure 2.13: Pressure distribution for a viscous cylinder (solid) versus inviscid (dashed). The parameter θ is the position on the cylinder, varying from $\theta = 0$ at the leading edge and traversing to the trailing edge.

streamlines are visualized in Fig. 2.14a, and the pressure coefficient along the surface of the airfoil is plotted in Fig. 2.14b (as generated by XFOIL). Unlike, the cylinder, the airfoil is not symmetric and so we see a pressure distribution for the upper and lower surfaces. Note that the negative of the pressure coefficient is plotted on the y-axis. We can now see the reason for this choice. The upper part of the curve (negative pressures) corresponds to the upper surface of the airfoil, and the lower part of the curve (positive pressures) corresponds to the lower surface of the airfoil.

The steepness of the adverse gradient is not as severe as the cylinder, and thus the flow can remain attached up to the trailing edge (likely there is a small amount of separation right at the trailing edge). This greatly reduces the pressure drag and is the reason the airfoil has the streamlined shape as opposed to the blunt shape of the cylinder. Notice that the pressure coefficient on the aft end of the airfoil recovers to a much higher value than does the cylinder (though it is not a perfect recovery and there is still pressure drag).

The pressure coefficient plot can tell us a lot about the performance of the airfoil. For example, we can show that the lift coefficient of the airfoil is



(a) Streamlines and pressure contours around an airfoil for inviscid potential flow. Image from [Thierry Dugnolle](#), Wikimedia, public domain.

(b) Pressure distribution around airfoil as generated by XFOIL.

Figure 2.14: Pressure around an airfoil.

directly related to the area between the two curves²:

$$c_l \approx \int_0^1 \left(C_{p_l} - C_{p_u} \right) d\left(\frac{x}{c}\right) \quad (2.18)$$

where C_{p_l} and C_{p_u} are the pressure coefficient on the lower and upper surface respective, and x/c is the normalized axial position along the airfoil from 0 to 1.

As mentioned, the steepness of the adverse pressure gradient is a critical factor in keeping the flow attached and thus reducing drag. Much of the airfoil shaping is done to provide as much lift as possible while maintain attached flow to keep drag down. Note that these are competing priorities. As lift is increased (area between the curves) there will generally be a steeper adverse pressure gradient eventually leading to stall (discussed more in the next section).

Figure 2.15 shows the pressure distribution for an airfoil flying at higher speeds. In this case, the airfoil is flying at subsonic speeds, but acceleration

²Actually, this formula gives the normal force rather than lift since we'd need the angle of attack for the latter, and it ignores the contribution of shear stress. But in most applications the normal force and lift force differ only slightly and the contribution of shear stress to lift is negligible.

over the airfoil causes the local Mach number to become supersonic over portions of the airfoil. In this figure the critical pressure coefficient (C_p critical) separates subsonic from supersonic flow. After the flow becomes supersonic a shock wave causes an abrupt increase in pressure during the recovery. A strong shock causes a lot of drag, and so airfoil shaping in this flow regime must take care to reduce shock wave strength while maintaining lift.

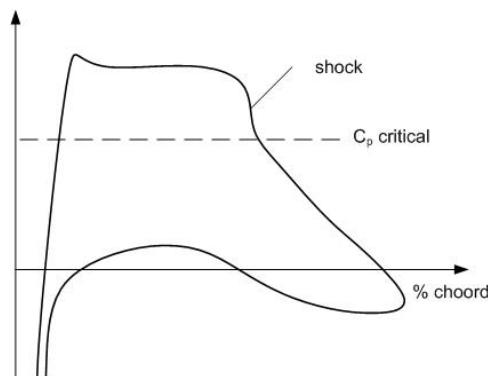


Figure 2.15: Pressure coefficient for an airfoil flying at subsonic speeds, but with local supersonic flow from acceleration over the airfoil. Figure by [philip poppe](#), Wikipedia, public domain.

2.6 Lift Curves and Drag Polars

As discussed previously, the force/moment coefficients for a given airfoil are a function of the angle of attack, Reynolds number, and Mach number. For many applications we might only operate within a narrow range of Re and M where performance does not vary significantly with these parameters. In these cases the lift, drag, and moment coefficients are effectively only a function of angle of attack. Or, variation with one or both of Re and M may be important but it can still be helpful to visualize these as separate curves (e.g., a separate plot for a set of discrete Re).

The variation of lift coefficient with angle of attack is called the *lift curve* and an example is shown in Fig. 2.16. This example is for an airfoil (2D). If the lift curve was for a full wing or full airplane then the y-axis would show C_L . Lift varies linearly with angle of attack until approaching *stall*. Stall occurs because the angle of attack becomes too large, the flow

separates from the airfoil and causes a significant increase in drag and a decrease in lift.

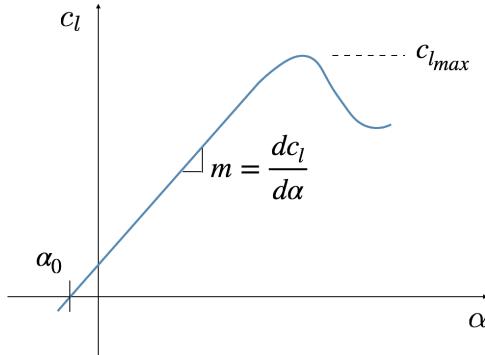


Figure 2.16: A notional lift curve slope highlighting the zero-lift angle of attack, lift curve slope, and maximum lift coefficient before stall.

The *maximum lift coefficient* that is achievable before stall is denoted as $c_{l_{max}}$ or $C_{L_{max}}$ in 3D. The slope of the linear portion of the lift curve is called the *lift curve slope*. It is usually denoted with a m or an a . Thin airfoil theory shows that the *lift curve slope of a thin 2D airfoil is theoretically 2π* . Real airfoils have a slope near this, but are slightly lower because of viscous effects.

A symmetric airfoil would have zero lift at an angle of attack of zero. Adding positive camber causes the *zero-lift angle of attack* (α_0) to become negative. That means that at zero angle of attack, a cambered airfoil produces lift. For a 2D airfoil, the lift coefficient prior to stall can be computed as:

$$c_l = m(\alpha - \alpha_0) \quad (2.19)$$

where m is the lift curve slope shown in Fig. 2.16.

If we add back in variation in Reynolds number the lift curve slope is generally negligibly affected. The main impact of Reynolds number is on stall ($c_{l_{max}}$), which is a viscous phenomenon. Mach number does not usually have a significant impact on lift, unless the change in Mach number is large.

The variation in drag coefficient with angle of attack (or lift coefficient) is called the *drag polar*. Figure 2.17 shows a notional drag polar, though more typically it is plotted as a function of lift coefficient rather than angle of attack. Again, for a full aircraft drag polar it would be plotted as C_D as a

function of C_L . Reynolds number is often important for drag polars (as would be expected since Reynolds number accounts for changes in viscosity). In this case a series of drag polars are shown at different Reynolds numbers. If the Mach number is less than 0.3 then there is generally no impact on drag, but for Mach numbers larger than 0.5 the impact of Mach number is significant.

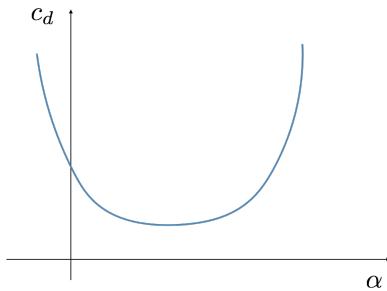


Figure 2.17: A notional drag polar. Plotted here versus angle of attack, but versus lift coefficient is more conventional.

Figure 2.18 shows the lift curve, drag polar, and moment curve for a NACA 2412 airfoil. Past stall the drag results are not reliable hence the very steep slope on the drag polar. Prior to stall, moment coefficient also has (theoretically) linear behavior and in practice is usually close to linear. By plotting the drag polar as shown in the figure (c_l on y-axis, c_d on x-axis), we can determine the maximum *lift to drag ratio* graphically. The lift to drag ratio L/D is an important measure of aerodynamic efficiency. If we take a straight line from the origin and tilt it until it just touches the drag polar, then the slope of that line is the max L/D , and the corresponding lift coefficient (and angle of attack) is the operating condition that maximizes L/D .

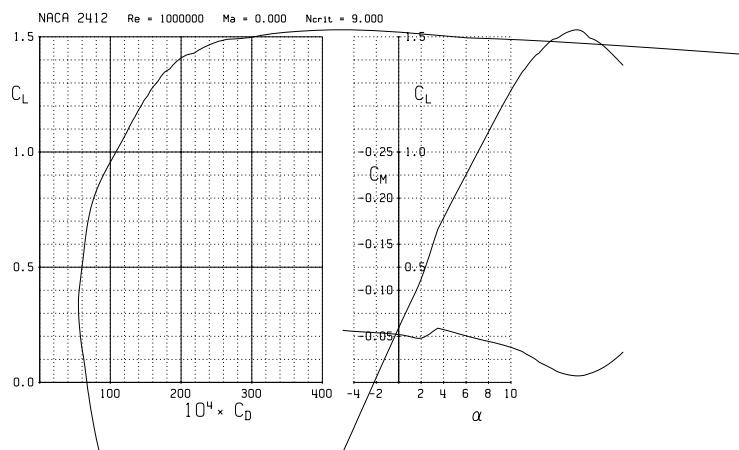


Figure 2.18: Lift curve, moment curve, and drag polar for a NACA 2412 airfoil at a Reynolds number of 1 million and a Mach number of 0. Figure is generated by XFOIL.

CHAPTER 3

Drag

The energy required by an aircraft is primarily used to overcome drag. Thus, drag estimation is a critical component of aircraft performance. This chapter provides an overview of some simplified methods to estimate drag. The interested reader is encouraged to study aerodynamics to learn more about the physics behind these methods and to be exposed to more accurate means to estimate drag.

3.1 Breakdown

Drag is a vector (i.e., a magnitude and a direction). As discussed in Chapter 2, drag is defined as the aerodynamic force in the direction of the freestream. Thus, the essence of drag estimation is determining its magnitude.

To make the process approachable, drag is usually decomposed into different components, a process called drag breakdown. There are many ways this can be done, the essential requirements are that all important sources of drag are accounted for, and that no source is double counted. In this text we will breakdown drag into three main sources:

1. *Parasitic* drag. Two types of stresses act on any body in a fluid: normal stresses (pressure) and shear stresses. Their contributions to

drag are known as pressure drag and skin friction drag, and their sum is called parasitic drag, zero-lift drag, or sometimes just viscous drag. For many aircraft, skin friction drag is the largest component of drag.

2. *Induced* drag. Lift cannot be created without expending energy. The associated additional drag caused by lift is called induced drag, vortex drag, or lift-dependent drag.
3. *Compressibility* drag. At high speeds shock waves are formed, and energy is carried away by these shock waves. This type of drag is called compressibility drag or wave drag. For Mach numbers less than 0.3 compressibility drag is generally negligible. However, for commercial aircraft, which typically fly around Mach 0.75–0.9, or jets which can fly well above the speed of sound, this is an important if not dominant source of drag.

We will use the subscripts p for parasitic, i for induced, and c for compressibility. The total drag is the sum of these components.

$$D = D_p + D_i + D_c \quad (3.1)$$

3.2 Parasitic Drag

Parasitic drag is a combination of skin friction and pressure drag. Each of these will be discussed separately, then combined into a total parasitic drag estimate.

3.2.1 Skin Friction Drag

Skin friction drag is the drag caused by viscous shear stresses acting over a body. In this section we will provide an overview of the physical mechanism and a means to estimate skin friction drag for a flat surface. From an introductory course in fluid mechanics or aerodynamics you may recall the concept of a boundary layer. For high Reynolds number flow, typical of most aircraft, the flow is essentially inviscid everywhere, except near the body. This region near the body where viscous effects are important is called the boundary layer.

At the body the flow must satisfy the *no-slip condition*. The no-slip condition states that the relative fluid velocity at the surface of a solid

object must be zero. Physically, this condition arises from molecular level mixing. In reality the fluid is not a continuum but rather a large collection of molecules. Molecules exchange energy, and thus speed with nearby molecules. Far away from the body the average speed of the molecules must be the freestream speed. But at the surface the average speed must be zero because the molecules are exchanging energy with the solid surface that is not moving (in the frame of reference of the solid object). Thus, a boundary layer forms to transition between the solid surface and the external velocity.

Figure 3.1 shows a diagram of typical boundary layer behavior over an airfoil. The boundary layer starts from a stagnation point, and at first the boundary layer is *laminar*. Lamina means layer, and a laminar flow moves in “layers”. What this really means is that mixing only occurs at a molecular level, and so the fluid behaves almost like layers of flow with different speeds. In contrast, turbulent flow is characterized by chaotic mixing across many scales.

At some point the laminar boundary layer becomes unstable and transitions to turbulent flow. Many things can cause these instabilities to occur sooner including increasing surface roughness, increasing Reynolds number, adverse pressure gradients, and surface heating. For the high Reynolds numbers of aircraft, laminar flow can typically only be sustained for about 10–20% chord, except for favorable conditions with specially shaped designs.

The boundary layer will generally not be able to stay attached to the body in an adverse pressure gradient and will separate, leaving a wake behind. Ideally, this wake is as small as possible as will be discussed in connection with pressure drag.

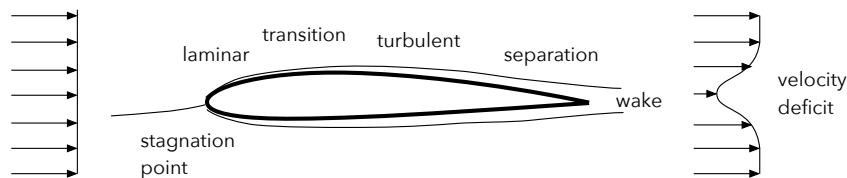


Figure 3.1: Depiction of a boundary layer in a high Reynolds number flow. Skin friction drag is caused by the interaction between the fluid and the body and pressure drag is caused by the momentum deficit in the wake.

Figure 3.2 compares a laminar and turbulent boundary layer. Some

important observations include: 1) the turbulent boundary layer is fuller closer to the wall (i.e., has higher speeds closer to the wall), and 2) the turbulent boundary layer is larger (i.e., it returns to freestream speed farther away from the wall).

The first observation means that a turbulent boundary layer has a higher wall shear stress, and thus skin friction drag. We can see this by recalling the definition of shear stress at the wall:

$$\tau_w = \mu \frac{du}{dy} \Big|_{y=0} \quad (3.2)$$

The turbulent boundary layer has a larger velocity gradient at the wall, thus higher shear stress.

Higher speeds near the wall isn't necessarily all bad. While it does lead to higher skin friction drag, the increased momentum is able to withstand an adverse pressure gradient for longer, and thus delay separation and decrease the wake size. For a blunt body, pressure drag is usually dominant and so a turbulent boundary layer will often produce less total drag. This is why, at least in part, a golf ball has dimples. The dimpled shape forces turbulence allowing the ball to fly further when hit. For a streamlined shape, like an airfoil, skin friction drag is usually the bigger contributor and so sustaining a substantial section of laminar flow is desirable. However, too much is not good either. A fully laminar section is prone to separation and premature stall.

The skin friction drag, per unit length, is the integral of the shear stress across the surface:

$$D' = \int_0^L \tau_w dx \quad (3.3)$$

This drag is normalized, in the same manner as the drag coefficient, and for this special case it is given the name skin friction coefficient:

$$C_f = \frac{D'}{q_\infty L} = \frac{1}{q_\infty} \int_0^1 \tau_w d\left(\frac{x}{L}\right) \quad (3.4)$$

For a laminar boundary layer the skin friction coefficient on a flat plate with no pressure gradient can be solved analytically (known as the Blasius solution, and something you would study in a more advanced fluids class):

$$C_f = \frac{1.328}{\sqrt{Re}} \text{ (laminar)} \quad (3.5)$$

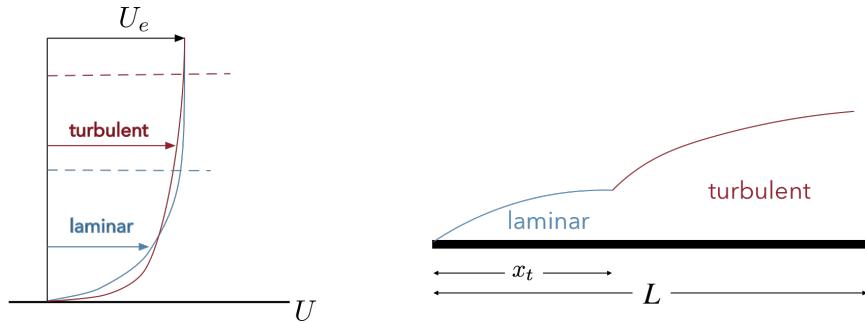


Figure 3.2: A comparison between laminar and turbulent boundary layer profiles. The edge velocity U_e is the velocity outside of the boundary layer, and x_t is the distance from the leading edge until transition.

For a turbulent boundary layer no analytic solution is possible. Various empirical fits exist, below is one of the simpler versions in wide use [1]:

$$C_f = \frac{0.074}{Re^{0.2}} \text{ (turbulent)} \quad (3.6)$$

Note that in both of these formulas the Reynolds number is based on the total length of the boundary layer.

For high speeds, a Mach number correction is needed. At high speeds the skin friction coefficient decreases (though not the skin friction drag because the dynamic pressure increases). The skin friction coefficient relative to the above incompressible formulas is given by [2]:

$$\frac{C_f}{C_{f,inc}} = (1 + 0.144M^2)^{-0.65} \quad (3.7)$$

This relationship is plotted in Fig. 3.3.

An airfoil is not a flat plate, but we can reasonably use the same approach using the length of the airfoil surface. An airfoil is generally not fully laminar or fully turbulent although we sometimes approximate it the way for conceptual design. A more accurate approach to computing the skin friction coefficient on an airfoil with both laminar and turbulent flow is discussed in Section A.1. The user must keep in mind that the above formulas are based on a smooth flat plate. Surface roughness can add considerable drag and so additional markup factors are necessary to account for roughness.

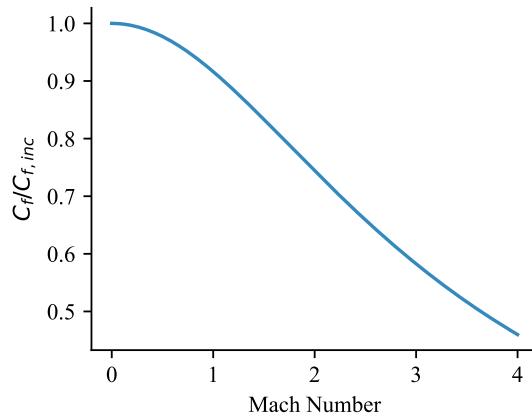


Figure 3.3: Reduction in skin friction coefficient at high Mach numbers.

3.2.2 Pressure Drag

Ideally an aircraft flies through though initially undisturbed air, and after it passes, the air is again at rest. Thus, no net energy was transferred to the air. Unfortunately, perfect efficiency is not possible, and a wake will be left behind. Figure 3.1 depicts this velocity deficit, which means that there is a pressure imbalance and thus a resultant drag.

A streamlined shape will generally have much less drag than a blunt shape. For example, both the airfoil and the cylinder shown in Fig. 3.4 have approximately the same drag for Reynolds numbers of low-speed aircraft [3]. This surprising result is why many early biplanes had such poor performance. The wire bracing, although seemingly small, created a large amount of drag. This is also why performance cyclists go to great lengths to use streamlined spokes, helmets, frames, etc.

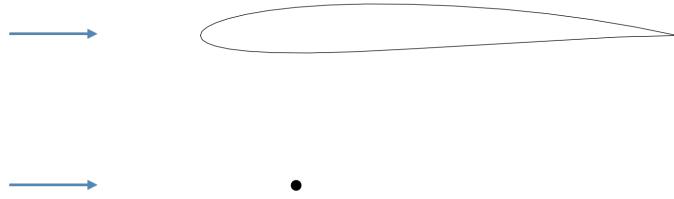


Figure 3.4: The larger airfoil and the small cylinder have approximately the same drag.

In general, predicting pressure drag is complex and requires solving the

flow field around an object. For our purposes we will use a simplified semi-empirical model for lifting surfaces (e.g., wings, tails) and bodies of revolution (e.g., fuselages, pods, nose cones). A flat plate, at zero lift, has no pressure drag. Adding thickness to an object will create pressure drag, and in this simple method we will multiply the skin friction coefficient by a form factor that accounts for the increase in drag due to thickness.

For a lifting surface many semi-empirical methods exist for the form factor (k) [4]. In this text we use Shevell's formula [5]:

$$k = 1 + Z \frac{t}{c} + 100 \left(\frac{t}{c} \right)^4 \quad (3.8)$$

where

$$Z = \frac{(2 - M_\infty^2) \cos \Lambda}{\sqrt{1 - M_\infty^2 \cos^2 \Lambda}} \quad (3.9)$$

In these formulas Λ is the quarter-chord sweep angle and t/c is the thickness-to-chord ratio. For incompressible flow $Z = 2 \cos \Lambda$. The wing form factor is depicted in Fig. 3.5 for three different sweep angles at $M_\infty = 0$.

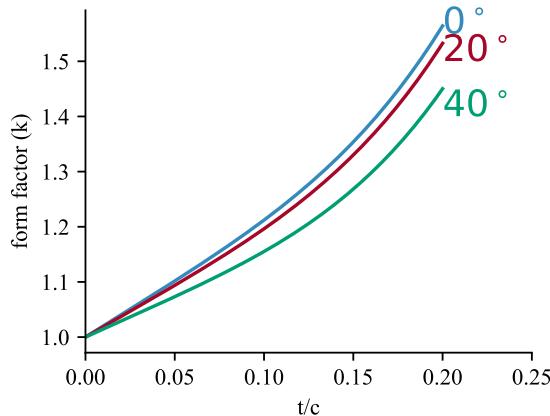


Figure 3.5: Form factor for a lifting surface as a function of thickness-to-chord ratio.

For a body of revolution (e.g., fuselage) various form factor methods also exist, though all are based on the fineness ratio (fr). The fineness ratio is defined as the length over the diameter $fr = l/d$ for the body of revolution as seen in Fig. 3.6a. Shevell provides a figure for form factor versus fineness

ratio, but not an explicit formula [5]. A naive fit to the data would lead to form factors less than 1 for large fineness ratios, which is not physically consistent. For this text we have developed the following expression using a simple quadratic fit, but with the constraint that it bottoms out at 1.

$$k = \begin{cases} 1.675 - 0.09fr + 0.003fr^2 & 5 < fr < 15 \\ 1 & fr \geq 15 \end{cases} \quad (3.10)$$

This form factor is seen in Fig. 3.6b. Note that this method is not valid for fineness ratios smaller than 5. This formula can also be used for non-circular cross sections. In that case one defines an effective diameter based on the maximum cross sectional area (S_{max}) of the body of revolution:

$$d_{eff} = \sqrt{\frac{4S_{max}}{\pi}} \quad (3.11)$$

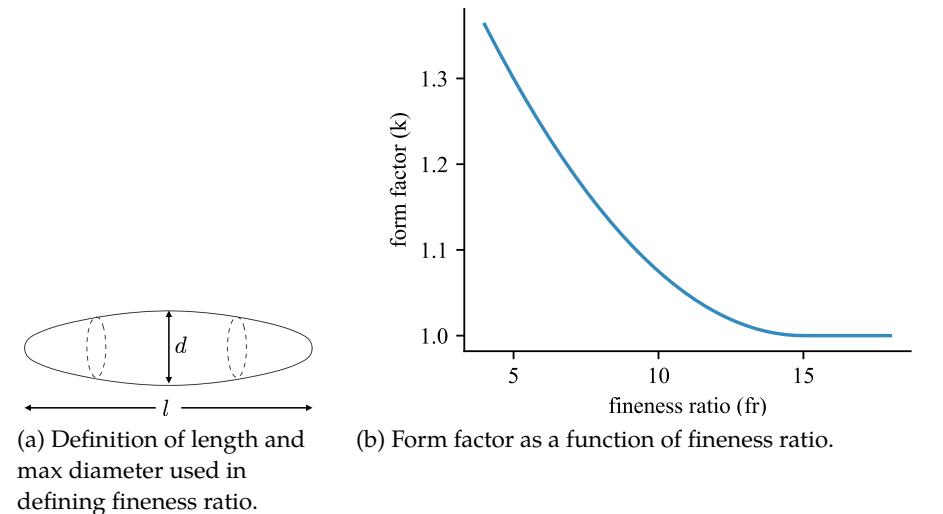


Figure 3.6: A body of revolution.

3.2.3 Putting it Together

We can now put the skin friction and pressure drag together to form parasitic drag. The total parasitic drag is:

$$D_p = kC_f q_\infty S_{wet} \quad (3.12)$$

The first factor accounts for the pressure drag, the second for the skin friction drag, and the last two are from the normalization of the skin friction coefficient. The wetted area S_{wet} is depicted in Fig. 3.7. The parasitic drag acts only on the portions of the body that are exposed to the fluid. It is called a wetted area because it is the part of the vehicle that would be wet if dipped in a fluid.

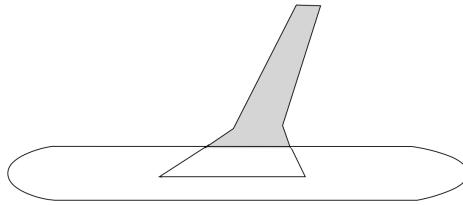


Figure 3.7: Wetted area is the area of that component that would get wet if dipped in a fluid. In this picture a portion of the wing carries through the fuselage (i.e., it is inside the fuselage and wouldn't "get wet".)

We can estimate wetted area for a lifting surface as:

$$S_{wet} \approx 2 \left(1 + 0.2 \frac{t}{c}\right) S_{exposed} \quad (3.13)$$

The factor of 2 accounts for both sides of the exposed planform area ($S_{exposed}$), and the inclusion of the thickness to chord ratio (t/c) is to approximate the additional area from curvature.

For a cylindrical section of a fuselage the area is simply

$$S_{wet} = \pi d l \quad (3.14)$$

For a rounded cylinder, like a nose cone, we can approximate the wetted area as:

$$S_{wet} = 0.75 \pi d l \quad (3.15)$$

The last step is to normalize the drag to form a parasitic drag coefficient. Using the standard definition of the drag coefficient:

$$C_D = \frac{D}{q_\infty S_{ref}} \quad (3.16)$$

where S_{ref} is some reference area for the aircraft that in general is different from the wetted area of the particular component. We need to normalize

by the same area for all components so that drag coefficients can be added together. The parasitic drag coefficient is then:

$$C_{Dp} = k C_f \frac{S_{wet}}{S_{ref}} \quad (3.17)$$

Additional sources of parasitic drag need to be added to this initial estimate. First, the method discussed estimates the parasitic drag of each component in isolation, but when assembled in an aircraft additional interference drag is created. It is possible that the interference could be beneficial, but it is almost always detrimental with a typical increase in drag of 3–8% [2]. Second, we have assumed ideal lifting surfaces and axisymmetric sections, but real aircraft require vary protuberances and hinges that create additional drag. For a jet transports these protuberances add about 2–5% additional drag, whereas a propeller-driven aircraft may have 5–10% additional drag [2]. A small RC aircraft is likely to have at least a 10% drag markup from protuberances and could have much more if not carefully designed. Other sources of additional parasitic drag include nacelle base drag, fuselage upsweep drag, control surface gap drag, etc.

3.2.4 Strip Theory

The proceeding method is easy to use, and appropriate in the early stages of design, but it is oversimplified for many configurations and does not address some fundamental considerations like how drag changes with angle of attack. An alternative approach to computing parasitic drag that is a bit more involved, but is generally more accurate, is strip theory. In this approach we subdivide the wing into a bunch of 2D strips, or in other words into airfoils (Fig. 3.8). Using 2D theory we compute the drag produced by each airfoil, and then we integrate across the wing to get the total drag.

For each section we find the angle of attack based on the local freestream and twist, and simply use the drag polar for that airfoil to compute the drag coefficient of that section:

$$c_d = f(\alpha, Re, M) \quad (3.18)$$

The drag polar could come from any valid source including simulation, wind tunnel data, or flight testing. Obviously the better the source the more accurate the drag prediction.

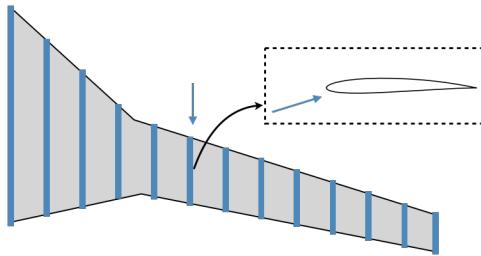


Figure 3.8: Strip theory breaks the wing up into strips (i.e., airfoils), depicted in blue.

One wrinkle, is that the angle of attack we need to use is what is called the effective angle of attack. This is the geometric angle of attack (what we normally think of as angle of attack) minus the induced angle of attack (which will be defined in the next section):

$$\alpha_{eff} = \alpha_g - \alpha_i \quad (3.19)$$

Computing the induced angle of attack generally requires a numerical solution rather than a closed-form solution.

Once we have the drag coefficient for each strip, we can get the total drag coefficient through an integral (or summation):

$$C_{Dp} = \frac{1}{S_{ref}} \int c_d c \quad (3.20)$$

where the above version of the formula assumes that the dynamic pressure is constant along the wing.

This type of approach is used in [XFLR5](#). Its main limitation is that it ignores the three-dimensional nature of the flow across the wing, including crossflow, but for many applications that is reasonable, and the flat-plate based method we introduced shares the same limitation.

3.3 Induced Drag

Induced drag is the additional drag necessary to generate lift. There are many ways to describe the physical mechanism. Perhaps one of the more fundamental is to consider Newton's third law. If lift is generated, that means that the surrounding air is generating a force on the vehicle to push

it upward. By Newton's third law, the vehicle must be creating an equal and opposite force pushing the air downward.

Indeed the air is pushed downward, but it does not continue downward indefinitely. Instead, because the way vorticity is distributed along the wing (discussed in an aerodynamics course) the air circles around as shown in Fig. 3.9a. Generally, this circulation of air forms two counter-rotating spiral vortices. One of these vortices is visualized from an aircraft flying through colored smoke in Fig. 3.9b

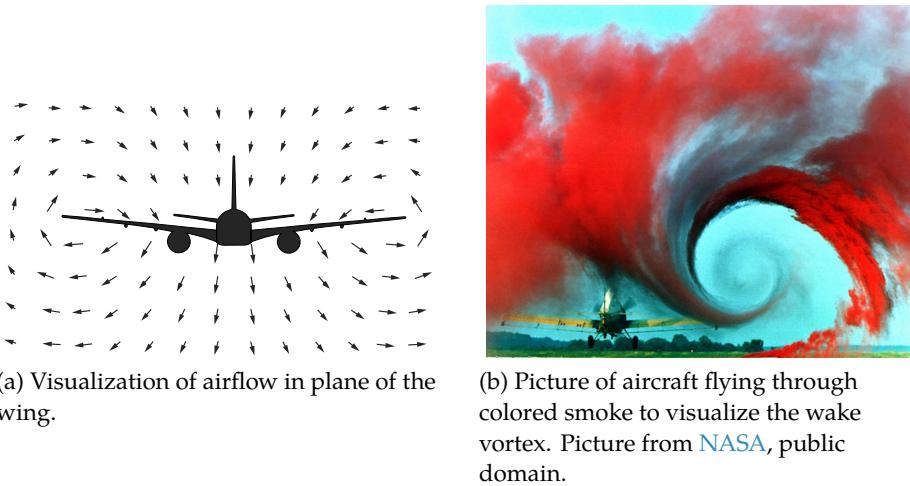


Figure 3.9: Visualization of wake vortex.

If we look at a line between the two vortices the velocity distribution looks like that shown in Fig. 3.10a. There is downward moving air between the two vortices, and upward moving air to the sides. As the vortices trail behind the aircraft in a wake it forms a horseshoe pattern like that shown in Fig. 3.10b. This velocity distribution (partially) explains why many migratory birds fly in a V shape (Fig. 3.11). If another bird flies behind and to the side of another bird then it can fly in a region of rising air and thus the energy it is required to expend is reduced. By flying in this manner the birds can increase their range (distance they can fly).

Many misconceptions about induced drag exist. A common explanation for the tip vortices is that there is high pressure below the wing and low pressure above the wing and so the flow leaks around the edge (Fig. 3.12). This is true, and helps explain why the vortices circle around, but it is

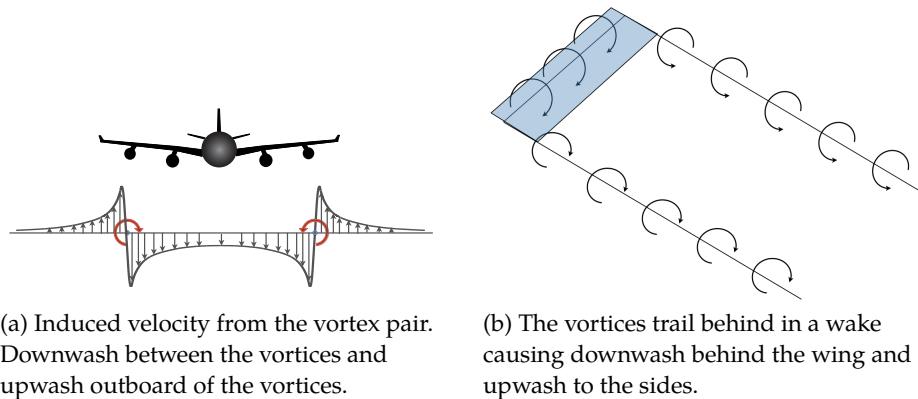


Figure 3.10: Visualization of the downwash/upwash from the trailing vortices behind an aircraft.



Figure 3.11: Migratory birds in a V formation. Picture from [Hamid Haji-husseini](#), CC BY 3.0.

incomplete. This perspective makes the process seem like a tip effect, and has lead to the idea that the aerodynamic purpose of a winglet is to suppress this tip vortex. This is misleading. By this logic a box wing (Fig. 3.13) should have no induced drag, as it has no tip. Lift is more fundamental than just a tip effect, and as pointed out by Newton's third law, to create lift there must be energy left behind in the air. Certainly, vorticity is highest at the tip (and thus the center of vorticity is near there), but the formation of the vortices and the production of induced drag is affected by the entire lifting surface and not just the tip. Winglets modify the lift distribution in such a way that it acts like a span extension. We will see shortly that increasing span decreases induced drag.



Figure 3.12: A simplified view of induced drag. The air leaks around the wing tip from the high pressure to low pressure side to form the vortex.

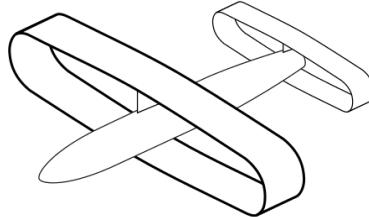


Figure 3.13: Box or annual wing. Image from [Steelpillow](#), Wikimedia, CC BY-SA 3.0.

Another way to look at induced drag is to use the Kutta-Joukowski theorem. This theorem says that the inviscid force per unit length generating by a lifting body is given by:

$$F' = \rho \vec{V} \times \vec{\Gamma} \quad (3.21)$$

where Γ is the circulation. Circulation is defined as a line integral of the velocity around a closed contour:

$$\Gamma = \oint \vec{V} \cdot d\vec{l} \quad (3.22)$$

For the airfoil in Fig. 3.14, if it is generating lift, its circulation will be into the page as shown. Any closed contour, of any shape, taken around the airfoil will give the same value for the circulation. This is an admittedly terse definition of circulation, but the important point for our purposes is that the generated force is given by the cross product defined in the Kutta-Joukowsky theorem, also shown in that same figure.

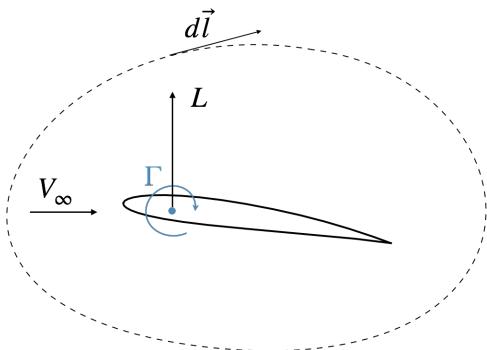


Figure 3.14: Circulation is found by integrating around any closed contour around the lifting body. Lift is given by the cross product of the freestream velocity and the circulation.

As discussed, when a wing generates lift it forces air downward. This downward moving air is called downwash (w) and is shown in Fig. 3.15. The effective incoming velocity is no longer V_∞ but rather V_r . Again, applying the Kutta-Joukowsky theorem we see that the force vector has tilted backwards creating a component of the force in the lift direction and a component in the drag direction (lift and drag are defined relative to V_∞). This component of drag is the induced drag.

The formula for the induced drag, and induced drag coefficient is given by:

$$D_i = \frac{L^2}{q_\infty \pi b^2 e_{inv}} \quad (3.23)$$

$$C_{Di} = \frac{C_L^2}{\pi A R e_{inv}} \quad (3.24)$$

where $AR = b^2/S_{ref}$ is the aspect ratio, and e_{inv} is the inviscid span efficiency that will be discussed in more detail shortly. Notice the strong dependence on lift, and on span. As expected as we increase lift, the induced drag increases. Increasing span allows for a decrease in induced

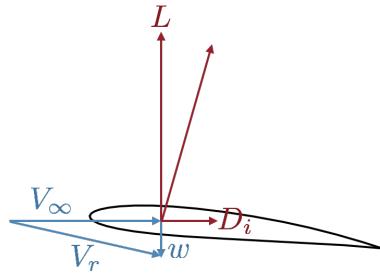


Figure 3.15: Induced drag from the viewpoint of the Kutta-Joukowski Theorem. The presence of downwash (w) tilts the incoming velocity vector and thus the resulting force. The component in the direction of the freestream is induced drag.

drag. This is why high-efficiency aircraft, or birds, have long wing spans although with the tradeoff of generally slower speeds.

It is important to note that this formula gives the *inviscid* induced drag. All the behavior we have discussed so far does not depend on the presence of viscosity. The inviscid span efficiency e_{inv} hides some complicated flow physics behind a single number. Effectively, it is a measure of how close the lift distribution is to the one that produces minimum induced drag. For a planar wing, which means a wing whose planform is contained in one geometric plane (e.g., no winglets), *the lift distribution for a fixed span planar wing with minimum induced drag is an elliptic shape* as seen in Fig. 3.16.

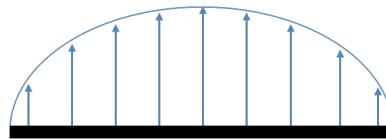


Figure 3.16: An elliptic distribution of lift.

Most aircraft do not have an exact elliptic lift distribution, but will often be fairly close. An elliptic distribution produces an inviscid span efficiency of 1.0 and a typical value for a wing is 0.98. The presence of the fuselage tends to decrease the lift produced at the center as shown in Fig. 3.17. For a wing/fuselage combination the inviscid span efficiency can be estimated as:

$$e_{inv} = 0.98 \left[1 - 2 \left(\frac{d_f}{b} \right)^2 \right] \quad (3.25)$$

where d_f is the diameter of the fuselage, and b the wingspan (a simplification of the formula in [6]).

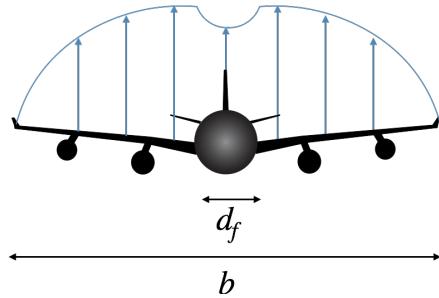


Figure 3.17: The presence of the fuselage causes a loss in lift production near the centerline.

The presence of lift also tends to increase the viscous component of drag. While this could potentially be book-kept with parasitic drag, we will include it with induced drag. Empirical data suggests that viscous induced drag also increases proportionally with the square of the lift. We can estimate the viscous induced drag coefficient as:

$$C_{D_i,v} = KC_{D_p} C_L^2 \quad (3.26)$$

Note that quadratic dependence on lift. We also assume that it is proportional to the parasitic drag. The remaining factor, K , is an empirical constant determined to be approximately $K = 0.38$ [5]. We can combine the inviscid and viscous components as follows:

$$\begin{aligned} C_{D_i} &= \frac{C_L^2}{\pi A R e_{inv}} + KC_{D_p} C_L^2 \\ C_{D_i} &= \frac{C_L^2}{\pi A R} \left(\frac{1}{e_{inv}} + KC_{D_p} \pi A R \right) \end{aligned} \quad (3.27)$$

We would like to return our equation to a simple form like before, so we define everything in parenthesis as $1/e$ where e is a new term called the Oswald efficiency factor. Now the induced drag coefficient is given by:

$$C_{D_i} = \frac{C_L^2}{\pi A Re} \quad (3.28)$$

where

$$e = \frac{1}{\frac{1}{e_{inv}} + KC_{D_p} \pi A R} \quad (3.29)$$

Notice that our total induced drag coefficient equation looks the same as the inviscid induced drag coefficient (Eq. (3.24)), except we have replaced e_{inv} with e . Whereas a typical value for the inviscid span efficiency (e_{inv}) is close to 1, a typical value for the Oswald efficiency factor (e) is between 0.7–0.9. The induced drag in Eq. (3.23) transforms similarly with e replacing e_{inv} .

The previous section mentioned the induced angle of attack, which we will touch on briefly. The presence of downwash changes the direction of the incoming velocity as explained earlier and shown in Fig. 3.15. The induced angle of attack is the angle created between the freestream and the new relative velocity as shown in Fig. 3.18.

$$\alpha_i = \tan^{-1} \left(\frac{-w}{V_\infty} \right) \quad (3.30)$$

On the right side of the figure we see that the induced angle of attack reduces the angle of attack experienced by the airfoil:

$$\alpha_{eff} = \alpha_g - \alpha_i \quad (3.31)$$

where α_g is the geometric angle of attack. Thus, the presence of downwash decreases the angle of attack experienced by each section.

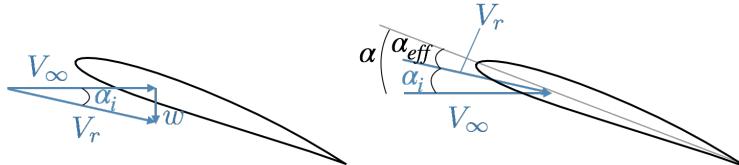


Figure 3.18: The induced angle of attack reduces the effective angle of attack of the section. Angle of attack of airfoil exaggerated for clarity of display.

For an elliptic lift distribution, the downwash is constant along the wing and the induced angle of attack is given (with a small angle approximation) by:

$$\alpha_i = \frac{C_L}{\pi AR} \quad (3.32)$$

In general, the downwash, and thus the induced angle of attack, varies along the wing and must be determined through a numerical solution such as lifting line theory or a vortex lattice method discussed in an aerodynamics course.

3.4 Tradeoffs in Parasitic and Induced Drag

For low Mach number (i.e., no compressibility drag) the total drag coefficient or total drag of an aircraft is given by:

$$C_D = C_{Dp} + \frac{C_L^2}{\pi A Re} \quad (3.33)$$

or

$$D = C_{Dp} q_\infty S_{ref} + \frac{L^2}{q_\infty \pi b^2 e} \quad (3.34)$$

At low speeds the drag is dominated by induced drag (see q_∞ in denominator), whereas at high speeds the drag is dominated by parasitic drag (see q_∞ next to C_{Dp}). The parasitic, induced, and total drag as a function of speed, for an example aircraft, is shown in Fig. 3.19. The point at which they cross, is the speed for minimum total drag, or in other words the speed that maximizes the lift-to-drag ratio.

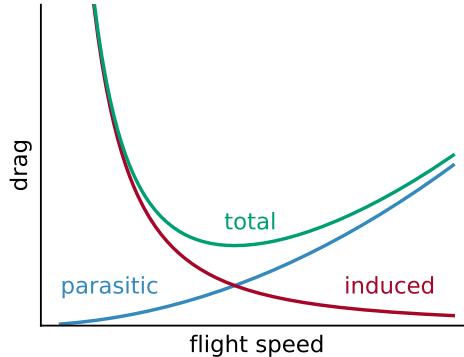


Figure 3.19: Tradeoff in parasitic and induced drag versus flight speed.

3.5 Compressibility Drag

Before discussing methods to estimate compressibility drag we need to understand the physical mechanism, and so we begin with sound waves. Disturbances in a fluid are communicated via sound waves—infinitesimally weak pressure waves that travel at the speed of

sound. A familiar example is a river flowing around a rock. The water “senses” the presence of the rock and starts diverting its path before it reaches the rock. The mechanism that allows the water to sense the rock is sound waves, or pressure waves, transmitted through the water. The increased pressure starts to divert the water well before it reaches the rock. Let’s reduce our rock to a point and rather than have the fluid move around the rock, let’s have the point move through the fluid. Figure 3.20 shows four scenarios with the point moving to the right. In the first scenario the point is not moving. Pressure waves travel in all directions at equal speed (the speed of sound). In the second scenario the point is moving to the right at subsonic speeds. The pressure waves in the direction of travel bunch up, and the separation between the pressure waves behind the point elongates. In the third scenario the point is moving at the speed of sound. The pressure waves coalesce at the front as they are all moving at exactly the same speed as the point. In the last scenario the rock is moving faster than the speed of sound. In one instant of time the point sends off a pressure wave, but in the next instant of time the rock has traveled past that sound wave. The envelope of pressure waves forms a cone, known as a Mach cone. See [this webpage](#) for an animated version of this concept.

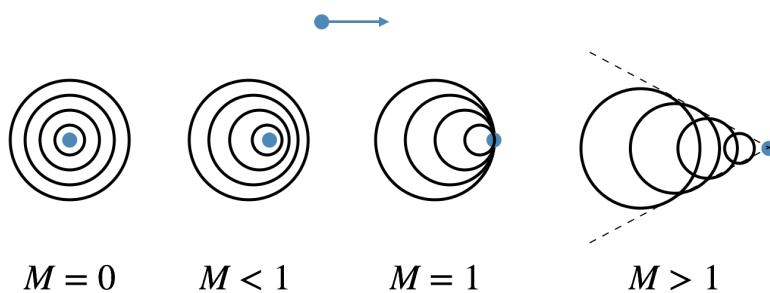


Figure 3.20: Sound waves emanating from a point moving to the right for four different Mach numbers.

Now let’s revisit our rock in the river example, but with an airplane and a small region of air. The following is only an analogy as fluid properties vary throughout the flow, but illustrates the general principle. If the airplane is flying subsonically the air will “sense” that the airplane is coming because of the pressure waves that emanate ahead of the aircraft. It will start moving ahead of time to flow around the airplane, like the river around the rock.

Now let's assume the airplane is flying supersonically. An air particle that is in front of the airplane won't know the airplane is approaching until it's already there. It can't know because the airplane is flying faster than the pressure waves. This is a problem because the air needs to move around the aircraft, but won't be able to until it's too late. Nature's solution is a shock wave.

A *shock wave* is a narrow region that causes an instantaneous change in flow conditions.¹ In an instant the pressure, density, and temperature all increase, the Mach number decreases, and the flow direction changes (if the flow is perpendicular to the shockwave then the flow direction does not change).

In the prior example the point was infinitely small and so it created an infinitely weak disturbance. The resulting wave front is called a *Mach wave*. For a finite disturbance the coalescence of sound waves creates a *shock wave*, or for an aircraft multiple shock waves (see Fig. 3.21).

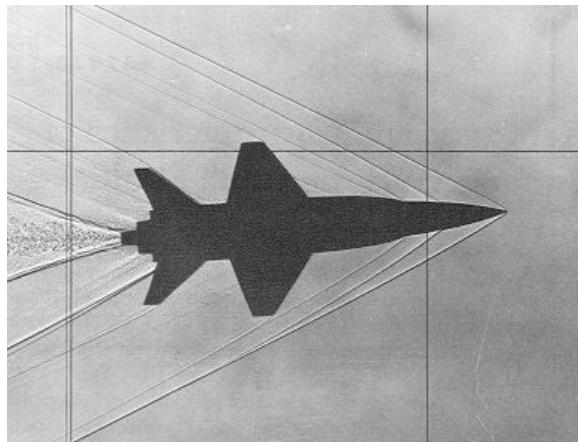


Figure 3.21: Shock waves around an aircraft in a wind tunnel visualized using [Schlieren photography](#). The Schlieren method takes advantage of the bending of light through density changes, and a shockwave produces a near instantaneous change in density. Image from [NASA](#), public domain.

A course in compressible flow will discuss shockwave theory in more

¹It's not actually instantaneous. The shock wave is like a boundary layer—a very thin region dominated by viscous effects. It's so thin that to the external flow it is effectively instantaneous. Note that even though shock waves have a viscous structure, they will still be correctly predicted in an inviscid solution.

detail. For aircraft design our main interest in shock waves is that they carry energy away and thus increase our drag. The stronger the shock wave the more the drag increases. Shock strength increases with blunter shapes and with increasing Mach number.

The following subsections discuss methods to estimate wave drag for transonic speeds and supersonic speeds separately. Recall that compressibility drag is often referred to as wave drag.

3.5.1 Transonic Wave Drag

Transonic mean Mach numbers close to 1. Usually transonic refers only to subsonic speeds, but sometimes transonic is used to mean both subsonic and supersonic speeds. In our case we mean the former, with a typical range of Mach 0.7–1.

Even though the freestream speed is subsonic, as air is accelerated around the aircraft regions of supersonic flow can occur. Figure 3.22 illustrates flow around an airfoil at three different Mach numbers. The first airfoil illustrates what is called the critical Mach number, or the Mach number where the maximum velocity on the aircraft is sonic. This flight speed is the point just before supersonic flow will appear. In aircraft design we usually care more about the *crest critical Mach number* (M_{cc}). This is the Mach number where the crest (top) of the wing first becomes sonic.

After passing the critical Mach number regions of supersonic flow exist on the airfoil, highlighted in dark blue in the figure. As the freestream Mach number increases the shock strength increases leading to higher compressibility drag. Viscous pressure drag may also increase. This is because speed drops significantly across the shock wave, thus making the boundary layer more likely to separate.

The formation of shock waves leads to a precipitous increase in drag as Mach 1 is approached. The rise is so sharp that it is often referred to as a “sound barrier” (Fig. 3.23). Not a barrier in a sense of impenetrability, but a barrier in the sense that a tremendous amount of energy is required to pass through. The Concorde, for example, had to use *afterburners* to pass through the sound barrier. Note that drag drops after Mach 1, though it does start increasing again (not shown on the plot).

One of the major keys to unlocking efficient transonic flight was the discovery by Adolf Busemann, and later independently by R. T. Jones, that a swept-back wing decreased transonic wave drag. In fact, this discovery is

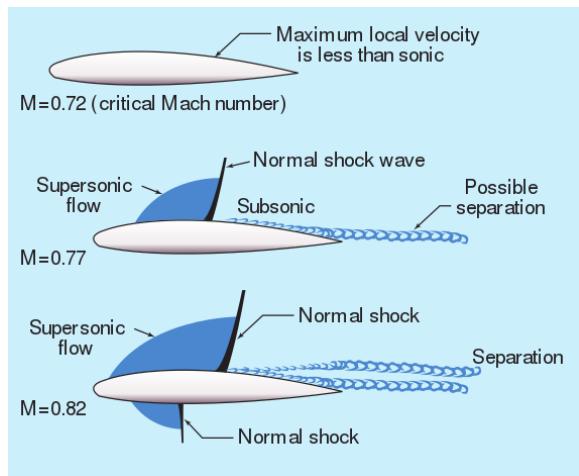


Figure 3.22: Visualization of regions of supersonic flow over an airfoil traveling at transonic speeds. Image from [FAA](#), Airplane Flying Handbook, public domain.

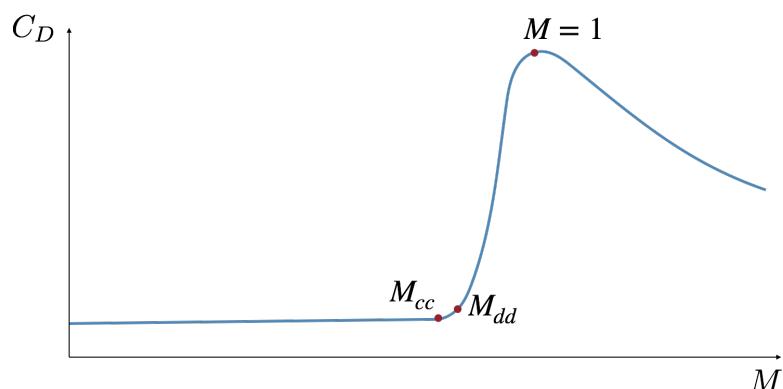


Figure 3.23: Notional plot of the steep compressible drag rise, or “sound barrier” prior to Mach 1. The crest critical Mach number and drag divergence Mach number are also shown.

considered one of the most important discoveries in aerodynamic history.² The concept can be motivated through a simple thought experiment. Imagine an infinite wing of constant cross-section in an inviscid flow. If the wing started moving sideways at constant speed, and the flow were inviscid, nothing would change. In a coordinate system moving at the wing speed, only the normal component of the inflow would exist. In other words, the normal and tangential components of velocity are independent. By adding a tangential component we have exactly recreated a swept wing flying obliquely (Fig. 3.24).

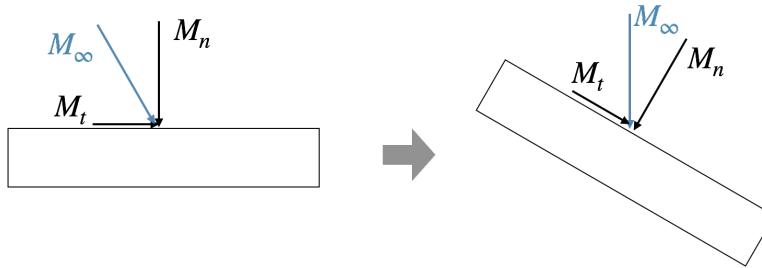


Figure 3.24: An unswept wing, on the left side, with normal (M_n) and tangential (M_t) components of Mach number, is exactly the same as a swept wing flying obliquely as shown on the right side.

In other words, as only the normal component of velocity matters for the straight wing, so too it is only the normal component of velocity that matters for the swept wing (Fig. 3.25). The implication of this idea is that sweeping the wing allows the airfoils to see a reduced effective Mach number. For example, if the flight Mach number was 0.8, and the sweep was 30 degrees, then the airfoils (normal to the sweep) could be designed to operate as if the incoming flow was $M = 0.8 \cos 30^\circ = 0.69$. That is a large reduction and is what has enabled transport aircraft to fly at much higher speeds. This concept is called *simple sweep theory* as it is a simplification of the real flow because it assumes two-dimensional flow behavior (that's what an finite wing implies), that all sections are identical, and assumes inviscid flow (viscosity can be important for the crossflow boundary layer development). Despite the simplifying assumptions, in most cases it approximates the real flow behavior quite well.

A forward swept wing produces the same transonic drag reduction benefits. However, forward sweep is aeroelastically unstable. For a

²<https://www.nasa.gov/langley/hall-of-honor/robert-t-jones>

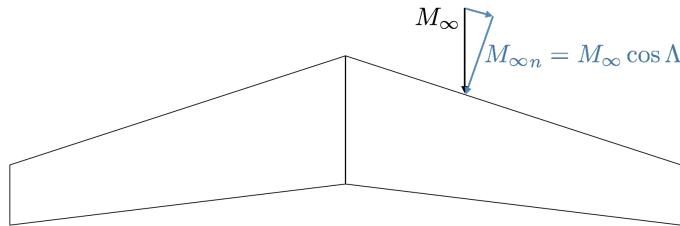


Figure 3.25: For a swept wing only the normal component of the velocity matters according to simple sweep theory.

forward swept wing an increase in lift at the tips typically causes them to twist up, which further increases lift creating a positive feedback loop leading to divergence. For a backward swept wing increased lift at the tip decreases the twist, leading to decreased lift and thus a stable configuration. This instability is one of the main reasons why forward sweep is much less common.

For transonic speeds we generally want to fly as fast as possible without incurring a significant drag rise penalty. In other words, most transport aircraft fly at the point just as the curve in Fig. 3.23 starts to turn upward. This Mach number is called the *drag divergence Mach number* (M_{dd}) and is labeled in that figure. The drag divergence Mach number can be estimated using a modification of the Korn equation with simple sweep theory [7]:

$$M_{dd} = \frac{0.95}{\cos \Lambda} - \frac{t/c}{\cos^2 \Lambda} - \frac{C_L}{10 \cos^3 \Lambda} \quad (3.35)$$

Notice that sweep plays a major factor, as expected, as does thickness and lift.

The compressibility drag coefficient can then be estimated using the following empirical relationship [8]:

$$C_{Dc} = \begin{cases} 0 & \text{for } M_\infty < M_{cc} \\ 20(M_\infty - M_{cc})^4 & \text{for } M_{cc} < M_\infty < 1 \end{cases} \quad (3.36)$$

Note the rapid rise in drag (quartic) with Mach numbers past M_{cc} . For Mach numbers less than M_{cc} there are no shock waves and thus no wave drag.

To complete the set of equations we need a way to relate M_{dd} (when drag starts to rise rapidly) to M_{cc} (when there is no compressibility drag).

Various definitions for the drag divergence Mach number exist. In the

following we use the definition that drag divergence occurs once the slope of the drag curve equals 0.1:

$$\frac{dC_D}{dM} = 0.1 \quad (3.37)$$

If we take the derivative of Eq. (3.36) and use the above definition of M_{dd} we find the following relationship:

$$M_{cc} = M_{dd} - 0.11 \quad (3.38)$$

3.5.2 Supersonic Wave Drag

For supersonic flight the presence of shock waves is unavoidable. To minimize drag, we generally want to have oblique shocks and this requires a sharp leading edge and slender shapes as depicted in Fig. 3.26 or as seen in a concept supersonic aircraft (Fig. 3.27).

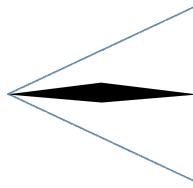


Figure 3.26: Depiction of an oblique shock around a sharp-nosed airfoil.

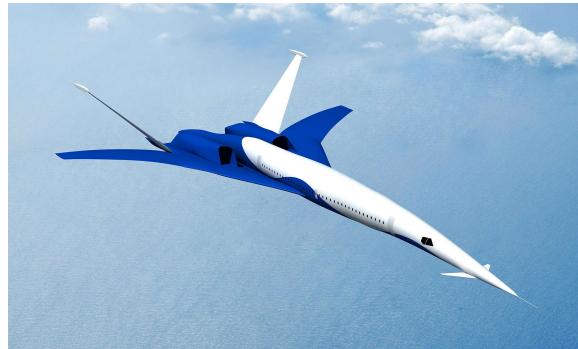


Figure 3.27: Supersonic transport concept showing sharp leading edges and slender shapes. Image from [NASA](#), Boeing concept supersonic transport, Icon-II, public domain.

The minimum wave drag for a wing can be broken up into two parts: one due to volume and one due to lift. The derivation assumes that the lift distribution is elliptic in both directions. In other words, this is the minimum drag and so for a given design the drag could be much higher. The formulas are derived by R. T. Jones [9], and are manipulated in Section A.2 to a form consistent with our usage. The results are shown below.

The wave drag of a wing due to volume is given by:

$$C_{Dc,v} = 4S_r \left(\frac{t}{c}\right)^2 \frac{(\beta^2 + 2r^2)}{(\beta^2 + r^2)^{1.5}} \quad (3.39)$$

where

$$\beta = \sqrt{M^2 - 1}, \quad (3.40)$$

$$r = \frac{\pi l^2}{4S_g} \quad (3.41)$$

and $S_r = S_g/S_{ref}$ is the ratio of the gross area to the reference area used in the drag normalization. The wave drag of a wing due to lift is given by:

$$C_{Dc,l} = \frac{C_L^2 r}{4S_r} \left[\sqrt{1 + \left(\frac{\beta S_r}{r} \right)^2} - 1 \right] \frac{S_{ref}}{S_g} \quad (3.42)$$

The wave drag of a fuselage due to volume is given by:

$$C_{Dc,v} = \frac{\pi^3}{4} \frac{d^2}{(l/d)^2 S_{ref}} \quad (3.43)$$

where l and d are the length and maximum diameter of the fuselage respectively.

In general, we can show that the volume dependent wave drag at supersonic speeds is proportional to volume squared over length to the fourth power:

$$D_w \propto \frac{volume^2}{l^4} \quad (3.44)$$

In other words, for low drag, a supersonic transport should be long and have as little volume as possible. The elongation effect is reflected in many designs, and is seen in the concept plane shown earlier (Fig. 3.27).

However, having low volume alone is not sufficient for low drag. The way that volume is distributed also matters. R. T. Whitcomb [made the important discovery](#) that reducing the volume of the fuselage near the wing led to a large reduction in drag [10]. This “coke-bottle” shape is seen in many supersonic aircraft such as the redesigned version of the Delta Dart (Fig. 3.28).



Figure 3.28: [NASA](#) QF-106 Delta Dart, NASA, public domain.

This idea became known as the *area rule* and enabled new high speeds for aircraft. A more detailed look at the theory shows that the second derivative of the area distribution appears in the drag calculation. This means that the area distribution should change smoothly—sharp changes produce high second derivatives in area and thus high drag. By decreasing the fuselage area, just as the wing area is increasing, a smoother cross-sectional area distribution can be maintained. The area rule is more complex than described. The distribution of cross-sectional areas perpendicular to the body (Fig. 3.29) only applies at Mach 1. In general, the cross-sectional areas are taken at the angle of the Mach wave and are thus speed dependent. Because the aircraft will generally be flown through many speeds, this means that the area distribution needs to be optimized for multiple conditions. Additionally,

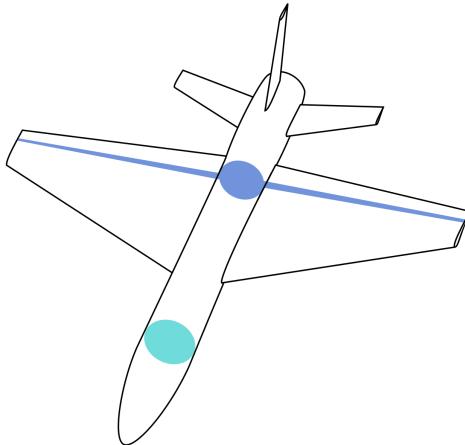


Figure 3.29: Cross sectional areas, as used in the area rule. Image by Petr Dlouhý, Wikimedia Commons, public domain.

cuts need to be taken for geometric planes spanning all 360 degrees around the aircraft.

3.5.3 High Mach or Blunt Bodies

As the Mach number is increased, or if the geometry is blunt rather than sharp, a detached bow shock will form (Fig. 3.30). A bow shock produces much higher drag than the oblique shocks discussed in the previous section. For aircraft this is generally undesirable, but for a reentry vehicle high drag is desirable. The high drag slows the vehicle down, and reduces the heat transfer to the vehicle. The temperature increase behind a bow shock is larger than that of an oblique shock, but because the shock detaches, the separation distance more than makes up for the increased temperature leading to overall lower heating.

Bibliography

- [1] Schlichting, H. and Gersten, K., *Boundary-Layer Theory*, Springer Science & Business Media, May 2003.
- [2] Raymer, D. P., *Aircraft Design: A Conceptual Approach*, American Institute of Aeronautics and Astronautics, 2012.

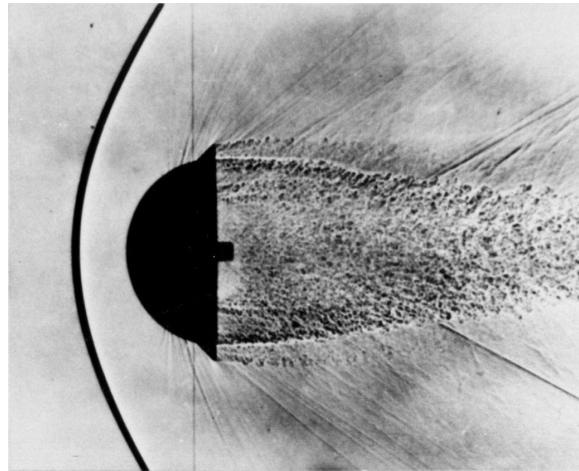


Figure 3.30: Detached bow shock in front of a blunt reentry body. Image by [NASA, public domain](#).

- [3] Talay, T. A., "Introduction to the Aerodynamics of Flight," NASA SP-367, National Aeronautics and Space Administration, Washington, D.C., 1975.
- [4] Takahashi, T. J., German, B. J., Shajanian, A., Daskilewicz, M. J., and Donovan, S., "Form Factor and Critical Mach Number Estimation for Finite Wings," *Journal of Aircraft*, Vol. 49, No. 1, jan 2012, pp. 173–182.
- [5] Shevell, R. S., *Fundamentals of Flight*, Pearson College Division, 1989.
- [6] Kroo, I., *Aircraft Design: Synthesis and Analysis*, Stanford University, 2006.
- [7] Malone, B. and Mason, W. H., "Multidisciplinary optimization in aircraft design using analytic technology models," *Journal of Aircraft*, Vol. 32, No. 2, mar 1995, pp. 431–438.
- [8] Hilton, W. F., *High-speed Aerodynamics*, 1951.
- [9] Jones, R. T., *Wing Theory*, Princeton University Press, Jul 2014.
- [10] Whitcomb, R. T., "A Study of the Zero-Lift Drag-Rise Characteristics of Wing-Body Combinations Near the Speed of Sound," NACA-TR-1273, National Advisory Committee for Aeronautics, Jan 1956.

CHAPTER 4

Wing Design

Wing design is at the heart of aircraft design. The wing (or sometimes wings) provides almost all of the lift of the aircraft, and its dimensions significantly affect the speed and performance of the aircraft. Appropriately designing the wing is critical.

4.1 Fundamental Parameters

One of the primary design parameters for a wing is its span. Its definition is usually straightforward (Fig. 4.1). Span refers to the projected span as shown in the figure. In other words, a vertical winglet does not increase span. Despite its apparent simplicity there are cases where the definition of span is ambiguous. Many modern aircraft have highly flexible wings and so the span will change depending on the flight conditions and/or fuel loading. In this case we would also need to specify the loading condition, for example span under a 1-g load (i.e., when lift equals weight). Often the most fundamental tradeoff in selecting span is balancing drag and weight (aerodynamics and structures). An increase in span decreases induced drag as $1/b^2$ as we saw in Chapter 3. Conversely, an increase in span increases bending stresses and thus structural weight. For some cases (e.g., a flying wing) bending stress is not the dominant structural constraint, but other structural constraints like [flutter](#) similarly require

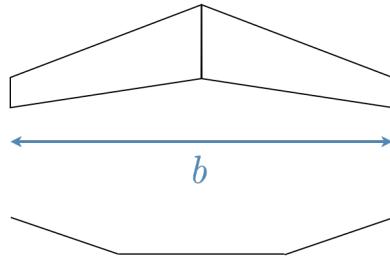


Figure 4.1: Depiction of wing span. The bottom shows a back view of the wing to demonstrate that span refers to the projected span, or the span projected onto a horizontal plane.

significant weight increases with increasing span.

If the wing area is fixed then increasing span decreases the internal volume (e.g., fuel/battery capacity). It also decreases the Reynolds numbers across the wing, which generally leads to increased parasitic drag and higher stall speeds because of reduced maximum lift coefficients.

If the wing is swept back then increasing span moves the aerodynamic center back, which may or may not be desirable. This will have a strong effect on stability as we will see in a later chapter.

Wing area is another fundamental parameter. Many definitions are used (e.g., wetted area, reference area, projected area) depending on the context. One of the most common reasons to increase wing area is to decrease stall speed. The stall speed for level flight is given by:

$$V_s = \sqrt{\frac{2W}{C_{L\max} \rho S_{ref}}} \quad (4.1)$$

However, increasing wing area leads to increased parasitic drag and wing weight. Wing area also affects the optimal cruise lift coefficient, which may or may not be desirable.

Sweep is often defined relative to the quarter chord line (Fig. 4.2), but is sometimes defined relative to the leading edge instead. The primary reason to use sweep is to decrease transonic wave drag as discussed in Section 3.5.1. The other reason to use sweep is as a means to achieve longitudinal stability for a flying wing as will be discussed in Section 5.2. Beyond those two use cases, there is rarely good reason to use sweep as there are many detrimental effects including an increase in structural span with a corresponding increase in weight, an increase in tip loads with a

corresponding increase in weight and a reduced margin for tip stall, and a reduction in C_{Lmax} .

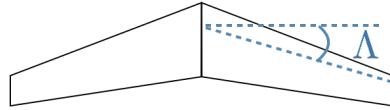


Figure 4.2: Sweep is generally measured about the quarter chord (but sometimes about the leading edge). It may be a distribution rather than a constant value.

Airfoil thickness is also primarily an aerodynamic/structural tradeoff. Thicker airfoils reduce wing weight if your wing is bending constrained (discussed in Section 8.2), usually increase C_{Lmax} (up to a point), and allow for more internal volume. Conversely, thicker airfoils increase parasitic drag and compressibility drag. Thus, aerodynamic considerations usually favor thinner airfoils, while structural considerations usually favor thicker airfoils.



Figure 4.3: Section thickness.

Chord and twist (Fig. 4.4) are strongly coupled with the lift and lift coefficient distributions as will be discussed in the next section. One of the main tradeoffs with chord is that more taper decreases wing weight, but can lead to tip stall. Although we often use taper ratio, modern wings increasingly use composites and have more complex shapes than linear taper.

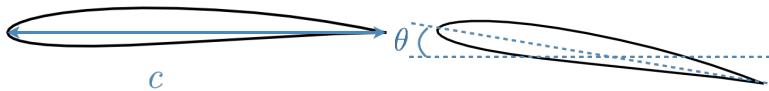


Figure 4.4: Chord and twist.

4.2 Lift Distributions and Lift Coefficient Distributions

As discussed in Section 3.3, an ideal lift distribution for a planar wing, from an aerodynamic point of view, is elliptic. When we consider not just aerodynamics but also structures the optimal lift distribution is generally loaded inboard further as shown in Fig. 4.5. By shifting more of the load inboard the wing weight can be reduced, primarily because of reduced bending loads, while only sacrificing a small amount of induced drag. For the aerostructural optimum a typical inviscid span efficiency is about 0.98 (compared to 1 for elliptic). We tradeoff a small increase in induced drag, for a decrease in weight, such that the overall system has improved.¹

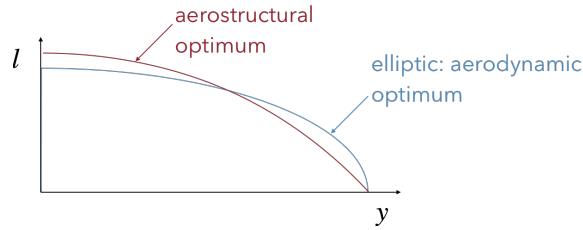


Figure 4.5: The elliptic distribution in blue is the aerodynamic optimum. The aerostructural optimum is loaded further inboard to balance tradeoffs in induced drag and wing weight.

Chord and twist are the biggest factors in determining the shape of the lift distribution with sweep next in importance. If the wing is untwisted, then an elliptic distribution of chord can be used to create an elliptic lift distribution. An example of an aircraft with an elliptic wing is the Supermarine Spitfire (Fig. 4.6). Although an elliptic wing can produce an elliptic load, there is no real advantage to this shape as an elliptic distribution of lift can be created with any wing, properly twisted. Unfortunately, the lift distribution cannot be determined in isolation. The lift coefficient is related to the lift distribution through its definition:

$$c_l = \frac{L'}{q_\infty c} \quad (4.2)$$

¹If interested in more details see this mathematical study by [R. T. Jones](#) or a study I worked on as a Masters student.



Figure 4.6: The Supermarine Spitfire is a well-known example of an aircraft with an elliptic wing. Image from [Adrian Pingstone](#), Wikimedia, public domain.

Note that the lift coefficient distribution is a function of the lift distribution and the chord distribution.

Before discussing an ideal lift coefficient distribution we first need to discuss *critical section theory*. Critical section theory is a simplified approach to predicting when a wing/aircraft will stall. We already know that an airfoil will stall at angles of attack past c_{lmax} , but how does that apply to a wing where angle of attack, c_l , and c_{lmax} can vary along the wing? The basic idea of critical section theory, is that the wing is predicted to stall when any one section would stall using 2D theory. Furthermore, the section that first stalls is predicted to be the location on the wing where stall initiates.

For example, from the lift distribution and Eq. (4.2) we can compute the lift coefficient distribution. From airfoil data we can extract a distribution for c_{lmax} along the span. Now we increase the angle of attack until the c_l curve just touches the c_{lmax} curve as shown in Fig. 4.7. The figure also points out the spanwise position where we expect stall to occur (at about 30% of the span). Critical section theory is a simplification of how stall develops. It works well if there is negligible crossflow across the wing, in other words if the flow is basically two-dimensional at each section.

At first glance it would seem that an ideal lift coefficient distribution

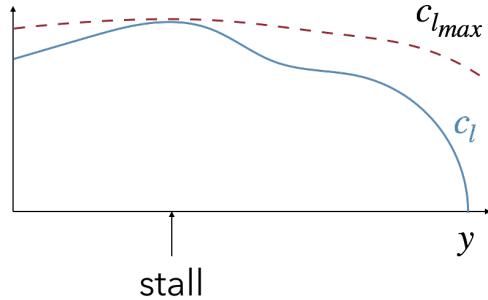


Figure 4.7: An illustration of critical section theory. Stall is predicted to occur at the angle of attack at which these two curves meet, and at the spanwise location where they meet.

would be constant across the span. That would mean that every section was equally contributing, and thus no section would stall prematurely. The problem with this distribution is that it doesn't prevent *tip stall*, which occurs when stall initiates near the wing tips. Tip stall is problematic for two reasons: 1) a loss of lift far from the centerline of the wing creates a large rolling moment that could be dangerous or at least degrade ride quality, and 2) ailerons (the control surfaces that affects roll) are generally located outboard to have maximum effect and if the ailerons stall then roll control is lost. For these reason a good design must ensure that stall occurs inboard first. Thus, an ideal c_l distribution is roughly constant inboard, then drops off towards the tip so that stall occurs inboard first.

Because both the lift and lift coefficient distribution are affected by chord and twist (at least indirectly), and the wing bending stress depends on the lift and chord distribution, creating an ideal balance can be challenging. Chord is perhaps the most impactful parameter. As discussed, by using twist we can create an elliptic lift distribution for any planform shape. However, that planform shape then dictates what the lift coefficient distribution looks like. For example, Fig. 4.8 shows the impact of three different planform shapes on the lift coefficient distribution, assuming an elliptically loaded lift distribution.

Let's examine each of these parameters in isolation. The top portion of Fig. 4.9 shows the impact of changing chord on the lift and lift coefficient distributions. A high taper ratio means that the root and tip chord are relatively close in size. That means that if there is no twist, the lift distribution is relatively uniform. In contrast, a small taper ratio (with no

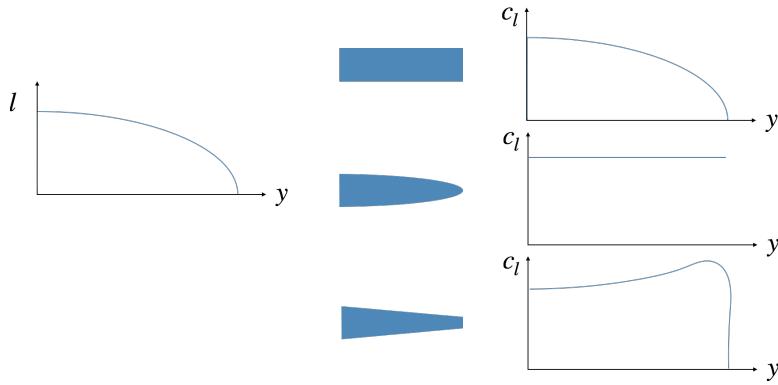


Figure 4.8: The lift coefficient distribution for three different planform shapes given an elliptic lift distribution.

twist) would produce higher lift at the root and lower lift at the tip. Because chord is also coupled with the lift coefficient, some potentially undesirable behavior is observed. For high taper ratio the chord is relatively uniform and so the lift coefficient distribution looks fairly similar to the lift. The low taper ratio on the other hand has small tip chords, and so according to Eq. (4.2) the c_l at the tip will be high. For structural loading a smaller taper ratio is desirable, but we see that it increases the likelihood of tip stall.

The second row of Fig. 4.9 shows the impact of changing twist on an elliptically loaded wing for fixed lift. Increased twist at the tip is called wash-in and decreased twist (i.e., increased negative twist) at the tip is called wash-out. Many wings have some wash-out in order to produce a more structurally efficient lift distribution. By reducing the twist at the tip the lift at the tip is correspondingly reduced. Note that the effect of twist is the same on lift and lift coefficient.

The last row of Fig. 4.9 shows the impact of changing sweep relative to an elliptically loaded wing. Like twist, sweep affects both distributions in the same way. Backward sweep increases the lift and the tip, whereas forward sweep decreases it.

One of the primary challenges of wing design is to create a chord/twist/sweep combination with a desirable lift distribution that balances induced drag and structural considerations, and a lift coefficient distribution that balances efficiency with preventing tip-stall. This is not always easy and like most engineering decisions requires making some

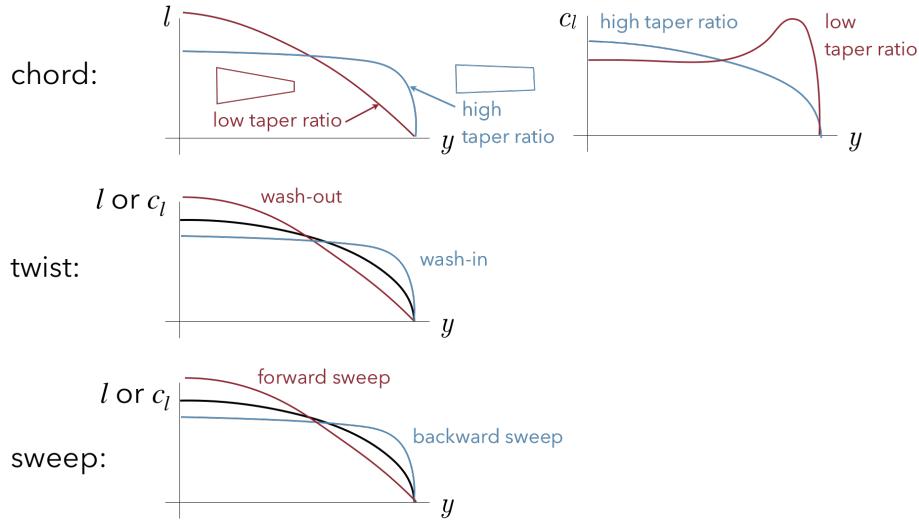


Figure 4.9: Notional depiction of how chord, twist, and sweep affects the lift and lift coefficient distributions.

tradeoffs.

4.3 Finite Wing Lift Curve Slopes

We discussed the differences between 2D and 3D lift coefficients in Chapter 2 and some in this chapter as well, but it is worth reemphasizing some of the differences. Recall that the 2D lift coefficient (e.g., an airfoil) is defined as:

$$c_l = \frac{L'}{q_\infty c} \quad (4.3)$$

whereas the 3D lift coefficient (e.g., wing or full airplane) is:

$$C_L = \frac{L}{q_\infty S_{ref}} \quad (4.4)$$

The total lift is related to the lift per unit span from the integral:

$$L = 2 \int_0^{b/2} L' dy \quad (4.5)$$

If we use this definition we can relate the 2D lift coefficient distribution to

the overall wing lift coefficient:

$$C_L = \frac{L}{q_\infty S_{ref}} \quad (4.6)$$

$$= \frac{2 \int_0^{b/2} L' dy}{q_\infty S_{ref}} \quad (4.7)$$

$$= \frac{2 \int_0^{b/2} c_l q_\infty c dy}{q_\infty S_{ref}} \quad (4.8)$$

$$= \frac{2 \int_0^{b/2} c_l c dy}{S_{ref}} \quad (4.9)$$

$$(4.10)$$

If we divide the reference area by the span to define the mean geometric chord \bar{c} then:

$$C_L = \frac{2 \int_0^{b/2} c_l c dy}{b \bar{c}} \quad (4.11)$$

$$C_L = \int_0^1 \frac{c_l c}{\bar{c}} dy^* \quad (4.12)$$

$$(4.13)$$

where $y^* = \frac{y}{b/2}$. What this equation shows is that for the case where the chord is constant ($c = \bar{c}$) then the wing lift coefficient is the average of the airfoil lift coefficient distribution. For cases where the chord is not constant, it is a chord weighted average.

In Chapter 2 we learned that the theoretical lift curve slope of an airfoil was 2π . The same is not true for a finite wing because of downwash. Let's denote the 2D lift curve slope as m . If we had an infinite wing, with a constant airfoil then the wing lift coefficient would be:

$$C_L = m(\alpha - \alpha_0) \quad (4.14)$$

and the lift curve slope would be unchanged.

However, for a finite wing there is downwash and so we need to include the induced angle of attack:

$$C_L = m(\alpha - \alpha_0 - \alpha_i) \quad (4.15)$$

From finite wing theory, we can show that the induced angle of attack for an elliptic lift distribution is given by:

$$\alpha_i = \tan^{-1} \left(\frac{C_L}{\pi AR} \right) \quad (4.16)$$

For small angles we can approximate this as:

$$\alpha_i \approx \frac{C_L}{\pi AR} \quad (4.17)$$

Substituting this in Eq. (4.15) gives:

$$C_L \approx m(\alpha - \alpha_0 - \frac{C_L}{\pi AR}) \quad (4.18)$$

If we solve this for C_L and take the derivative with respect to angle of attack in order to get the wing lift curve slope we find:

$$\frac{dC_L}{d\alpha} = \frac{m}{1 + \frac{m}{\pi AR}} \quad (4.19)$$

where m is the 2D lift curve slope, typically a value close to 2π . Notice that as the aspect ratio goes to infinity we recover the 2D slope as expected. However, for finite wings the lift curve slope is reduced compared to that of its airfoils (Fig. 4.10). For example, an aspect ratio 8 wing would have a predicted lift curve slope of $0.8(2\pi)$ or 80% of the 2D value. Keep in mind that this is a simplified formula assuming constant airfoils, an approximately elliptic load, and low Mach number.

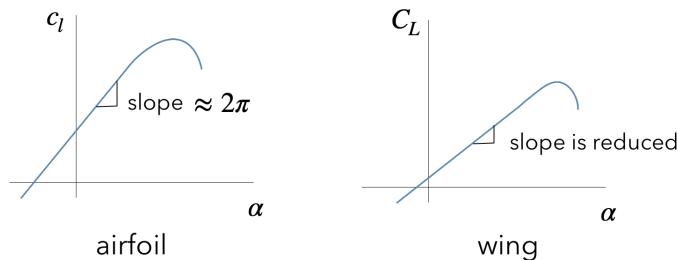


Figure 4.10: Lift curve slope for airfoil and wing.

CHAPTER 5

Stability

The best performing aircraft won't be of much use if it is too unstable to fly. This chapter provides an introduction to static and dynamic stability of flight vehicles. For many aircraft the longitudinal and lateral dynamics are effectively decoupled, and will be presented separately in this chapter.

5.1 Definitions and Fundamentals

We begin with a look at the forces and moments for a three-dimensional aircraft. In Chapter 2 we discussed the difference between wind and body axes for an airfoil. In three dimensions, we need one extra angle to define the rotation between wind and body axes (sideslip angle). Recall that lift and drag are defined in the wind axes, and that the wind axes are aligned with the freestream direction as shown in Fig. 5.1. The angle of attack defines one of the rotation angles between the body axes and wind axes. As we move to 3D, we need two additional parameters: the sideslip angle β and the side force Y .

In two-dimensions the only moment was the pitching moment. In three-dimensions we have three moments as shown in Fig. 5.2: roll (l), pitch (m), and yaw (n). We define these moments about the wind axes. Conventionally, the letter l is used for roll even though this can easily be confused with lift. In this text we will spell out *roll* in the coefficient.

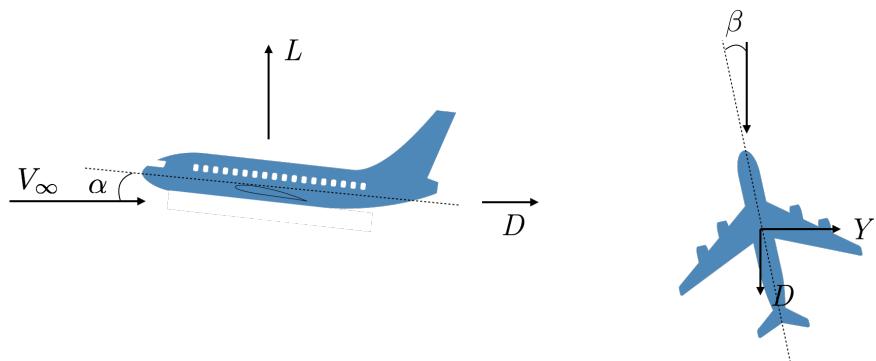


Figure 5.1: Wind axes. Lift and drag are defined relative to these axes. Definitions for positive angle of attack and sideslip also shown.

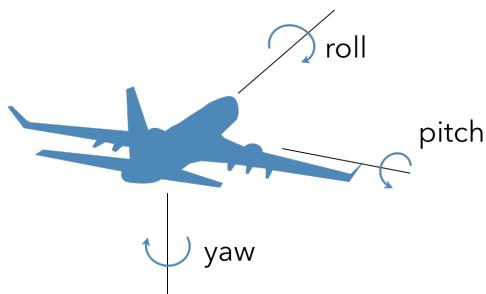


Figure 5.2: The three moments for a flight vehicle have specific names: roll, pitch, and yaw.

The forces and moment coefficients are defined as follows:

$$C_D = \frac{D}{q_\infty S_{ref}} \quad (5.1)$$

$$C_Y = \frac{Y}{q_\infty S_{ref}} \quad (5.2)$$

$$C_L = \frac{L}{q_\infty S_{ref}} \quad (5.3)$$

$$C_{roll} = \frac{roll}{q_\infty S_{ref} b_{ref}} \quad (5.4)$$

$$C_m = \frac{m}{q_\infty S_{ref} c_{ref}} \quad (5.5)$$

$$C_n = \frac{n}{q_\infty S_{ref} b_{ref}} \quad (5.6)$$

(5.7)

Notice that roll and yaw use span as the second reference dimension, whereas pitch uses chord. This is because the dynamics of roll and yaw are primarily effected by the change in inertia from span, whereas pitch is primarily affected by changes in chord.

The rotation rates (e.g., rad/s) about the three axes are: roll rate (p), pitch rate (q), and yaw rate (r). These rotation rates are normalized as follows:

$$\hat{p} = \frac{pb_{ref}}{2V_\infty} \quad (5.8)$$

$$\hat{q} = \frac{qb_{ref}}{2V_\infty} \quad (5.9)$$

$$\hat{r} = \frac{rb_{ref}}{2V_\infty} \quad (5.10)$$

(5.11)

A wide variety of control surfaces exist, but the most common are: ailerons to control roll, elevators to control pitch, and rudder to control yaw (see Fig. 5.3). An aircraft is said to be *trimmed* if the sum of its forces and moments are zero. In other words it is in static equilibrium. The control surfaces must have adequate response to trim the aircraft throughout its flight profile.

Figure 5.4 overviews the different types of stability. The x -axis in each figure is time, and the y -axis is some response (e.g., deflection). The upper

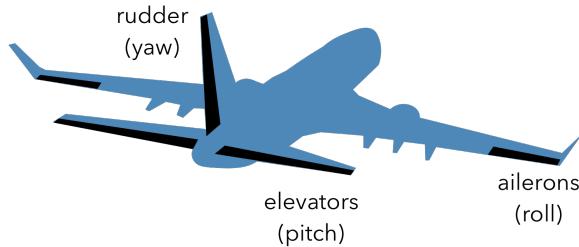


Figure 5.3: Typical control surfaces and the moments they control.

left corner represents a statically unstable system, also known as divergence. A disturbance from equilibrium causes the system to diverge further and further from its equilibrium point. The bottom left corner represents a statically stable system, also known as subsidence. A disturbance is damped out and the response converges to a steady state condition.

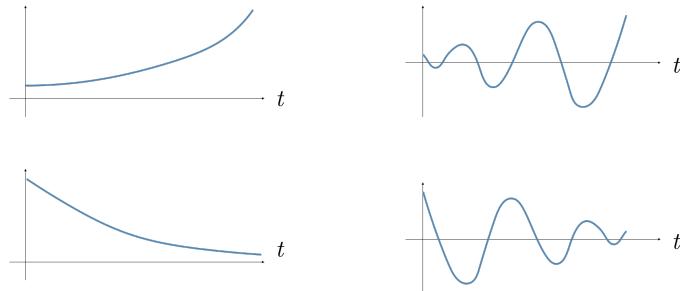


Figure 5.4: Four cases demonstrating static and dynamic stability/instability.

The upper right figure is statically stable, but dynamically unstable. This is also known as a divergence oscillation. The static stability causes the system to return towards equilibrium, but it overshoots each time leading to a dynamic instability. The bottom right figure is statically and dynamically stable, also known as a damped oscillation. One important takeaway from these figures is that it is possible to be statically stable and still dynamically unstable. Thus, if stability is desired, it is important to look at both the static and dynamic case.

Most aircraft are designed to be both statically and dynamically stable. However, too much stability is generally undesirable as it makes it difficult to maneuver. In fact, for some applications a small amount of instability is

desirable. For example, many fighter aircraft are designed to be unstable. This instability allows the fighters to be highly maneuverable, but it also requires fly-by-wire stabilization otherwise the pilot workload would be unmanageable (kind of like balancing a broomstick upside down). The tradeoff is perhaps more familiar by comparing a bicycle and a car. A bicycle is laterally unstable, whereas a car is laterally stable. The instability of a bicycle allows it to quickly maneuver side to side (conversely, it wouldn't be so great if your car was at risk of tipping on to its side whenever it stopped). For a transport aircraft, strong stability and slow movements are desirable, whereas for a fighter some instability with quick movements are needed.

Before proceeding there are two more concepts we need to understand. The first is the center of pressure. From statics you learned that the distributed loads acting on a body can always be resolved in an equivalent set of point forces and moments, at any location. The *center of pressure* is the location where the resultant moment is zero. Sometimes it is thought of as the point through which the resultant pressure field acts. For example, Fig. 5.5 shows a pressure distribution acting along an airfoil (just the upper surface distribution is shown for simplicity), and three different locations where we have resolved the distributed load into an equivalent set of forces and moments. Note that in the middle figure the chosen location has zero moment. This location is the center of pressure.

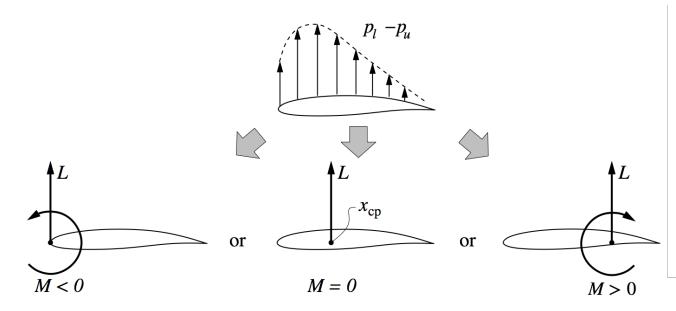


Figure 5.5: The distributed loads can be converted to a set of point forces/moments at any location, though the moment will differ depending on the chosen location. Figure by Joaquim Martins, used by permission.

While the center of pressure seems like a reasonable place to resolve the distributed load, it turns out to not be that useful for aircraft analysis. The problem with the center of pressure is that it moves around as the airplane

changes angle of attack. For example, Fig. 5.6 shows the airfoil at three different angles of attack, and thus with three different load distributions. The resulting center of pressure for each case is at three different locations (note that for the last case the center of pressure is not even on the airfoil). Because an airplane must fly through a range of angles of attack, particularly in a dynamics analysis, the center of pressure concept is cumbersome. A more useful point is the aerodynamic center. The *aerodynamic center* is the location on the vehicle where the moment does not change with angle of attack. In other words, it is the point where:

$$\frac{dC_m}{d\alpha} = 0 \quad (5.12)$$

Unlike the center of pressure, the aerodynamic center is independent of angle of attack (by definition). For most airfoils, its aerodynamic center is at or near the quarter chord ($c/4$). For a wing, its aerodynamic center is near the quarter-chord of the spanwise location where the chord equals the mean aerodynamic chord. The aerodynamic center, as defined, generally doesn't exist past stall where the lift and moment behavior is nonlinear.

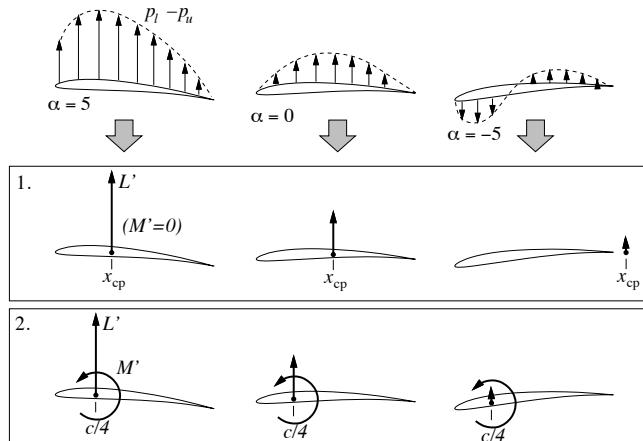


Figure 5.6: As the angle of attack changes the location of the center pressure also changes. The location for the aerodynamic center, however, stays constant. Figure by Joaquim Martins, used by permission.

5.2 Longitudinal Static Stability

The aircraft shown in Fig. 5.7 is flying at static equilibrium. If the aircraft encounters a disturbance that increases its angle of attack, α , then for it to be statically stable it requires a negative pitching moment, C_m , about the aircraft's center of gravity. Conversely, if the angle of attack was decreased, then a positive pitching moment would return it to its starting point.

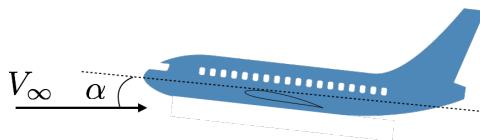


Figure 5.7: An aircraft in static equilibrium, used to illustrate longitudinal static stability.

Mathematically this condition can be expressed as:

$$\frac{\partial C_{mcg}}{\partial \alpha} < 0 \quad (5.13)$$

Indeed, this is the definition of longitudinal static stability. This partial derivative is called a stability derivative. There are many of them: all of the forces/moments with respect to various flight conditions including α, β, p, q, r and sometimes M (Mach number) and $\dot{\alpha}$ (rate of change of angle of attack). There are also *control derivatives*, which are the same except the derivative is with respect to control inputs like δ_a (aileron deflection), δ_e (elevator deflection), and δ_r (rudder deflection). Because there are many stability/control derivatives, a shorthand is often used. The above condition is written equivalently as:

$$C_{m,\alpha} < 0 \quad (5.14)$$

Vocally, this notation is read as: $C_m \alpha$ is less than zero. The fact that the moment is about the center of gravity is implied.

Consider the longitudinal forces and moments acting on an aircraft expressed generically in Fig. 5.8. We have not yet made any assumptions on the signs for any quantity in the figure (other than weight). A drag force also exists, but its effect on aircraft dynamics is generally negligible.

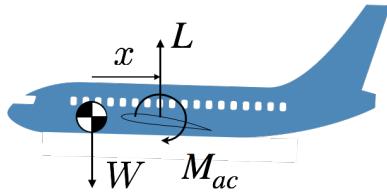


Figure 5.8: Longitudinal forces and moments acting on a generic aircraft.

The circle symbol with four quadrants is the symbol for the center of gravity (cg).

A force balance in the vertical direction shows that lift must be positive.

Now let us sum moments about the cg.

$$M_{cg} = M_{ac} - xL \quad (5.15)$$

We now divide everything by $q_\infty S$, and c in order to normalize.

$$C_{mcg} = C_{mac} - \frac{x}{c} C_L \quad (5.16)$$

Longitudinal static stability requires that $dC_{mcg}/d\alpha < 0$. If we take the derivative of the previous equation to apply the stability criteria, and use the definition of the aerodynamic center we have:

$$\frac{dC_{mcg}}{d\alpha} = -\frac{x}{c} \frac{dC_L}{d\alpha} < 0 \quad (5.17)$$

For an aircraft $dC_L/d\alpha > 0$ (lift increases with angle of attack), and so for stability we require $x > 0$. In other words, we require the center of gravity to be in front of the aerodynamic center. This is a general conclusion and so bears repeating: *for longitudinal static stability the aircraft's center of gravity must be forward of the aircraft's aerodynamic center*.

This quantity x/c is important enough that it has its own name: *static margin* (note that c is generally the mean aerodynamic chord, unless a different value of c was used in normalizing the pitching moment coefficient). The static margin is a measure of the stability of the aircraft.

Using the above expression we can also express the static margin as:

$$\text{static margin} = \frac{x}{c} = -\frac{dC_{mcg}/d\alpha}{dC_L/d\alpha} = -\left(\frac{dC_m}{dC_L}\right)_{cg} \quad (5.18)$$

The *neutral point* is the cg location on the aircraft where the static margin is zero. In other words, this is the limit to the farthest aft position you can place the cg and still have stability. The neutral point is the same thing as the aerodynamic center for the full aircraft. The term aerodynamic center can be used for components (e.g., the aerodynamic center of the tail, or the aerodynamic center of the wing), whereas the neutral point always refers to the full aircraft.

A typical static margin is 10–15% for general aviation, 5–10% for commercial/military transports, and < 0 for modern fighters. Sufficient stability is important, but more stability is not always better as it decreases responsiveness and because it incurs greater trim drag. Trim drag is the extra drag required to trim the aircraft. In other words, the deflection of control surfaces to zero out forces/moment causes extra parasitic and lift-dependent drag, which are bookkept as trim drag. Too much static stability can also sometimes lead to a dynamic instability (see upper right figure of Fig. 5.4).

Unfortunately, static stability is not enough. We also need the aircraft to be trimmed (in this case $C_{mcg} = 0$). Otherwise, the aircraft is not in equilibrium. If we apply the trim condition then we have:

$$C_{mcg} = C_{mac} - \frac{x}{c} C_L = 0 \quad (5.19)$$

or:

$$C_{mac} = \frac{x}{c} C_L \quad (5.20)$$

We already know that C_L must be positive, and for stability we just determined that x must be positive. That means that we require C_{mac} to be positive. Unfortunately, C_{mac} is negative for most any airfoil, and by extension most any wing. How then we do we trim the aircraft?

For a flying wing (tailless aircraft) there are three options. The first is to use a reflexed airfoil. A reflexed airfoil has its aft end curved upward. This shape is very inefficient and sacrifices quite a bit of lift (the aft end actually lifts downward), and it usually has a poor maximum lift coefficient. But it does allow for a positive C_{mac} , which is what was needed to trim. The second, and most common option, is to sweep the wing and use washout (twist tips down). At zero lift, the tips will actually lift down (the tips won't, or shouldn't, lift down at the cruise angle of attack). It can be a bit tricky to design well, but with an appropriate combination of sweep, twist, taper, and aspect ratio, one can relatively efficiently achieve stability and

trim. The third option is to use conventional airfoil/wing design, but just let the airplane be unstable ($x/c < 0$) and instead rely on active controls to fly the airplane.

Instead of using a flying wing, the more common solution is to add a tail. We analyze this scenario using the forces and moments shown in Fig. 5.9. Again, we neglect the minor impact of drag, and assume the offset of the tail relative to the wing is small. Usually, tails use symmetric airfoil sections so $M_t = 0$ and is thus not shown. The fuselage, nacelles, and other components of the aircraft do usually contribute a non-negligible moment, but don't change the basic conclusions for this analysis so they are omitted. Generally, the lift from these other components is negligible.

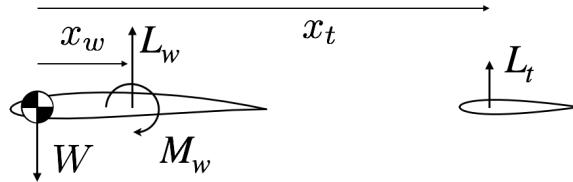


Figure 5.9: Primary longitudinal forces and moments acting on a wing/tail combination. The w subscripts correspond to the wing, and the t subscripts correspond to the tail.

The stability condition is unchanged (center of gravity must be forward of aerodynamic center), but we now need to assess trim. Computing moments about the center of gravity gives:

$$M_{cg} = M_w - x_w L_w - x_t L_t \quad (5.21)$$

The normalizing parameters used for the wing lift coefficient, are different from those used in the tail lift coefficient so we expand each term:

$$C_{mcg} q_\infty S_w c_w = C_{mw} q_\infty S_w c_w - x_w C_{Lw} q_\infty S_w - x_t C_{Lt} q_t S_t \quad (5.22)$$

Notice also that the dynamic pressure experienced by the tail is not q_∞ as the tail is generally in the wake of the wing. We now divide through by $q_\infty S_w c_w$:

$$C_{mcg} = C_{mw} - \frac{x_w}{c_w} C_{Lw} - C_{Lt} \eta \left(\frac{S_t x_t}{S_w c_w} \right) \quad (5.23)$$

where $\eta = q_t/q_\infty$, the ratio of dynamic pressures, is called the tail efficiency. Typically it is about 0.85–0.9 [1]. The tail experiences a lower

dynamic pressure because it is the wake of the wing and/or the boundary layer of the fuselage. For a T-tail, the tail is moved up high out of the wake and so $\eta = 1$. The quantity in parenthesis $(S_t x_t)/(S_w c_w)$ is known as the tail volume coefficient and is often used for preliminary sizing of the tail (as we will see in Section 5.5).

Recall that we need the pitching moment about the cg to be zero to trim:

$$C_{mw} - \frac{x_w}{c_w} C_{Lw} - C_{Lt} \eta \left(\frac{S_t x_t}{S_w c_w} \right) = 0 \quad (5.24)$$

The pitching moment on a typical wing C_{mw} is negative as discussed, and x_w is positive for static stability (the wing provides most of the lift so the aerodynamic center will be near here) and so the second term is negative as well. Thus, for trim we require the last term to be positive. This is most easily accomplished by making C_{Lt} negative. Indeed this is how a conventional configuration operates. The purpose of the horizontal tail is to lift downward and balance the moment generating by the wing, allowing for the aircraft to be trimmed. The downside is that the tail is lifting down, and so there is increased drag both from the tail, but also from the extra incremental lift you need to carry in the wing. This lift-dependent drag can be reduced by moving the tail further aft, so as to make the moment arm larger, and thus reducing the amount of downward lift the tail must create. The tradeoff is the increase weight and drag of the structure that allows the tail to move aft.

The other alternative is to allow x_t to be negative, or in other words to put the “tail” in front of the wing. This type of configuration is called a canard (see for example Fig. 5.10). The advantage of this configuration is that the canard can generate positive lift, allowing for a smaller main wing.

However, the wake from the canard significantly affects the wing and so the interference drag is larger. Additionally, to make the aircraft stable the canard must stall before the main wing, which means that the main wing must be larger than it would be otherwise—counteracting the main benefit of the configuration. In short, there is not an obvious aerodynamic net benefit, but it can be more efficient in some cases. The design is more complex as the aerodynamic interference and cg sensitivity is usually greater. However, one nice operational benefit is that a canard configuration can be designed to be virtually stall proof. This is accomplished by making the canard stall before the main wing. When the canard stalls the lift drops forward of the cg and so the airplane returns

back to a stable configuration (conversely if the airplane were designed so that the main wing stalled first, then the aircraft would pitch up further deepening the stall, which could be catastrophic). This safety feature is an appealing benefit for home-built aircraft like the Long-EZ in Fig. 5.10.



Figure 5.10: The Rutan Long-EZ, a canard configuration designed to be stall-proof. Photo by [Adrian Pingstone](#), public domain.

Flying wings, conventional tail, and canards are not the only configuration options to address longitudinal stability. Various configurations exist including: a tandem wing, three surfaces, joined wing, box wing, twin fuselage with dual tails, etc.

For longitudinal static stability we have learned that placing the center of gravity ahead of the aerodynamic center is critical. Finding the center of gravity is fairly straightforward, but how do we find the aerodynamic center? To figure this out, we will use Fig. 5.11. On the right is the aerodynamic center. We don't know where it is located, but we know it exists (as long as we are not stalled). On the left is some arbitrary point that we are going to take moments about. We want to solve for x which will tell us exactly where the aerodynamic center is relative to point we are taking moments about.

First, we take moments about the arbitrary point p :

$$M_p = M_{ac} - xL \quad (5.25)$$

Now we take the derivative of the above equation with respect to angle of attack

$$\frac{dM_p}{d\alpha} = \frac{dM_{ac}}{d\alpha} - x \frac{dL}{d\alpha} \quad (5.26)$$

By definition, the derivative of M_{ac} with respect to α is 0. We can now solve for x :

$$x = -\frac{dM_p/d\alpha}{dL/d\alpha} \quad (5.27)$$

or normalizing:

$$\frac{x}{c_{ref}} = -\frac{dC_{mp}/d\alpha}{dC_L/d\alpha} \quad (5.28)$$

where c_{ref} is the chord used in normalizing the pitching moment coefficient, normally the mean aerodynamic chord.

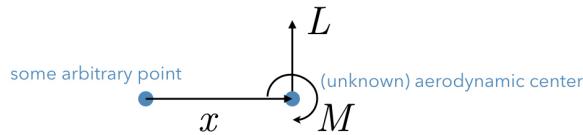


Figure 5.11: Diagram used in determining location of aerodynamic center.

Thus, we just need our full aircraft lift curve slope and the moment slope about any arbitrary point (although this equation could be a little more general by adding drag and vertical offsets). Notice, that if I happened to already be at the aerodynamic center then $dM_p/d\alpha$ is 0 as so x is 0 as expected.

If you also want to find the value of M_{ac} an easy way to do so is to refer to Eq. (5.25) and note that if the lift is zero, then the corresponding moment is the moment at the aerodynamic center. In other words, the zero-lift pitching moment (C_{m0}) is the pitching moment at the aerodynamic center.

5.3 Lateral Static Stability

Let's now determine the criteria for yaw stability, also known as directional stability, by examining Fig. 5.12. Based on the sign conventions show in the figure we see that if a disturbance increases the sideslip angle β , then we need a positive yawing moment to suppress the disturbance. In other words:

$$\frac{dC_n}{d\beta} > 0 \quad (5.29)$$

or in the simplified notation:

$$C_{n,\beta} > 0 \quad (5.30)$$

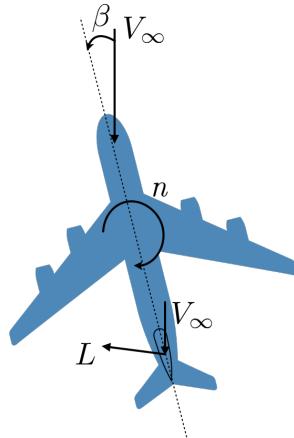


Figure 5.12: Illustration of criteria for yaw (directional) stability through a disturbance in β .

Be careful with the sign conventions. Recall that a positive α is in the direction of a positive pitching moment, but a positive β is in the direction of a negative yawing moment. A typical desired value for $C_{n,\beta}$ is around 0.06–0.15 (with β in radians) [2, 3]. This is one stability derivative where bigger is usually better. This is not to suggest that a bigger tail is always desirable, but unlike many of the other stability characteristics, if you could get more directional stability for free then you would want it as it improves handling.

The vertical tail provides most of the directional stability, although the vertical tail is usually sized by control requirements rather than static stability. Figure 5.12 also shows a vertical tail and its resultant lift force. If the vertical tail is behind the cg then this lift provides a restoring moment to reduce sideslip.

Roll stability is a little trickier to visualize. Figure 5.13 shows the forces acting on an aircraft with a disturbance in the bank angle. The tricky aspect is that bank angle (like pitch and yaw angles) does not cause a change in aerodynamic forces directly. Aerodynamic forces only depend on the relative wind, and bank angle is independent of wind. Weight also does not cause a change in moments as it always acts through the cg. The reason a bank angle disturbance causes a potential instability is that the bank angle tilts the lift vector away from vertical and so a sideways component of force exists (to the right in the figure). That sideways force

causes the aircraft to move to the right, and so it begins to sideslip. Based on the sign convention, this is a positive sideslip angle (if the aircraft is moving from left to right, then the relative wind is from right to left). To reduce the sideslip we need a negative rolling moment. Thus, the criteria for roll stability is:

$$\frac{dC_{roll}}{d\beta} < 0 \quad (5.31)$$

or

$$C_{roll,\beta} < 0 \quad (5.32)$$

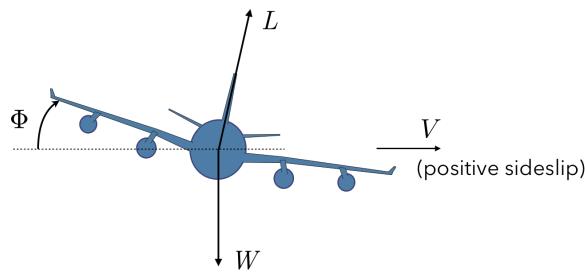


Figure 5.13: Illustration of criteria for roll stability through a disturbance in roll angle. The aircraft is flying into the page.

Too much roll stability is undesirable as the aircraft becomes overly sensitive when turning. An ideal range for this stability derivative is much more variable, but as a rough reference point, a typical value for $C_{roll,\beta}$ is between -0.1–0 [2].

This stability derivative is also known as “effective dihedral”. That is because dihedral increases roll stability. To see this, consider the scenario shown in Fig. 5.14. The scenario is identical to that shown in the previous figure, namely a disturbance has created a positive bank angle. However, this time the wings have some dihedral. Like before, the relative wind from sideslip comes from right to left shown in black (shown below the airplane so that the arrow directions are easier to see). We can break up each black velocity vector into its components, the red vectors, parallel and perpendicular to the wing dihedral. The right wing experiences an upward velocity, whereas the left wing experiences a downward velocity. The net effect is a restoring rolling moment to reduce the bank angle. Thus, adding dihedral increases roll stability. Winglets also help to increase roll stability. Negative dihedral, called anhedral, has the opposite effect and will reduce roll stability.

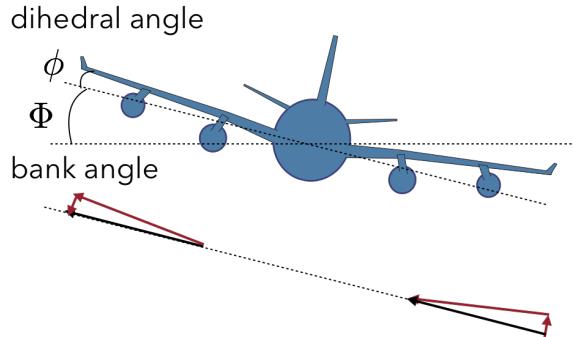


Figure 5.14: Positive dihedral increases roll stability.

5.4 Coordinated Turns and Adverse Yaw

As noted in Fig. 5.3 turning primarily occurs with ailerons, or in other words through a rolling moment. The turn is initiated through antisymmetric aileron deflection. A right turn, for example, has the left aileron deflect down and the right aileron deflect up. The aileron on the left wing increases the local camber and thus lift, whereas the aileron on the right wing has the opposite effect and decreases lift. The increase in lift on the left, and decrease in lift on the right, causes a rolling moment (Fig. 5.15) and thus a component of lift to the right that allows for a right turn.

One of the challenges with this maneuver is that because the lift on the left side increases, the induced drag on the left side also increases. Conversely, the induced drag on the right side decreases. The net effect is a yawing moment opposite the turn direction that creates sideslip (see right side of Fig. 5.15). This phenomenon is called adverse yaw.

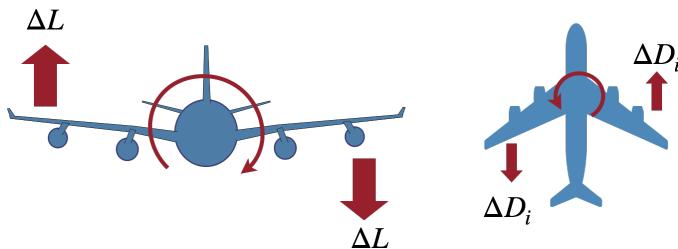


Figure 5.15: An illustration of adverse yaw.

Typically, the way this problem is addressed is to deflect the rudder at the same time in order to eliminate the yawing moment (generally the elevators would be deflected as well because tilting the lift vector during the roll causes some loss in lift). Such a maneuver is called a *coordinated turn*. Stated another way, a coordinated turn is a turn that occurs without any sideslip. Later, in the performance chapter, we will assume that turns are coordinated.

Rudder is not used to turn aircraft. Generally its only purpose is to eliminate yawing moments. Turning with rudder will induce a rolling moment and the amplitude of that rolling moment is difficult to control with rudder. Furthermore, it will create a sideslip. Generally, this is an inefficient and (if carrying passengers) uncomfortable approach to turning. Small RC airplanes have low moments of inertia, and so with sufficient roll stability, sometimes just rudder is used for turning. The benefit is simplicity in controls for a beginning RC pilot.

5.5 Tail Types and Statistical Tail Sizing

The combination of horizontal and vertical tails is referred to as the empennage. There are many types of empennages, some of which are shown [here](#).

A T-tail is used to move the horizontal tail out of the engine exhaust and fuselage boundary layer. The effectiveness of its vertical tail is increased because it is end-plated, similar to a winglet. However, a T-tail is susceptible to flutter problems (think of a weight at the end of a beam), and deep stall problems (where the aircraft is at a high angle of attack, and the elevators on the horizontal tail becomes waked making it difficult or impossible to escape the stall).

A cruciform tail is a compromise between the conventional tail and the T-tail. The structural issues are mostly avoided, but you don't get the end-plating benefit. A dual tail, also known as an H-tail when the vertical tails extend below, increases the effectiveness of the horizontal tail through end-plating. Having two vertical tails allows for greater yawing moments and can prevent issues with the rudder lining up with the propulsion system wake and becoming ineffective. The downside of the dual tail is that the control linkage is more complex, and, in the case of a true H-tail, ground clearance is reduced.

A V-tail can increase ground clearance, but increases control complexity as the rudder and elevator must be mixed. A Y-tail (not shown in that link) uses downward surfaces to protect a pusher propeller from ground strike. Tail sizing is complex, and includes more factors than just stability, such as control, drag, and weight. However, as a starting point, statistical methods can be used to provide rough estimates for proper tail sizing. This type of approach assumes that you are designing a conventional aircraft, or more precisely one that is like the past aircraft that the fits are based on. Because tails add both weight and wetted area, the designer will generally prefer the smallest tail possible that provides the necessary level of stability and control.

Some guidelines as an initial starting point are provided below, primarily based on commercial transports and general aviation aircraft [4, 3]:

- A typical aspect ratio is 3–5 for a horizontal tail, and 1.3–2 for a vertical tail.
- A typical taper ratio for a tail is 0.4–0.6.
- T-tails tends to have a higher aspect ratios (5–5.5) for the horizontal tail to avoid wake effects, and lower aspect ratios for the vertical tail along with a higher taper ratio (0.85–1).
- A horizontal tail area is typically 20–25% of the wing area.
- The downward lift produced by the horizontal tail may be as high as 5% of the aircraft weight.
- In the longitudinal stability analysis we saw a grouping of coefficients we referred to as the tail volume coefficient. There is actually a coefficient for both the horizontal and vertical tail:

$$V_h = \frac{S_h l_h}{S_w c_w}, \quad V_v = \frac{S_v l_v}{S_w b_w} \quad (5.33)$$

where the w subscript corresponds to the wing, h to the horizontal tail, and v to the vertical tail. Note the use of span, rather than chord, in the normalization of the vertical tail coefficient. A typically horizontal tail coefficient is 0.5–0.7 and a typical vertical tail coefficient is 0.02–0.04.

- The tail arm is typically 45–65% of the fuselage length.

- Most tails use symmetric airfoil sections.

It should be emphasized again that these are only statistical correlations and may or may not apply to an individual aircraft. These types of fits should only be used as starting points and are not a replacement for important sizing constraints like: stability, nose-wheel rotation, control authority, etc.

5.6 Dynamic Stability

As discussed in the first part of this chapter, static stability is important but not enough. We must also examine dynamic stability. From a dynamics course you might recall that the stability of a dynamic system can be assessed from its eigenvalues. In general, these eigenvalues are complex, meaning that they have a real and an imaginary part. The real part of the eigenvalues tells us whether the response will decay or amplify. If the real part (σ) is negative, then that particular mode is dynamically stable. The complex part is the damped natural frequency. The larger the value, the higher the frequency of oscillation in the dynamic response. Referring back to Fig. 5.4 we can associate eigenvalues with each case. The upper left is real and positive, the lower left is real and negative, the upper right is complex with the real part positive, the lower right is complex with the real part negative.

Typically, these eigenvalues are plotted in the complex plane, like that shown in Fig. 5.16. All eigenvalues on the left-half plane are dynamically stable. The further to the left they are, the faster they decay. The plot also helps us visualize other parameters like the undamped natural frequency (ω_n) or the damping ratio (ζ), both of which are shown in Fig. 5.16. These parameters can tell us about things like settling time, overshoot, etc.

However, for this discussion our primary interest is in dynamic stability, which is indicated by the real part of the eigenvalue (σ).

For an aircraft, the primary modes have names associated with them.

Figure 5.17 shows the eigenvalues associated with the longitudinal dynamics. There are two modes: the short period and the phugoid.

The short period and phugoid modes are visualized in Fig. 5.18a and Fig. 5.18b respectively. The short period mode is generally highly damped, and thus not objectionable. Most pilots don't even feel it. It results in oscillations caused by changes in angle of attack and altitude, but occurs

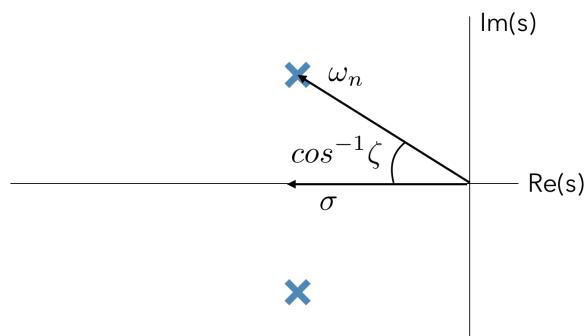


Figure 5.16: A pair of eigenvalues for a particular mode plotted in the complex plane.

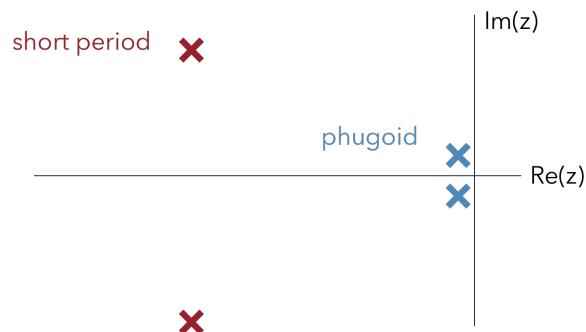


Figure 5.17: Longitudinal dynamic modes for an aircraft.

mostly at constant airspeed. The frequency of the short period mode is strongly related to the aircraft's static margin.

In contrast, the phugoid mode is lightly damped and has a low frequency (or long period—in fact sometimes it is referred to as the long period mode). It is associated with gradual changes in altitude and airspeed, roughly at constant angle of attack. This mode is generally only very lightly damped, or sometimes it is somewhat unstable. If piloting under visual flight rules (VFR, which means that visibility is good and the pilot can fly just with external cues) then some phugoid instability is generally not a problem. This is because the pilot response time is much faster than the instability period so it can be easily corrected, often without even noticing. However, if piloting under instrument flight rules (IFR, which means that visibility is poor and the pilot must rely on instrumentation), then low phugoid damping can be problematic.

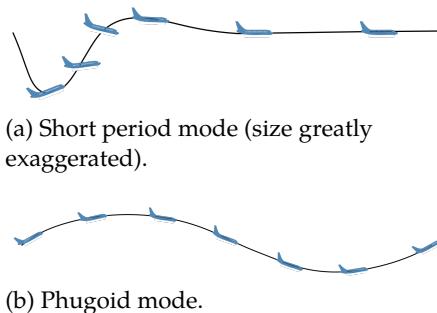


Figure 5.18: Visualization of longitudinal dynamic modes.

For lateral dynamics there are three modes: roll subsidence, spiral mode, and dutch roll (Fig. 5.19). The roll subsidence mode (Fig. 5.20a) is usually uninteresting. Typically it is highly damped with no oscillatory component. The spiral mode (Fig. 5.20b) is slow and lightly damped. Because it is slow it can be corrected easily when detected, but if not corrected the spiral continues to tighten with decreases in altitude, and eventually enters a high-speed dive.

The dutch roll mode is a coupling between yawing and rolling. It can be quite noticeable in a flight and generally leads to passenger discomfort. Thus, for ride quality, higher damping is desirable even though it is often a challenging mode to dampen sufficiently.

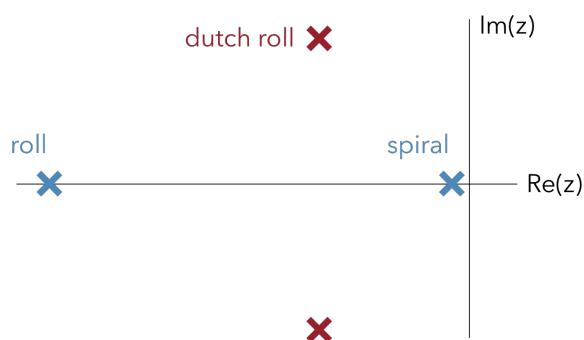


Figure 5.19: Lateral dynamic modes for an aircraft.

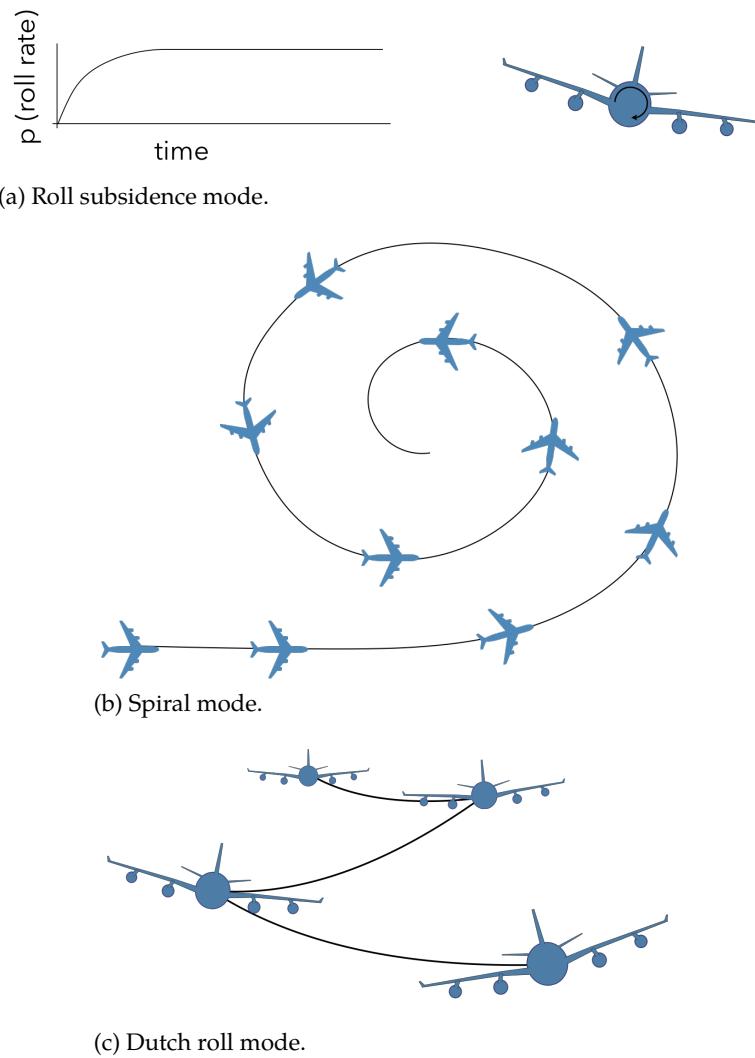


Figure 5.20: Visualization of lateral dynamic modes.

Bibliography

- [1] Shevell, R. S., *Fundamentals of Flight*, Pearson College Division, 1989.
- [2] Phillips, W. F., *Mechanics of Flight*, John Wiley & Sons, Jan 2004.
- [3] Raymer, D. P., *Aircraft Design: A Conceptual Approach*, American Institute of Aeronautics and Astronautics, 2012.
- [4] Kroo, I., *Aircraft Design: Synthesis and Analysis*, Stanford University, 2006.

CHAPTER 6

Propulsion

Without a propulsion system an aircraft is just a glider. Historically, improvements in propulsion systems have contributed about half of the fuel-burn reductions for aviation. We start with electric motors and propellers, which are used for small UAVs and are enabling emerging applications like vertical takeoff and landing “air taxis”. Next, we discuss some of the basics of gas turbine engines. Finally, we end with a brief discussion on placement of the propulsion system.

6.1 DC Electric Motor

Figure 6.1 shows a diagram of a DC motor. A battery supplies a voltage (v_b) across a motor with some resistance (R). The back emf (v_m) is the electromotive force caused by the spinning shaft. The end product of the motor is a torque (Q) at some rotation speed (Ω).
The rotation speed is proportional to the back emf, for a constant magnetic field:

$$\Omega = K_v v_m \quad (6.1)$$

The *motor velocity constant* K_v is typically given in RPM/Volts. The torque is generated by the magnetic field and so is proportional to current. To be more accurate, there is a *no-load current* (i_0) required just to overcome

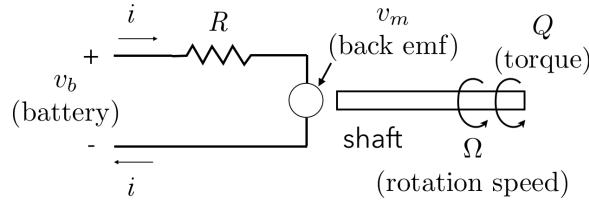


Figure 6.1: Diagram of a DC motor.

friction before a positive torque is generated:

$$Q = K_t(i - i_0) \quad (6.2)$$

From an energy balance we can show that the torque constant, K_t , must be equal to $1/K_v$ ¹. Thus:

$$Q = \frac{1}{K_v}(i - i_0) \quad (6.3)$$

The battery must supply enough voltage to overcome the back emf and the resistance:

$$v_b = v_m + iR \quad (6.4)$$

If we solve this expression for v_m , substitute v_m into Eq. (6.1), and rearrange the expression in terms of current we have:

$$i = \left(v_b - \frac{\Omega}{K_v} \right) \frac{1}{R} \quad (6.5)$$

With the current known, we can now solve for the torque using Eq. (6.3). The power delivered to the shaft is then:

$$P_{shaft} = Q\Omega \quad (6.6)$$

The motor efficiency is the power out (shaft power) divided by power in (batter power):

$$\eta_m = \frac{Q\Omega}{iv_b} \quad (6.7)$$

Thus, given the three motor constants: K_v , R , i_0 , and the battery voltage, we can solve for the motor current, torque, power, and efficiency as a

¹See [derivation](#) and a [more sophisticated model](#) that includes heat transfer from Mark Drela

function of Ω . Motor constants can be found online for most electric motors, but it can be worthwhile to measure these values for yourself. For example, consider a motor with the constants $K_v = 3,691 \text{ RPM/V}$, $i_0 = 0.28 \text{ amps}$, $R = 1.04 \text{ Ohms}$, and a battery voltage of 6V. The resulting motor torque, power, and efficiency as a function of rotation speed is shown in Fig. 6.2.

6.2 Propellers

A detailed look at propeller aerodynamics is beyond the scope of this text. For the purposes of design we just need a functional relationship between inputs like the operating condition and outputs like thrust and torque. In general, we could write the thrust as a function of many parameters like the airspeed, rotation rate, propeller diameter, air density, air viscosity, speed of sound, blade shape, and pitch:

$$T = f(V, \Omega, D, \rho, \mu, a, c, \theta, \theta_p) \quad (6.8)$$

As discussed in Chapter 2, nondimensional analysis is fundamental to many aerodynamic analyses, and we apply the same principles here. There are many ways the relevant parameters could be normalized, in the following we use standard normalization for propellers. The thrust coefficient is defined as:

$$C_T = \frac{T}{\rho n^2 D^4} \quad (6.9)$$

where n is the propeller rotation rate in revolutions per second, and D is the propeller diameter. In other words if Ω is in radians per second then:

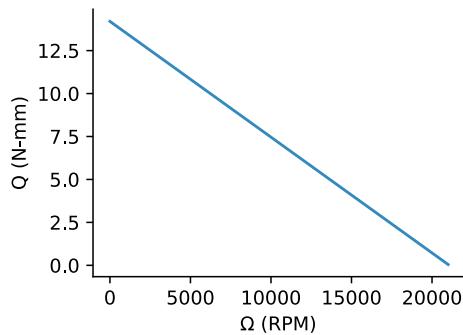
$$n = \frac{\Omega}{2\pi} \quad (6.10)$$

and is given in revolutions per second. The torque coefficient is defined as:

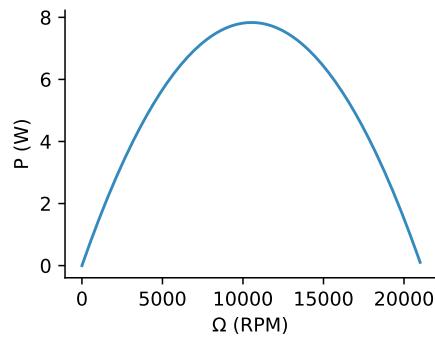
$$C_Q = \frac{Q}{\rho n^2 D^5} \quad (6.11)$$

and the power coefficient is defined as:

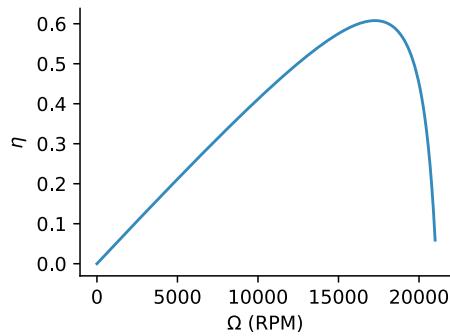
$$C_P = \frac{P}{\rho n^3 D^5} \quad (6.12)$$



(a) Motor torque.



(b) Motor power.



(c) Motor efficiency.

Figure 6.2: Electric motor performance for an example configuration as a function of RPM.

Based on these definitions you can show that these last two coefficients are directly related by:

$$C_Q = \frac{C_P}{2\pi} \quad (6.13)$$

Propeller efficiency is already nondimensional, and is a ratio of the useful power out from thrust, divided by the power in from the shaft (which is the power out from the motor, assuming no additional losses):

$$\eta = \frac{P_{out}}{P_{in}} = \frac{TV}{Q\Omega} \quad (6.14)$$

For the inputs, most of the normalizations will already be familiar and include the Reynolds number, and the Mach number. One new normalization is that of the rotation speed. The *advance ratio* is defined as a ratio of forward speed to rotation speed (times a constant factor):

$$J = \frac{V}{nD} \quad (6.15)$$

Using these nondimensionalizations we can express thrust (or any of the other outputs) in a functional form with fewer inputs:

$$C_T = f(J, Re, M, geometry^*) \quad (6.16)$$

where *geometry** means a normalized geometric description of the blade. For some applications the Reynolds number doesn't change appreciably across the flight envelope, or the variation in performance with Reynolds number is less significant. For smaller UAVs, the Mach number is insignificant. In these cases, for a given propeller, we can determine thrust, torque, power, and efficiency as just a function of advance ratio:

$$C_T = f(J) \quad (6.17)$$

Figure 6.3 shows the thrust coefficient, power coefficient, and efficiency of an example propeller as a function of advance ratio.

From experimental data, or an aerodynamics simulation, this type of data can be tabulated for a given propeller. The [UIUC Propeller Data Site](#), for example, contains this kind of data for a large number of propellers used for small UAVs.

Given the data for a propeller(s), and a known operating condition (freestream speed, and propeller rotation speed), the thrust coefficient,

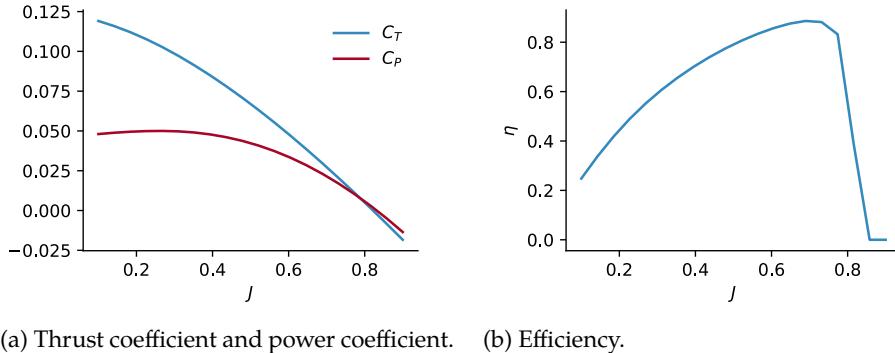


Figure 6.3: Example propeller data as a function of advance ratio.

power coefficient or efficiency can be found as a function of advance ratio using interpolation. Then the coefficients can be unnormalized to get thrust and torque for that operating condition. The difficulty is knowing what rotation speed the propeller will operate at, and this is the subject of the next section.

6.3 Motor-Prop Matching

The difficulty in choosing the right motor-propeller combination (actually the right motor-propeller-battery combination) is that their performance is coupled. The rotation speed, for example is not an independent parameter. The motor shaft rotates the propeller so they must both have the same rotation speed, and the same torque.

For example, Fig. 6.4 shows the torque of a motor using the methods outlined in Section 6.1 as a function of the rotation speed and the battery voltage. It also shows the torque of the propeller using the methods outlined in Section 6.2 as a function of rotation speed and freestream velocity. The operating torque and rotation speed, will be where these two curves intersect. Thus the rotation speed is an implicit function of the freestream speed and the voltage across the battery.

Furthermore, that rotation speed then determines the motor efficiency, propeller efficiency, and propeller thrust. This is illustrated in Fig. 6.5. The dashed line shows the rotation speed at which the two torque curves cross, and thus the operating rotation speed. The efficiency of the motor and the

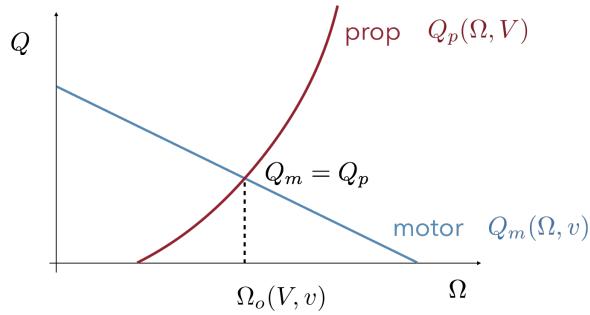


Figure 6.4: The rotation rate and torque of the motor/prop combination is determined by the intersection of the motor and propeller torque curves. In other words, the rotation rate and torque are implicit functions of velocity and voltage.

propeller are then determined from this speed. Note that the resultant efficiency for both the motor and propeller is poor compared to their peak efficiencies. We would call this system *poorly matched*. Finally, the thrust of the propeller is determined from this rotation speed, which may or may not be sufficient for the aircraft.

Let's discuss the matching problem in more detail. Figure 6.6 shows two motor/prop combinations. The one on the left is poorly-matched. No matter what speed we operate at, the efficiency will be relatively low because the peaks for the two curves occur at very different locations. In contrast, a *well-matched* motor/prop system is one where their peak efficiencies line up closely. The right hand graph in that figure is a much better match as the peak efficiencies more closely line up.

The currently outlined approach gives the outputs (thrust, torque, and efficiency) as a function of freestream velocity and battery voltage. For the purposes of design we may want to look at the limiting case where we are using full throttle (i.e., the full voltage of the battery). This means that the battery sizing is also coupled in to this problem (both in terms of max voltage, but also in terms of weight). Note that if you wish to explore cases for partial throttle, for conceptual design we will assume that the electronic speed control scales throttle linearly with voltage. However, not all speed controllers have a linear scaling. In any case, using a known battery voltage reduces to computation of thrust, torque, and efficiency to

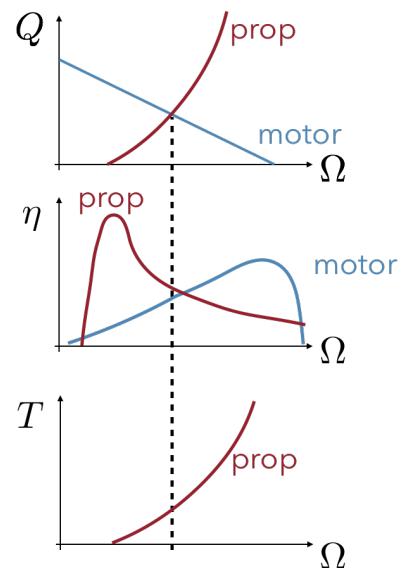


Figure 6.5: The rotation speed from the intersecting torque curves in turn determines the efficiency of both the motor and the propeller, as well as the thrust of the propeller.

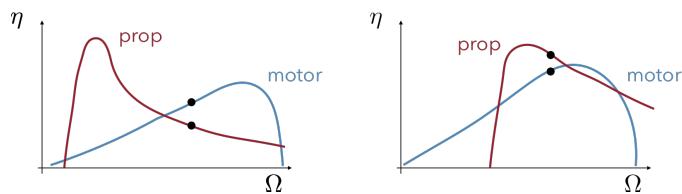


Figure 6.6: An example of a poorly matched, and better (if not well) matched motor-prop system.

just a function of flight speed.

$$T, Q, \eta = f(V) \quad (6.18)$$

To make sure the design is adequate we need to compare thrust across the flight speed range, with drag across the flight speed range. As discussed in Chapter 3 we can compute the drag as a function of airspeed:

$$D(V) = C_{Dp} q S_{ref} + \frac{L^2}{q \pi b^2 e} \quad (6.19)$$

Clearly, we need $T > D$ across the speeds we want to fly at. Equality is sufficient to maintain level flight, but in general we need excess thrust so that we can climb (and to provide some margin for uncertainties in these estimates).

Choosing the right motor/prop/battery combination might seem daunting, and indeed will require some iteration, but can be made easier by thinking about the most important parameters. For example, most fixed-pitch small UAV propellers have a peak efficiency at around the same advance ratio, generally between 0.4–0.7. If you know your design flight speed, and approximate propeller size, then you can compute the corresponding rotation speed. For a small UAV you will likely use a 2- or 3-cell LiPo battery, which has a voltage of 7.4 or 11.1 V respectively. With a known battery you can compute the K_v value that will give you a well-matched RPM at say 70% max voltage. This thought process ignores many coupling effects, but is a starting point for further iteration. On our website² there is a script to help you analyze how well your motor/prop/battery system is matched, and to plot thrust and drag across a range of speeds.

One way to help size the motor is to look at the power-to-weight ratio. In Chapter 7 we will learn more about why this is an important parameter for electric-powered aircraft. For small UAVs a reasonable target is 150 W/kg, or for an aerobatic UAV perhaps as much as 300 W/kg. Modern RC aircraft almost always use **brushless motors**. The electronic speed controller (ESC) is sized by making sure it is rated for the maximum current of the motor and the maximum voltage of the battery (with some margin).

²<http://flow.byu.edu/me415/>

6.4 Gas Turbine Engines

A propulsion system, from a propeller to a rocket, can be represented generically as shown in Fig. 6.7. Air may enter the system at a mass flow rate of \dot{m} at the freestream speed of the vehicle, and leave at some exit velocity V_e . In between, fuel may be injected at a mass flow rate of \dot{m}_f . The thrust generated by this system is given by a momentum balance (derivation omitted):

$$T = \dot{m}(V_e - V_\infty) + \dot{m}_f V_e + (p_e - p_\infty)A_e \quad (6.20)$$

where p is pressure, and A_e is the exit area. For a rocket engine $\dot{m} = 0$ as there is no air entering. For an electric motor there is no combustion so $\dot{m}_f = 0$. For a jet engine, all of these terms are important. For subsonic applications, or a fully expanded rocket nozzle (which is the ideal condition), $p_e = p_\infty$. Also for a jet engine the mass fuel rate is generally much lower than the mass flow rate of the air. Thus, for an electric motor and (approximately) for a jet engine at subsonic conditions the thrust equation reduces to:

$$T = \dot{m}(V_e - V_\infty) \quad (6.21)$$

We now have the intuitive result, that to generate thrust, an air-breathing propulsion system must accelerate the incoming air.

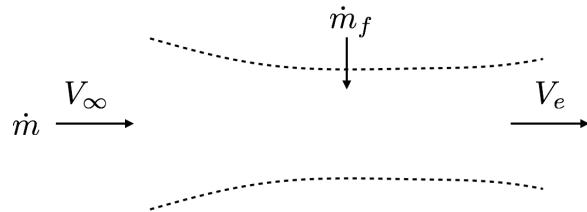


Figure 6.7: A generic representation of a propulsion system with incoming air flow rate \dot{m} , fuel mass flow rate \dot{m}_f , and exit velocity V_e .

The propulsion system must not only generate sufficient thrust, but should do so efficiently. To analyze efficiency, we will look at the propulsion system in the ground frame of reference Fig. 6.8. The propulsion system, depicted as the blue box, travels at some speed V_∞ . Because we are in the ground frame of reference, there is still (not-moving) air ahead of it, and in its wake the relative velocity is $V_e - V_\infty$. This energy in the wake represents wasted energy. Like the discussion with drag, if possible, we

would like the propulsion system to pass through and leave no wasted energy in the wake. Unfortunately, this is not compatible with creating thrust as seen in our thrust equation (Eq. (6.21)).



Figure 6.8: A propulsion system flying by in the perspective of the ground frame of reference.

The overall efficiency of a propulsion system is given by the product of the propulsive efficiency and the thermal efficiency, and is the ratio of power in the fuel (enthalpy times mass flow rate) to useful work out to the vehicle:

$$\eta = \eta_t \eta_p = \frac{TV_\infty}{\dot{m}_f h_f} \quad (6.22)$$

where h_f is the enthalpy of the fuel. The thermal efficiency is the efficiency of converting the power in the fuel to mechanical power in the working fluid (power to vehicle plus power lost in wake):

$$\eta_t = \frac{TV_\infty + \frac{1}{2}(\dot{m} + \dot{m}_f)(V_e - V_\infty)^2}{\dot{m}_f h_f} \quad (6.23)$$

For our purposes, we are more interested in the propulsive efficiency. *Propulsive efficiency* is given by a ratio of the useful power out (thrust times forward speed) to all the mechanical power produced (useful power out plus the power wasted in the wake):

$$\eta_p = \frac{TV_\infty}{TV_\infty + \frac{1}{2}(\dot{m} + \dot{m}_f)(V_e - V_\infty)^2} \quad (6.24)$$

If we use the simplified form of thrust (Eq. (6.21)) then the propulsive efficiency reduces to:

$$\eta_p = \frac{2}{1 + \frac{V_e}{V_\infty}} \quad (6.25)$$

We can now see that there is a fundamental tradeoff between high thrust and high efficiency. To produce lots of thrust we want a large exit velocity (V_e) relative to V_∞ (Eq. (6.21)). Unfortunately, the larger the exit velocity the worse the efficiency (Eq. (6.25)). For ideal efficiency a propulsion

system's exit velocity would equal V_∞ (and thus there would be no wake), but there would also be no thrust.

An alternative approach to increasing thrust, without penalizing efficiency, is to increase the mass flow rate (Eq. (6.21)). This means that we need a larger propeller or a larger engine inlet. Unfortunately, larger propellers/inlets have higher tip speeds that eventually will create shock waves with associated drag increases and efficiency losses. Furthermore, the faster we fly the worse this problem becomes. In other words, at higher speeds we are forced to have smaller inlets to avoid shock waves, and thus smaller mass flow rates.

Ultimately, these tradeoffs lead to the two ends of the spectrum depicted in Fig. 6.9. A turboprop, on the left, uses a small velocity increase with a large mass of air (high efficiency) to produce sufficient thrust. But because the propeller is large, flight speeds must be relatively slow. Conversely, a turbojet, on the right, uses a large velocity increase with a small mass of air (low efficiency). The small inlet allows for high flight speeds. A turbofan falls somewhere in the middle and is what is used on most commercial transports.

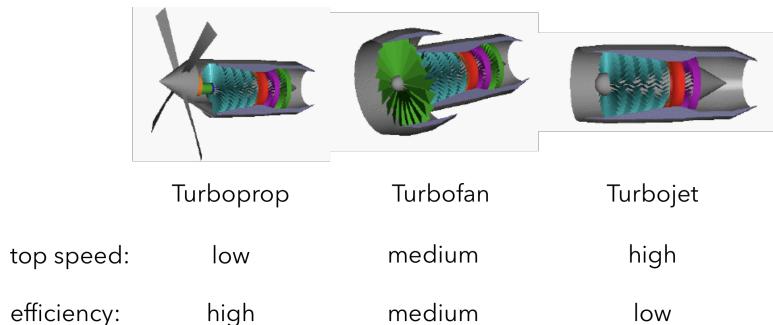


Figure 6.9: The range of gas turbine engines from turboprops to turbojets. Image adapted from [NASA](#), public domain.

In a turbofan, some of the air passes through the turbine and some passes through the fan but bypasses the core (Fig. 6.10). The *bypass ratio* is defined as:

$$\frac{\dot{m}_{bypass}}{\dot{m}_{core}} \quad (6.26)$$

the ratio of the mass flow rate that bypasses the core, over the mass flow rate through the core. The turboprop and turbojet are effectively extremes

of a turbofan where the turboprop has a very large bypass ratio and the turbojet has a bypass ratio of zero. Commercial transports continue to push the turbofan bypass ratio higher to increase efficiency while maintaining effectively the same incoming Mach number.

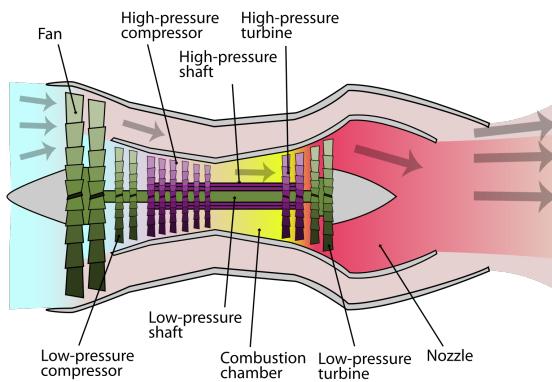


Figure 6.10: In a turbofan some of the air passes through the fan and into the core, and some passes through the fan but bypasses the core. Image by [K. Aainsqatsi](#), Wikimedia, CC BY-SA 3.0.

A gas turbine engine contains a compressor and a turbine and their blades are shaped very differently as depicted in Fig. 6.11. Furthermore, there are many stages in the compressor and only a few stages in the turbine as seen in Figs. 6.9 and 6.10. The reason for this is that air passing through the compressor is moving from a low pressure region to a high pressure region, or in other words through a strongly adverse pressure gradient. The airfoils can't carry much lift without separating, hence the lightly cambered airfoils and the need for many stages, each stage creating a small increase in pressure. Conversely, air passing through the turbine is moving from high to low pressure with a very favorable pressure gradient. The airfoils can be very aggressive with camber and thickness in order to generate high lift and extract maximum power from the airflow. The compressor performs work on the air and turns the shaft, whereas the turbine extracts work from the fluid. Net work is achieved because the work extracted at the turbine is at a much higher temperature (see the [Brayton cycle](#)).

We mentioned a turbojet as the extreme end of a turbofan, but there are actually more extreme options. As the Mach number gets past about Mach 3 a turbojet is no longer a useful option. A ramjet (Fig. 6.12) gets rid of the

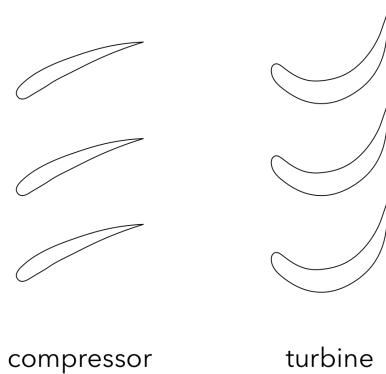


Figure 6.11: Typical blade shapes for the compressor and turbine sections of a gas turbine engine.

turbine and compressor altogether, and uses the high speed of the aircraft to compress the air. This configuration removes a lot of moving parts, but is relatively inefficient as high drag is created by the shock waves needed to slow the air down to subsonic speeds for combustion. Also, it only works at high Mach numbers so some other form of propulsion or assisted takeoff is needed to get the vehicle up to speed.

As the speed increases even further (approximately more than Mach 5) less compression is desirable so that the temperatures don't get too high and enough useful work can be extracted, and so a scramjet (supersonic combustion ramjet) is used. In this case the air stays supersonic all the way through. As you can imagine, maintaining combustion at supersonic speeds presents its own unique challenges. The Mach numbers that correspond to these various propulsion devices are summarized in Fig. 6.13. Earlier we defined the efficiency of an engine, however a more commonly used metric is the *specific fuel consumption* (*sfc*), also known as thrust specific fuel consumption. It is the ratio of fuel weight burned per unit time relative to thrust:

$$sfc = \frac{\dot{m}_f g}{T} \quad (6.27)$$

A lower *sfc* means that it takes less fuel to produce a specified amount of thrust. Typically *sfc* is expressed in units of $1/hr$. (As we will see, rockets use a metric called specific impulse which is the inverse of *sfc*, except usually expressed in seconds rather than hours).

We can see how *sfc* is related to its overall efficiency by using the

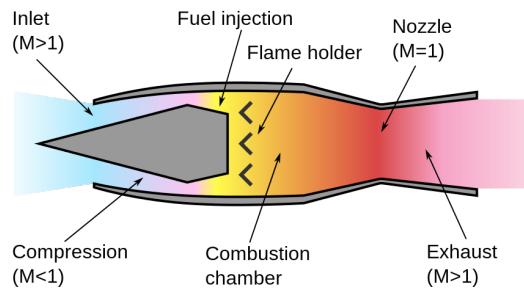


Figure 6.12: Depiction of a ramjet. Image by [Cryonic07](#), Wikimedia Commons, CC BY-SA 3.0.

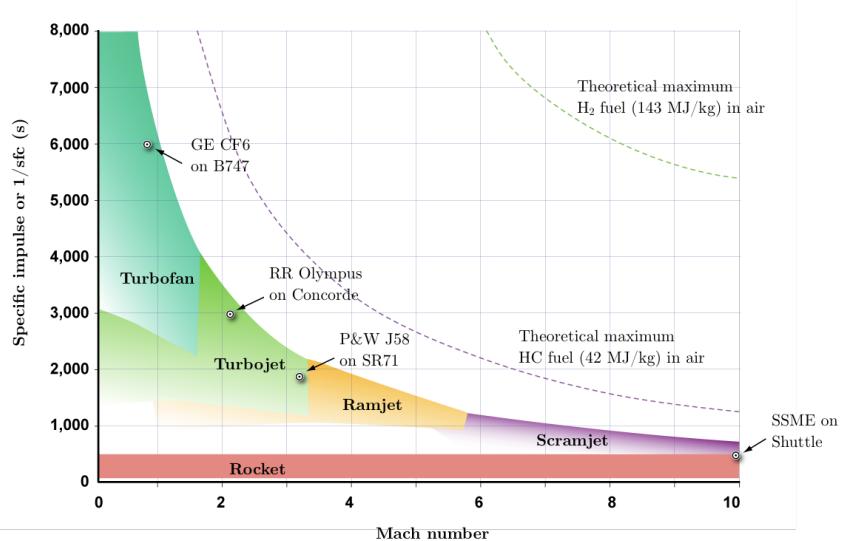


Figure 6.13: Mach number range for different propulsion devices. Modified from a figure by [Kashkhan](#), Wikimedia Commons, CC BY-SA 3.0

definition of overall efficiency (Eq. (6.22)) and replacing the thrust with that of the Eq. (6.27):

$$\eta = \frac{V_\infty g}{sfc h_f} \quad (6.28)$$

Thus, sfc and η are inversely related as we would expect.

Jet engines have enabled the wonders of modern air transport in part because jet fuel (and petroleum-based fuels in general) is so energy dense. A typical energy density is approximately 42 MJ/kg. That means just 1 kg of jet fuel contains approximately the amount of energy needed to send 1 kg of mass into orbit.

6.5 Propulsion System Placement

Beyond selecting a suitable motor/propeller combination, or a suitable gas turbine engine, one needs to determine how many propulsors to use and where to place them. Figure 6.14 shows two configurations, the one on the left is called a *single* and the one on the right a *twin*. The primary benefit of a single is simplicity. The main benefit of a twin is some redundancy allowing for the possibility of a recovery from an engine failure. You can also use counter-rotating props, which provide some benefit in terms of lateral trim drag. For small UAVs a twin is usually more crash resistant and for hand launches can allow you to be out of the blade path (especially for a pusher, discussed next). For piloted aircraft, a twin moves the propulsion systems out of the primary field of view.

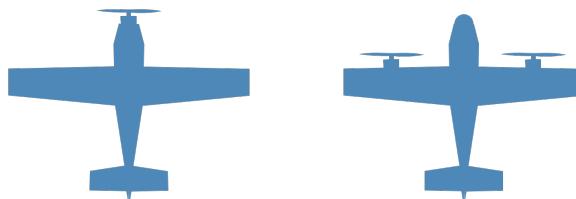


Figure 6.14: A single versus a twin propulsion system configuration.

While the historical trend has been to move from a large number of engines to a small number (generally 1–2), recent advances in battery technology, motor power density, IMU capabilities, and autonomous control are reversing that trend for electric-powered aircraft applications. Whereas jet engines are more efficient with fewer large systems, electric

motors can be (approximately) just as efficient with a large number of small motors. Having a large number of small motors can allow for greater redundancy, control, and effective mechanical actuation for applications like vertical takeoff. This type of system is called distributed propulsion (Fig. 6.15), and is an active area of research.



Figure 6.15: NASA X-57 Maxwell concept aircraft with 14 electric motors.
Image from [NASA](#), public domain.

Another major design decision, for a propeller-driven aircraft, is whether to use a *pusher* or *tractor* configuration. Figure 6.16 shows these two configurations with a pusher on the left and a tractor on the right. A pusher allows for improved visibility and directional stability (the propulsion system acts kind of like a vertical tail). Some primary disadvantages of a pusher are that it moves the center of gravity farther aft and is usually noisier because the prop ingests the wing wake. Efficiency tradeoffs are less clear. The wings on a pusher will be more efficient because they are not in the prop wash, but the propellers will be less efficient because they are in the wing wake (and vice-versa for a tractor).

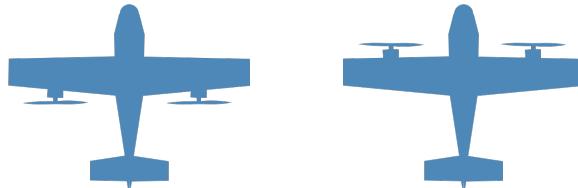


Figure 6.16: A pusher versus a tractor configuration.

Finally, placement of the propulsion system can be on the fuselage, on the

wing, in a pod, in the tail, etc. Determining proper placement requires an integrated look at aerodynamic performance, structures, stability, and other system-level considerations.

CHAPTER 7

Performance

Performance is an integration of the various subsystems into high-level metrics that describe how well the aircraft performs. Various metrics are important to the aircraft designer depending on the mission and flight envelope of the vehicle. This chapter discusses some of the more common performance metrics.

7.1 Range

Range is the maximum distance that an aircraft can fly. We will develop two expressions to estimate range: one for propulsion systems at constant thrust and one for systems at constant power. Recall from Chapter 6 that there are basically two extremes for propulsion systems: imparting a large velocity change to a small mass of air, or imparting a small velocity change to a large mass of air. The former corresponds to jet engines. Because the exit velocity is so large, the thrust is relatively insensitive to the freestream speed (Eq. (6.21)). In other words, the thrust is relatively constant with airspeed. The latter option corresponds to propeller systems. The thrust of propeller-driven systems varies significantly (generally decreasing with airspeed), but its power is relatively constant with changes in airspeed. While there are gas-powered piston engines with propellers, most modern uses of propellers are with electric power. Thus, we will only analyze

electric propeller-driven aircraft. In the following sections we will refer to jet powered and electric powered for convenience, but really mean constant thrust with fuel burn and constant power with no fuel burn respectively.

7.1.1 Jet Powered

We begin with estimating range for a jet-powered aircraft. Because the amount of fuel in the aircraft changes during the flight we need to start with the specific range. The *specific range* is the distance flown per unit weight of fuel burned. We can express the specific range as follows:

$$\text{specific range} = \frac{\text{distance flown}}{\text{fuel burned}} = \frac{\text{distance flown / time}}{\text{fuel burned / time}} \quad (7.1)$$

The numerator is the flight speed, and the denominator is the weight flow rate (mass flow rate times acceleration of gravity).

$$\text{specific range} = \frac{V_\infty}{\dot{m}_f g} \quad (7.2)$$

In Chapter 6 we saw the following definition for the specific fuel consumption (Eq. (6.27)):

$$sfc = \frac{\dot{m}_f g}{T} \quad (7.3)$$

Thus, we can express the fuel burned per unit time ($\dot{m}_f g$) as $sfc T$. So, the specific range is given by:

$$\text{specific range} = \frac{V_\infty}{sfc T} \quad (7.4)$$

If we assume steady, level flight then thrust equals the drag. Next, we multiply the equation by lift over weight, which should be equal for steady, level flight.

$$\text{specific range} = \frac{V_\infty}{sfc D} \frac{L}{W} \quad (7.5)$$

The range is the integration of specific range (distance/weight change) over the total change in weight across the flight. For a jet-powered aircraft the weight of the aircraft decreases through the flight, so the range is given by integrating the above expression from the initial weight (W_i) to the final weight (W_f):

$$R = \int_{W_i}^{W_f} \frac{V_\infty}{sfc D} \frac{1}{W} (-dW) \quad (7.6)$$

where a minus sign is used because the weight decreases from initial to final.

In the following derivation we will make the simplifying assumption that the specific fuel consumption, lift-to-drag ratio, and velocity are constant during the flight. The latter is equivalent to constant Mach number in the isothermal portion of the atmosphere. Making these assumptions allows us to pull those constants out of the integral.

$$R = \frac{V_\infty}{sfc} \frac{L}{D} \int_{W_f}^{W_i} \frac{dW}{W} \quad (7.7)$$

After performing the integral we have an expression that allows us to estimate range:

$$R = \frac{V_\infty}{sfc} \frac{L}{D} \ln\left(\frac{W_i}{W_f}\right) \quad (\text{jet powered}) \quad (7.8)$$

This is the well-known *Breguet range equation*

For comparing with electric vehicle later, it will be easier to specify the weights as a fuel fraction (weight of fuel relative to max total takeoff weight) and to use masses rather than weights. The initial mass is the total takeoff aircraft mass including fuel (m_{TO}), whereas the final mass is that minus the fuel mass $m_{TO} - m_f$. So the jet engine range estimates looks like:

$$R = \frac{V_\infty}{sfc} \frac{L}{D} \ln\left(\frac{1}{1 - m_f/m_{TO}}\right) \quad (\text{jet powered}) \quad (7.9)$$

As an aside, there are different range equations depending on what you choose to hold constant. In this case we have held velocity and C_L constant (so that L/D is constant). That means that the density is not constant ($C_L = W/(\frac{1}{2}\rho V^2 S_{ref})$) and would need to decrease through the flight as the weight decreases. In other words, the aircraft would need to increase in altitude while flying (a cruise-climb). This is usually a preferably way to fly as it is closer to constant Mach number, which is why we have used this form of the equation. Commercial transports can't perform a continuous cruise-climb because of air traffic control requirements, but for long flights do approximate it by shifting to higher altitude levels as the flight progresses. An alternative approach is to hold C_L and density (altitude) constant, leading

to a different estimate for range:

$$R = 2 \frac{V_\infty}{sfc} \frac{\sqrt{L}}{D} \left(\sqrt{W_i} - \sqrt{W_f} \right) \quad (7.10)$$

The Breguet range equation is simple to use and demonstrates that need for balance between aerodynamic efficiency (L/D), propulsive efficiency (sfc), and structural efficiency (m_f/m_{TO}). The first two sub-metrics we have seen already, but the “structural efficiency” is new. The most ideal structure would be one that is 100% fuel (i.e., no structure) and at the end of flight your airplane would weigh nothing. That would make W_i/W_f infinite, which means that your range was infinite in the sense that if your airplane needed no structure, just fuel, you could add an infinite amount of fuel to fly forever. Of course, that is not possible, some finite amount of structure is always needed, and an all-fuel aircraft would be useless because it couldn’t carry any payload. For a given amount of fuel, the lighter the structure is, then the better the structural efficiency is.

Another way to express this range equation is in terms of the overall efficiency of the propulsion system. Recall from Chapter 6 that the overall efficiency was related to sfc by (Eq. (6.28)):

$$\eta = \frac{V_\infty g}{sfc h_f} \quad (7.11)$$

where h_f is the enthalpy of the fuel. We now solve this for V_∞/sfc and substitute that into the Breguet range equation:

$$R = \frac{h_f}{g} \eta \frac{L}{D} \ln \left(\frac{1}{1 - m_f/m_{TO}} \right) \quad (7.12)$$

This form of the equation is arguably more useful as the enthalpy and max efficiency of jet engines are more universal than V_∞ or even sfc .

These equations are very useful for preliminary design, but recall that multiple assumptions were made. In detailed design we could relax the assumptions on steady-level flight and on constant properties. Instead, we would perform a numerical integration through the flight trajectory to estimate range.

7.1.2 Electric Powered

Estimating the range of an electric powered vehicle is a little more straightforward because the mass of the vehicle doesn't change through the flight. Batteries are often characterized by their specific energy, which is their energy per unit mass (e_b). The total energy available from the battery then is the specific energy, times its mass, times its thermal efficiency (as there are some losses).

$$E_{available} = e_b m_b \eta_b \quad (7.13)$$

Note that e_b should be an effective specific energy for the battery that accounts for losses from temperature, depth of discharge, cycle life, etc. The energy expended from the propulsion system¹

$$E_{expended} = \frac{TR}{\eta_p \eta_m} \quad (7.14)$$

where R is the range, T is the average thrust across the flight, η_p is the propeller efficiency, and η_m is the motor efficiency. For steady-level flight the thrust equals the drag so

$$E_{expended} = \frac{DR}{\eta_p \eta_m} \quad (7.15)$$

We can now equate the energy available to the energy expended and solve for range:

$$R = \frac{e_b m_b \eta}{D} \quad (7.16)$$

where $\eta = \eta_b \eta_m \eta_p$ is the total efficiency of the propulsion system (battery, motor, propeller, etc.). To put this in a similar form to that of a jet engine we will do the same thing we did in the previous section: multiply by lift over weight assuming steady, level flight:

$$R = \frac{e_b m_b \eta}{D} \frac{L}{W} = \frac{e_b m_b \eta}{D} \frac{L}{m_{TO}} \quad (7.17)$$

where W is the total mass of the aircraft including battery. Rearranging we have:

$$R = \frac{e_b}{g} \eta \frac{L}{D} \frac{m_b}{m_{TO}}$$

(7.18)

¹Actually, this is the minimum expended energy as there are various subsystems like lights, instrumentation, etc. that are also powered by the battery.

There are a lot of similarities between this equation and that of a jet-powered aircraft (Eq. (7.12)). Both have an efficiency term, lift-to-drag ratio, and a structural efficiency. The structural efficiency follows essentially the same role here: an ideal structure would be all battery (excluding some mass for payload). However, the jet engine has some advantage here as burning fuel during the flight decreases the aircraft weight. Figure 7.1 compares the two “efficiencies” (note this isn’t really an efficiency). A typical fuel fraction is around 40% or less so this generally isn’t a major difference.

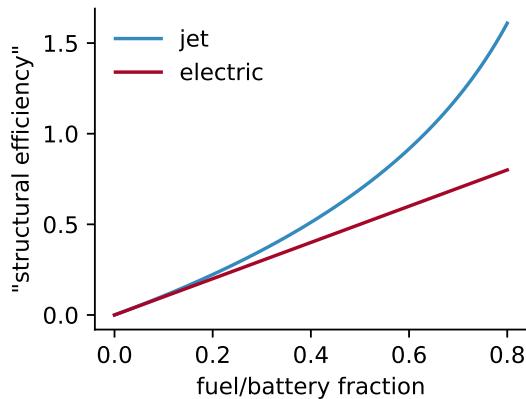


Figure 7.1: Comparison of m_b/m_{TO} for a battery versus $\ln(1/(1 - m_{fuel}/m_{TO}))$ for various mass fractions of battery/fuel.

Peak efficiencies for the two systems also differ. An efficient jet engine has around a 40% total efficiency, whereas an electric system can be approximately twice as efficient.

However, both of these differences are relatively small compared to the last difference: the enthalpy of jet fuel compared to battery specific energy. The enthalpy of jet fuel is about 42.8 MJ/kg. Today’s best batteries (at scale) have a specific energy of approximately 300 Wh/kg or 1.08 MJ/kg. In other words, jet fuel is amazingly energy dense, about 40 times that of a battery! This factor explains the major difference in range capabilities between the two types of propulsion.

Electric propulsion can’t compete on range, but it does have its advantages: no emissions, less noise, cheaper, fewer moving parts, higher efficiency, efficient at small sizes to allow for distributed propulsion, etc.

Thus, the current interest for short range applications, especially with vertical takeoff and landing (VTOL).

7.2 Endurance

Endurance is similar to range, but rather than a measure of flight distance, *endurance* measures flight time. For applications like search and rescue, time in the air is more important than distance flown.

Specific endurance is time flown per unit weight of fuel burned:

$$\begin{aligned}
 \text{specific endurance} &= \frac{\text{time}}{\text{fuel burned}} \\
 &= \frac{1}{\text{fuel burned} / \text{time}} \\
 &= \frac{1}{\dot{m}_f g} \\
 &= \frac{1}{sfc T}
 \end{aligned} \tag{7.19}$$

Note that this is the same as specific range divided by V_∞ . Following the same procedure as the previous section yields:

$$t = \frac{1}{sfc D} \frac{L}{g} \ln \left(\frac{1}{1 - m_f/m_{TO}} \right) = \frac{h_f}{g} \frac{\eta}{V_\infty} \frac{L}{D} \ln \left(\frac{1}{1 - m_f/m_{TO}} \right) \quad (\text{jet power}) \tag{7.20}$$

$$t = \frac{e_b}{g} \frac{\eta}{V_\infty} \frac{L}{D} \frac{m_b}{m_{TO}} \quad (\text{electric}) \tag{7.21}$$

where t is the endurance.

7.3 Optimal Lift Coefficients

In this section we will explore the approximate optimal lift coefficient to fly at for different metrics. For a given altitude, this corresponds to an optimal speed. This is easy to see using the definition of the lift coefficient

and rearranging:

$$C_L = \frac{L}{\frac{1}{2}\rho V^2 S_{ref}} \quad (7.22)$$

$$V = \sqrt{\frac{2L}{\rho S_{ref} C_L}} \quad (7.23)$$

Lift coefficient is a more fundamental parameter, because it is nondimensional, so all derivations are in terms of lift coefficient even though we often speak in terms of speed. All the estimates in this section are simplified to allow for analytic formulas. This type of approach is helpful to build intuition and to understand what parameters matter most, but as the aircraft design progresses a numerical computation of optimal lift coefficients should be used rather than these analytic formulas.

Lift to Drag Ratio. The minimum drag speed of an aircraft is of course zero. More interesting is the minimum drag speed for a fixed lift, which is equivalent to maximizing the lift-to-drag ratio. If the compressibility drag is negligible then the drag coefficient can be expressed as:

$$C_D = C_{Dp} + \frac{C_L^2}{\pi A Re} \quad (7.24)$$

and the lift to drag ratio is then:

$$\frac{L}{D} = \frac{C_L}{C_D} = \frac{C_L}{C_{Dp} + \frac{C_L^2}{\pi A Re}} \quad (7.25)$$

To find the lift coefficient that maximizes this expression we need to take the derivative of this equation, with respect to C_L , and set the result equal to zero (note that it's a little simpler to instead minimize D/L rather than maximize L/D).

$$\frac{d(L/D)}{dC_L} = \frac{C_{Dp} + \frac{C_L^2}{\pi A Re} - C_L \frac{2C_L}{\pi A Re}}{\left(C_{Dp} + \frac{C_L^2}{\pi A Re}\right)^2} = 0 \quad (7.26)$$

Thus,

$$C_{Dp} + \frac{C_L^2}{\pi A Re} - 2 \frac{C_L^2}{\pi A Re} = 0 \quad (7.27)$$

$$C_{Dp} = \frac{C_L^2}{\pi A Re} \quad (7.28)$$

$$(7.29)$$

Or in other words, the lift coefficient that maximizes L/D occurs when the parasitic drag and induced drag are equal. Solving for the lift coefficient yields:

$$C_{L\max L/D} = \sqrt{C_{Dp} \pi A Re} \quad (7.30)$$

The above derivation assumed that both e and C_{Dp} are independent of C_L , though neither is strictly true. The inviscid span efficiency changes with lift coefficient, but usually not by much for modest changes in lift coefficient. The parasitic drag coefficient isn't a function of C_L per se. However, both depend on Reynolds number and so because lift is fixed they are implicitly related:

$$\frac{dC_{Dp}}{dC_L} = \frac{\partial C_{Dp}/\partial Re}{\partial C_L/\partial Re} \quad (7.31)$$

Both depend on Mach number as well, but we assume that is negligible consistent with our assumption that compressibility drag is negligible. A numerical simulation, evaluating drag at each Reynolds number and Mach number will yield a more accurate estimate, but the above formula is useful for conceptual design purposes.

Range Jet Power. Recall that the specific range for a jet aircraft is:

$$V_\infty \frac{L}{D} \frac{1}{sfc W} \quad (7.32)$$

Both weight and sfc are independent of the lift coefficient (if at constant altitude— sfc changes some with density), and so to maximize range we need to maximize

$$V_\infty \frac{L}{D} = V_\infty \frac{C_L}{C_D} \quad (7.33)$$

We can infer already that the speed for maximum range will be faster than the speed for minimum drag. That is because the presence of the V_∞ term means that increasing speed has a linear effect whereas the decrease in L/D for increasing speed beyond the max L/D speed will initially be small. Using Eq. (7.23) and assuming steady level flight ($L = W$) the metric we need to maximize becomes:

$$\sqrt{\frac{W}{\frac{1}{2}\rho S_{ref} C_L}} \frac{C_L}{C_D} \quad (7.34)$$

For a given altitude, removing terms that are independent of C_L reduces the above expression to:

$$\frac{\sqrt{C_L}}{C_D} \quad (7.35)$$

Thus, to maximize range we need to maximize the above expression, or minimize its inverse, with respect to C_L . We will do the latter:

$$\frac{d \frac{C_D}{\sqrt{C_L}}}{d C_L} = \frac{d}{d C_L} \left(\frac{C_D p}{C_L^{1/2}} + \frac{C_L^{3/2}}{\pi A Re} \right) \quad (7.36)$$

$$= -\frac{1}{2} \frac{C_D p}{C_L^{3/2}} + \frac{3}{2} \frac{C_L^{1/2}}{\pi A Re} = 0 \quad (7.37)$$

$$C_D p = 3 \frac{C_L^2}{\pi A Re} \quad (7.38)$$

In other words, for maximum range the parasitic drag is three times the induced drag.

$$C_{L \text{max range, jet}} = \sqrt{\frac{C_D p \pi A Re}{3}} = \frac{1}{\sqrt{3}} C_{L \text{max L/D}} \quad (7.39)$$

We see that indeed the optimal lift coefficient is lower than of max L/D , or referring to Eq. (7.23), the maximum range speed is $3^{1/4}$ times the maximum L/D speed. In other words, the maximum range speed is about 32% faster.

Again, keep in mind, that some simplifying assumptions were made like assuming constant sfc , and so a numerical solution is preferable when possible. Still these analytic solutions are useful in lending insight into what parameters matter most and in determining approximate optimal conditions.

Range Electric Power. For an electric powered vehicle all the terms in the range equation (Eq. (7.18)) are constant with lift coefficient except $\eta L/D$. Unfortunately, the efficiency term is not easily amenable to an analytic expression. Generally it varies weakly with flight speed, as compared to L/D , and so for the purpose of this analysis we assume constant efficiency. If efficiency is constant with C_L , then C_L for maximum range under electric power is the same as the C_L for maximum L/D .

$$C_{L\max \text{ range, elec}} \approx C_{L\max L/D} \quad (7.40)$$

Again, remember that this expression is only true if propeller efficiency is constant. In practice, the maximum range speed is usually somewhat faster than the maximum L/D speed. If the propulsion system has been identified then a numerical solutions should be sought.

Endurance Jet Power. For a jet, the only terms in the endurance equation (Eq. (7.20)) that vary with C_L are L/D (again ignoring small variations in sfc). That means that speed for max endurance is the same as the speed for max L/D .

$$C_{L\max \text{ endurance, jet}} = C_{L\max L/D} \quad (7.41)$$

Endurance Electric Power. For an electric powered aircraft to maximize endurance we must maximize the quantity

$$\frac{\eta}{V} \frac{L}{D} \quad (7.42)$$

Again, we make the same approximation that η is constant. In that case, to maximize endurance we need to maximize the quantity:

$$\frac{1}{V} \frac{L}{D} \quad (7.43)$$

Using a similar procedure to that of maximizing range of a jet-powered aircraft, we find that for optimizing endurance we need to minimize the quantity:

$$\frac{C_D}{C_L^{3/2}} \quad (7.44)$$

Taking the derivative with respect to C_L gives:

$$\frac{d \frac{C_D}{C_L^{3/2}}}{d C_L} = \frac{d}{d C_L} \left(\frac{C_{Dp}}{C_L^{3/2}} + \frac{C_L^{1/2}}{\pi A Re} \right) \quad (7.45)$$

$$= -\frac{3}{2} \frac{C_{Dp}}{C_L^{5/2}} + \frac{1}{2} \frac{C_L^{-1/2}}{\pi A Re} = 0 \quad (7.46)$$

$$C_{Dp} = \frac{1}{3} \frac{C_L^2}{\pi A Re} \quad (7.47)$$

In other words, for maximum endurance the parasitic drag is one third of the induced drag.

$C_{L_{\text{max endurance, elec}}} \approx \sqrt{3 C_{Dp} \pi A Re} = \sqrt{3} C_{L_{\text{max L/D}}} \quad (7.48)$

We see that indeed the optimal lift coefficient is higher than of max L/D , or referring to Eq. (7.23), the maximum endurance speed is $3^{-1/4}$ times the maximum L/D speed. In other words, the maximum endurance speed is about 24% slower. As noted previously, a numerical solution is preferably if the various in efficiency is known, though in general the max endurance speed will be slower than the max L/D speed.

7.4 Rate of Climb

When climbing, the lift, drag, and thrust vectors tilt as shown in Fig. 7.2, whereas weight remains pointed downward. A force balance along the direction of flight yields:

$$T = D + W \sin \gamma \quad (7.49)$$

where γ is the climb angle. Rearranging the expression gives:

$$\sin \gamma = \frac{T - D}{W} \quad (7.50)$$

The *rate of climb* is the speed we climb vertically. This is given by the component of our flight speed in the vertical direction:

$$V_\infty \sin \gamma \quad (7.51)$$

Using the previous equation gives the rate of climb:

$$R/C = \frac{V_\infty(T - D)}{W} \quad (7.52)$$

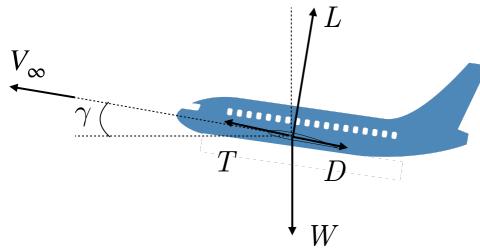


Figure 7.2: Depiction of forces on aircraft during a steady climb..

As discussed in the beginning of this chapter, for a jet-powered aircraft thrust is mostly independent of speed, so a good metric of climb performance is the thrust-to-weight ratio: T/W . For electric-powered aircraft power is mostly independent of speed, so a good metric of climb performance is the power-to-weight ratio: P/W or TV/W . For jet power, the speed that maximizes the rate of climb is faster than the speed for max L/D . For an electric powered aircraft, since power is constant, the speed that maximizes rate of climb is the minimum power speed (which is the same as the speed for maximum endurance). As long as the flight path angle, γ , is not large, in conceptual design we usually assume that the aircraft drag is unchanged from its level flight value.

Maximizing rate of climb means minimizing the time it takes to reach a certain altitude. However, various other climb metrics are sometimes used including maximizing the gradient of climb (a slower speed) or minimizing the fuel/energy used to climb (a faster speed with a lower rate of climb).

Another related metric is the ceiling. The *absolute ceiling* is the highest altitude the aircraft can fly at, and is the altitude at which the rate of climb equals zero. As the altitude increases, density decreases, and the speed to maintain level flight increases. This increase in speed requires more power. Additionally, for jet engines the maximum power available tends to decrease with altitude, and propeller performance is affected by density as well. In either case, there is some altitude at which there will be no excess power and the aircraft will no longer be able to climb. A related, and

perhaps more useful metric is the *service ceiling*. This is the highest altitude at which a given rate of climb can be sustained. The target climb rate is often around 30 meters/minute for larger aircraft.

7.5 Glide Ratio

Gliding flight if when a heavier-than-air aircraft flies without any thrust. The *glide ratio* refers to the ratio of forward distance traveled relative to vertical distance descended. A high glide ratio is desirable because that means for a given starting altitude, the aircraft can glide further. We can determine the glide ratio starting from the same equation used previously in climbing (Eq. (7.50)). The only difference is that this time there is no thrust:

$$\sin \gamma = -\frac{D}{W} \quad (7.53)$$

The negative sign denotes that the flight angle is decreasing rather than increasing as in the climb case. We will drop the negative sign going forward, with the understanding that the angle is now positive down. Referring to Fig. 7.2 again, a force balance perpendicular to the flight direction yields:

$$L = W \cos \gamma \quad (7.54)$$

Solving for weight and substituting into the previous equation gives:

$$\sin \gamma = \frac{D}{L \cos \gamma} \quad (7.55)$$

After rearranging we have:

$$\tan \gamma = \frac{D}{L} \quad (7.56)$$

The glide ratio as stated earlier is the ratio of forward distance relative to vertical distance. From Fig. 7.3 we can see that this is precisely $1/\tan \gamma$:

$$\text{glide ratio} = \frac{1}{\tan \gamma} = \frac{L}{D} \quad (7.57)$$

In other words the glide ratio is the same as the lift-to-drag ratio.

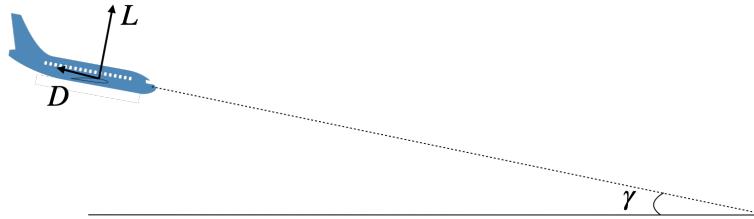


Figure 7.3: Depiction of a gliding aircraft showing that the glide ratio (forward distance / vertical distance) is exactly the lift to drag ratio.

7.6 Turning Radius

As discussed in Section 5.4, turning generally occurs by banking the aircraft in a coordinated turn. Figure 7.4 depicts how tilting at a bank angle (Φ) allows a component of the lift to turn the aircraft through a centripetal acceleration. A force balance in the vertical direction gives:

$$L \cos \Phi = mg \quad (7.58)$$

Applying Newton's second law in the horizontal direction, using the form for centripetal acceleration, gives:

$$L \sin \Phi = m \frac{V^2}{R_t} \quad (7.59)$$

where R_t is the turning radius. If we combine these two equations we can solve for the turning radius:

$$R_t = \frac{V^2}{g \tan \Phi}$$

(7.60)

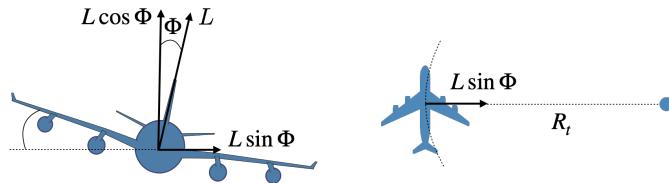


Figure 7.4: Depiction of an aircraft turning to allow determination of the turning radius.

We can't do much about the acceleration of gravity so if we want to able to make tighter turns we need to either fly slower (decrease stall speed) or increase our bank angle. For small UAVs a fairly extreme bank angle is acceptable, but would be much more limited in passenger carrying aircraft.

7.7 High Lift Devices

Takeoff and landing performance are primarily functions of stall speed. Using Eq. (7.23) we see that the stall speed occurs once C_L reaches its highest value: C_{Lmax} .

$$V_{stall} = \sqrt{\frac{2W}{\rho S_{ref} C_{Lmax}}} \quad (7.61)$$

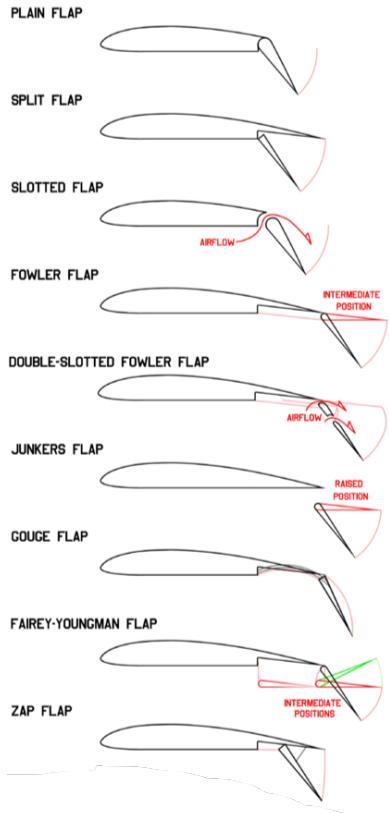
This equation tells us what we can do if we want a slower stall speed and thus a shorter runway for takeoff/landing. We can't control air density, but this equation does explain why we need a longer runway at airports at high altitudes. Decreasing stall speed is a common reason for increasing wing area, but that choices comes with other design compromises. Clearly, increasing C_{Lmax} has a major effect on decreasing stall speed.

Unfortunately, designing for high lift means that drag will be high too. What we really want is variable geometry: a high-lift configuration for takeoff/landing, and a streamlined configuration for cruise. This is exactly how transport aircraft operate.

There are two primary mechanisms for changing the geometry of a wing to increase its maximum lift coefficient: *flaps* and *slats*. Both are used in combination on transport aircraft. A variety of flaps are depicted in Fig. 7.5a, with a picture of one in Fig. 7.5b.

Flaps increase the effective camber of the airfoil, and thus the zero lift angle of attack (Fig. 7.6). Notice that for a given angle of attack the lift is increased, and that the angle of attack for maximum lift is reduced. Notice also that using flaps has almost no effect on c_{lmax} . The primary benefit of flaps is that maximum lift occurs at a lower angle of attack, which helps with pilot visibility and decreases the amount of necessary fuselage upsweep (the aft end of the fuselage curves up to increase ground clearance during takeoff rotation).

A slotted flap enables higher lift through two mechanisms. First, it reenergizes the boundary layer allowing for attached flow along the flap at



(a) A variety of flap options. Figure by NiD.29, Wikimedia Commons, CC BY-SA 3.0.



(b) A flap system on a transport aircraft. Picture by David Monniaux, Wikimedia Commons, CC BY-SA 3.0.

Figure 7.5: Trailing edge flaps.

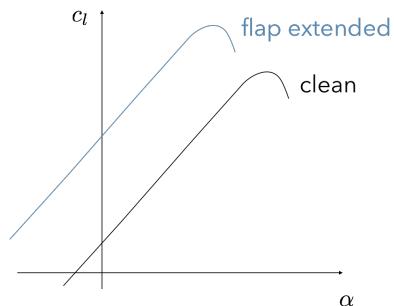
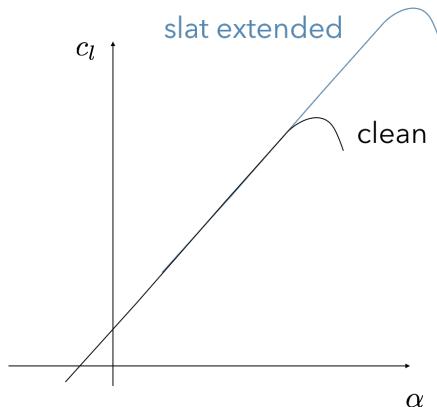


Figure 7.6: Flaps change the effective camber, and thus the zero lift angle of attack. They have almost no effect on c_{lmax} however.

higher angles of attack. Second, it increases the pressure experienced at the aft end of the main wing allowing for a less adverse pressure gradient. The other common high-lift device is a slat, which occurs on the leading edge (Fig. 7.7a) rather than the trailing edge. As seen in Fig. 7.7b, slats have little effect over most of the lift curve. Instead, they extend the range of angles of attack over which the flow remains attached, thus delaying stall. Like the difference between a pointed airfoil and a round-nose airfoil, the leading edge shape allows for attached flow across a wider range of angles of attack. Slats work well in combination with flaps by extending the maximum lift coefficient, while keeping the angle of attack manageable. Flaps generally do not extend along the full wing, but slats do. The reason for this is that slats change the maximum lift coefficient and the aircraft can't take advantage of that extra lift unless it is extended across the full span, especially at the tips. For example, Fig. 7.8 shows an example where slats did not extend all the way to the wingtips. The extra possible lift at the root can't be utilized because the tip will stall first.



(a) Slat on the leading edge of a wing.
Picture from [Arpingstone](#), Wikipedia,
public domain.



(b) Slats have almost no effect on camber,
instead the increase the maximum section
lift coefficient.

Figure 7.7: Picture of a slat and a notional diagram of its effect on lift.

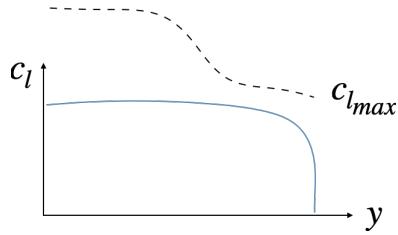


Figure 7.8: An example where slats extended only over a portion of the wing. In that case the extra increase in c_{lmax} is unusable.

7.8 Takeoff

We will discuss only the main principles involved in determining takeoff field length. For further details you are referred to the Federal Aviation Regulations (FARs) See for example, [Part 25 for transport aircraft](#).

One critical speed during takeoff is called the *decision speed* and is denoted as V_1 . If you have an engine failure or other critical failure before the decision speed then you abort the takeoff and put on full brakes. This is called the accelerate-stop distance. If you have a failure after the decision speed you must takeoff (there won't be enough runway to brake), circle around and land or divert to an alternate airport. This is called the accelerate-climb distance. Takeoff field length considers not just the runway, but includes the total distance until the airplane is some height above the ground (Fig. 7.9), which is 35 feet for transport category airplanes and 50 feet for general aviation.

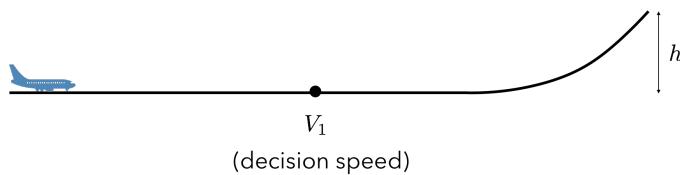


Figure 7.9: A critical factor in takeoff field length is the decision speed between braking and climbing. Takeoff distance is defined as not just the ground roll portion, but also the distance required to clear an obstacle of height h .

The takeoff distance is determined from the balanced field length. The *balanced field length* is the point at which the accelerate-stop distance and accelerate-climb distance are equal. This is the shortest possible field

length for safe takeoff of that aircraft.

Figure 7.10 illustrates this concept. The gray line shows how much runway you would need to accelerate to the candidate value of V_1 . The vertical distance from the gray to the blue curve corresponds to the extra distance required to stop assuming you decided to stop at the decision speed. Intuitively it makes sense that if we decide to stop at a later (higher) speed, then we will need more distance to come to a stop. hence the blue line moves further and further away from the gray.

The vertical distance from the gray to the red curve corresponds to the distance required to clear the required obstacle height assuming you decided to takeoff. This line drops down because the decision point corresponds to an engine failure event. If we fly further with two engines, it'll take less runway to reach takeoff speeds with our remaining engines. Conversely, if we have a failure very early, and decide we're still going to takeoff anyway, it'll take more distance to get up to speed on our one engine. Hence the red line is further away at low speeds.

Clearly, it is not ideal to decide to takeoff if an engine failed early while still traveling slowly. Conversely, it is not ideal to decide to brake if we are moving quickly. The minimum distance that balances these considerations is the crossover point of the two curves and is called the balanced field length. At that point the distance to stop and the distance to takeoff are identical. This point defines the actual value of V_1 . If we are below this speed then its requires less distance to stop, and if we are above this speed then it takes less distance to continue the takeoff.²

There are too many uncertainties to predict takeoff field length accurately from an equation, but we can work out what parameters matter and provide an expression that is useful in making relative comparisons in takeoff field length. Recall that we assume that jet powered aircraft operated at constant thrust, whereas electric propeller-driven aircraft operate at constant power.

²This article provides a good overview of these concepts if you'd like a little more detail without going all the way to the FARs.

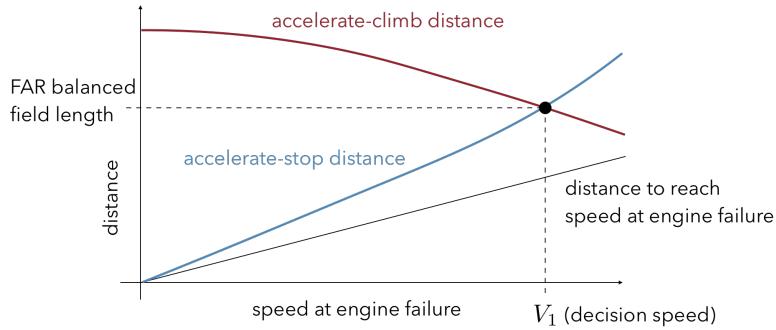


Figure 7.10: An illustration of the balanced field length where the distance to stop and the distance to climb are equal.

Takeoff Jet Power. We start by rearranging Newton's second law and setting up an integral from the start of the ground run to some time t :

$$F = m \frac{dV}{dt} \quad (7.62)$$

$$dt = \frac{m}{F} dV \quad (7.63)$$

$$\int_0^t dt = \int_0^V \frac{m}{F} dV \quad (7.64)$$

We make the assumption that the net force is independent of speed, which is reasonable if thrust is much larger than drag (drag itself is a strong function of speed). Then the integral simplifies to:

$$t = \frac{m}{F} V \quad (7.65)$$

We now perform a second integral to get distance (d) traveled and integrate from the start of the ground run until lift off (LO). We also make use of Eq. (7.65) in order to express velocity as a function of time:

$$d_{LO} = \int_0^{t_{LO}} V dt \quad (7.66)$$

$$d_{LO} = \int_0^{t_{LO}} \frac{F t}{m} dt \quad (7.67)$$

Again, we make the same assumption of constant force (net thrust) when performing the integral:

$$d_{LO} = \frac{F}{m} \frac{t_{LO}^2}{2} \quad (7.68)$$

Or, using Eq. (7.65) the distance to liftoff is:

$$d_{LO} = \frac{mV_{LO}^2}{2F} \quad (7.69)$$

We can reexpress this equation in terms of variables more usable to the aircraft designer by noting that $m = W/g$, the force is the net thrust ($F = T - D$), and the minimum lift-off speed is 20% greater than V_{stall} :

$$V_{LO} = 1.2 \sqrt{\frac{2W}{\rho S_{ref} C_{Lmax}}} \quad (7.70)$$

Plugging these alternate parameters into our ground roll expression gives:

$$d_{LO} = 1.44 \frac{W^2}{g \rho S_{ref} C_{Lmax} (T - D)} \quad (7.71)$$

Drag in this case is particularly difficult to predict. It's not the same as drag during level flight, but needs to include ground effect and rolling friction including some alleviation in the normal force from lift:

$$D = \mu(W - L) + D_{aero} \quad (7.72)$$

Instead, we will assume that thrust is much bigger than drag (which is usually true and which we have already implicitly assumed by using a constant force in the integrals), then takeoff distance should be proportional to the remaining parameters (removing constants):

$$d_{LO} \propto \frac{W^2}{\rho S_{ref} C_{Lmax} T} \quad (7.73)$$

Again, this procedure does not give us an accurate predictive function per se, but it does, for example, tell us that an aircraft that is twice as heavy will approximately require four times the takeoff length. Note that this discussion has only focused on the ground roll distance (omitting the climb over obstacle portion). For a transport aircraft the ground roll is most of the total length (about 80%) so often it is the ground roll that is focused on during conceptual design.

Electric Powered Takeoff. We follow the same procedure as before, but now assume constant power rather than constant thrust. In the first integral:

$$\int_0^t dt = \int_0^V \frac{m}{F} dV \quad (7.74)$$

the force is not constant, but if we multiply top and bottom by velocity we will have power:

$$\int_0^t dt = \int_0^V \frac{mV}{P} dV \quad (7.75)$$

Assuming power is constant during takeoff:

$$t = \frac{m}{P} \frac{V^2}{2} \quad (7.76)$$

The second integral for distance, replacing V with the above expression gives:

$$d_{LO} = \int_0^{t_{LO}} V dt \quad (7.77)$$

$$= \int_0^{t_{LO}} \sqrt{\frac{2Pt}{m}} dt \quad (7.78)$$

$$= \sqrt{\frac{2P}{m}} \frac{2}{3} t_{LO}^{3/2} \quad (7.79)$$

Using Eq. (7.76) to replace time in the above expression gives:

$$d_{LO} = \frac{mV_{LO}^3}{3P} \quad (7.80)$$

We make the same substitutions as before for stall speed giving:

$$d_{LO} = 1.629 \frac{W^{5/2}}{g(\rho S_{ref} C_{Lmax})^{3/2} P_{net}} \quad (7.81)$$

where P is the net power including drag, though again we will usually assume the power from thrust is much larger.

$$d_{LO} \propto \frac{W^{5/2}}{(\rho S_{ref} C_{Lmax})^{3/2} P} \quad (7.82)$$

Unlike transport aircraft, for electric powered aircraft the climbing phase may be longer than the ground roll phase and thus would be important to include even in conceptual design. However, the details of transitioning from ground roll to climb are outside our scope.

7.9 Landing

Landing field length is perhaps even more difficult to predict because of the large number of uncertainties, but as we will see it is primarily a function of stall speed. Landing consists of three phases: glide, deceleration, and ground roll (Fig. 7.11).

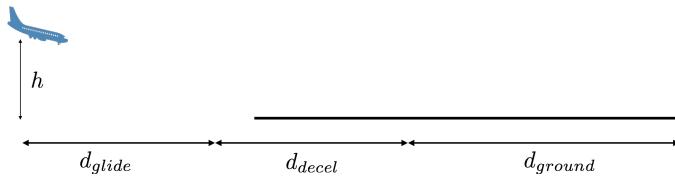


Figure 7.11: Notional depiction of landing with a gliding phase, deceleration phase, and a ground roll phase.

During the gliding phase, we can use the glide ratio we derived earlier:

$$d_{glide} = h \left(\frac{L}{D} \right)_{eff} \quad (7.83)$$

where h is the obstacle height (50 feet for transport aircraft under Part 25), and $(L/D)_{eff}$ is an effective lift-to-drag ratio including ground effect.

The deceleration phase is where the aircraft is just above the ground, but not yet touching the ground. Deceleration occurs from aerodynamic drag. If we assume constant deceleration:

$$d_{decel} = \frac{V_g^2 - V_L^2}{2a} \quad (7.84)$$

$$= \frac{V_g^2 - V_L^2}{2 \frac{D_{eff}}{W/g}} \quad (7.85)$$

$$= \left(\frac{L}{D} \right)_{eff} \frac{1}{2g} (V_g^2 - V_L^2) \quad (7.86)$$

$$(7.87)$$

where V_g is the speed at the end of the glide phase, and V_L is the landing speed before touchdown.

The ground phase is where the wheels are on the ground and the brakes

are applied. Assuming constant deceleration from friction:

$$d_{decel} = \frac{V_L^2}{2a} \quad (7.88)$$

$$= \frac{V_L^2}{2(\mu(W - L) + D)/m} \quad (7.89)$$

$$= \frac{V_L^2 W}{2g(\mu(W - L) + D)} \quad (7.90)$$

$$(7.91)$$

where the drag is now higher with spoilers deployed.

Spoilers are a critical component of a landing system (Fig. 7.12). They increase aerodynamic drag, but more importantly they drastically decrease lift allowing for a bigger normal force and thus higher rolling resistance from the brakes. The total drag is a combination of rolling drag and aerodynamic drag. The rolling drag has the potential to be much bigger so it is important to decrease lift quickly to allow the brakes to slow the aircraft down.



Figure 7.12: Spoilers on an aircraft. Picture from [Richardgm](#), Wikimedia Commons, public domain.

Adding all the terms we see that landing distance is of the form (where A and B are some constant independent of speed):

$$d = A + BV_L^2 \quad (7.92)$$

and because the landing speed is proportional to the stall speed (where C is a different constant):

$$d = A + CV_s^2 \quad (7.93)$$

In the end, landing distance is primarily a function of stall speed. In practice, pilot skill in touching down quickly so that spoilers and brakes can be deployed, perhaps makes the biggest difference.

CHAPTER 8

Structures

The aircraft structure must be designed to safely carry all loads experienced during flight. At the same time the structure should be as small and light as possible so as to not increase the required lift or wetted area. Aerodynamics and structures are often at odds and so careful design is necessary to achieve an optimal balance.

8.1 Typical Structural Elements

Most aircraft components are built using a *semi-monocoque* structure. A monocoque structure is one in which loads are carried through an external skin (e.g., an egg shell). A semi-monocoque structure uses a shell structure but with reinforced frames. This type of design can be seen in the fuselage shown in Fig. 8.1.

Figure 8.2 highlights some of the major structural components of the fuselage. The frames provide the shape and also increase buckling strength. The long skinny members between frames carry longitudinal loads and are called stringers or longerons. Often the terms are used interchangeably, but do have a small difference in meaning.¹ The skin (besides providing a covering), primarily carries torsion and shear loads.

¹This [Wikipedia](#) article discusses the difference.



Figure 8.1: Semi-monocoque structure visible inside aircraft fuselage. Picture by [YSSYguy](#), Wikimedia Commons, CC BY-SA 4.0.

These same elements are visible in the picture of the fuselage seen previously (Fig. 8.1).

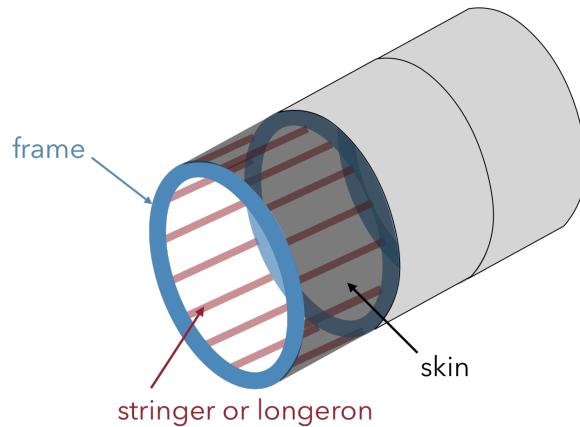


Figure 8.2: A depiction of some of the major elements in a typical fuselage structure.

Figure 8.3 shows a simplified wing structure. Like the frames on the fuselage, the ribs provide shape and buckling strength. Even though pictured as solid, ribs are generally not solid in order to save weight. Though not pictured, stringers (also called stiffeners) are usually used as well, and of course skin. One major difference for a wing is that it must carry the lift of the airplane, which equals or exceeds the weight of the entire airplane. To help carry those loads, a spar or spars are used. Spars help carry the vertical loads and also increase torsional rigidity. The top

and bottom of the spars, called the spar caps, carry bending loads.

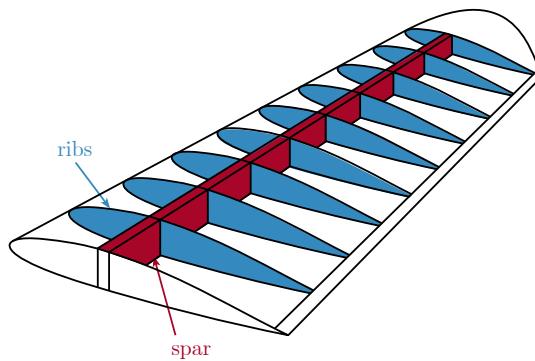


Figure 8.3: Some elements in a simple wing structure. Figure adapted from [this figure](#) and [this figure](#), both by MLWatts, Wikimedia Commons, public domain.

Large transports typically use a box beam structure with a front spar and rear spar (depicted as 1 and 2 in Fig. 8.4). The region between these spars forms a box beam structure. A cross-sectional view can be seen in Figure 5 of [this webpage](#).

8.2 Wind Bending

For many aircraft the critical loads in the wing come from bending. Figure 8.5 shows a typical load distribution across a conventional aircraft. Most of the weight (inertial load) is concentrated near the centerline, whereas the lift is distributed across the wing. Furthermore, the weight is distributed throughout the entire aircraft whereas the lift is concentrated primarily in the wing. In other words, the inertial loads from the wing are relatively small compared to the lift in the wing. The wing, mounted to the fuselage, acts like a cantilevered beam, and the distributed aerodynamic loads create large bending moments in the wing. Flying wings are an exception where bending may not be a critical case. For a flying wing the aerodynamic and inertial loads are all contained in the wing so they tend to cancel out to a much greater degree.

In the previous section, we discussed how transport wings typically use a

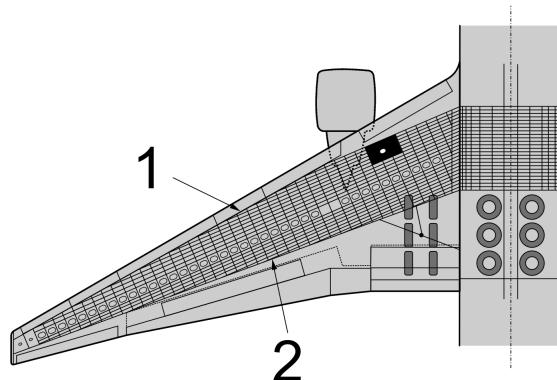


Figure 8.4: A more complex wing structure has a front and rear spar, ribs, and stringers, in a box beam construction. Figure from [Tosaka](#), Wikimedia Commons, CC BY 3.0.

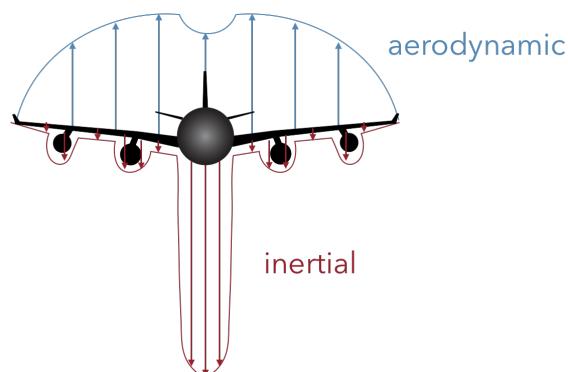


Figure 8.5: Typical aerodynamic and inertial load distribution for a transport aircraft.

box beam structure. A simplified representation of one is depicted in Fig. 8.6. The dashed portions of the wing contain just a thin skin, that is effectively non-structural, with the primary purpose of creating the aerodynamic shape. A cross-section of the wing, with positive loading, experiences a bending moment like that shown in the figure. The upper surface is in compression, while the lower surface is in tension.

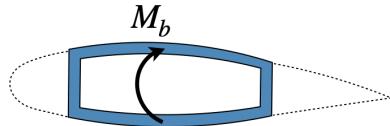


Figure 8.6: A notional depiction of a box beam often used in large aircraft wing structures. It consists of a forward and rear spar, and spar caps on top and bottom.

In the following, we derive a simplified estimation of the mass of material required to resist bending. This is not meant to be a highly accurate methodology, but rather a simple analytic expression that can give insight into what parameters most affect the structural weight of the wing. The wing can be approximated as a beam, and the bending stress can then be estimated from the beam bending equation (assuming uni-axial bending):

$$\sigma = \frac{M_b c}{I} \quad (8.1)$$

where M_b is the bending moment, c is the distance from the neutral axis, and I is the moment of inertia about the neutral axis. In this case, the distance from the neutral axis to the upper (or lower) surface where the stress is highest is approximately $t/2$ where t is the airfoil thickness. We can approximate the moment of inertia as two rectangles each with area $A/2$ as seen in Fig. 8.7. The sides of the box beam would contribute some to the moment of inertia, but the biggest contribution comes from the spar caps (similar approximation often used in I-beams). Furthermore, we assume that the spar cap is thin relative to the airfoil thickness, which is usually a good approximation, so that the moment of inertia about its own center line is negligible. Using the parallel axis theorem, and multiplying by two to account for both upper and lower surfaces, we estimate the box beam moment of inertia as: $I = 2(A/2)(t/2)^2$. Putting these terms together gives:

$$\sigma = \frac{2M_b}{At} \quad (8.2)$$

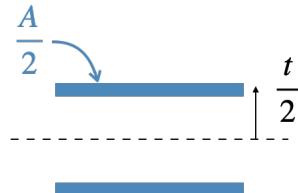


Figure 8.7: Simplified representation of the box beam for the purposes of estimating its moment of inertia about the neutral axis.

The mass of bending structural material in the wing is given by integrating the area A of material across the span:

$$m_{bending} = \int \rho Ady \quad (8.3)$$

where ρ is the density of the structural material. If we solve the stress equation for A and plug it into this expression we find that the mass of bending structural material can be estimated by the following expression:

$$m_{bending} = \int \frac{2\rho M_b}{\sigma t} dy \quad (8.4)$$

If we further assume that the structure is as efficient as possible, and thus every section is near its maximum stress (a constant), then the structural bending weight of the wing is proportional to the local bending moment over the local airfoil thickness.

$$m_{bending} \propto \int \frac{M_b}{t} dy \quad (8.5)$$

This expression explains why a thicker airfoil generally results in a lighter wing (if bending constrained). A thicker airfoil increases the moment of inertia of the section, like separating the two ends of an I-beam, reduces stress, and thus reduces the amount of material needed in the wing to resist bending.

In conceptual design it can be useful to constrain $\int M_b/t$ as a surrogate for structural wing weight. The bending moment varies throughout the wing, but the maximum is generally the *root bending moment*, which is the bending moment at the root of the wing. Sometimes root bending moment is used as a simpler surrogate for structural wing weight.

Although root bending moment has been frequently used in past aircraft conceptual design studies, it's often a poor metric for wing bending weight. One problem with root bending moment is that it ignores airfoil thickness so that the optimal spans are far too large (for a fixed area as span increases the chords shrink, and thus for fixed airfoils the thicknesses shrink also). Root bending moment is also especially problematic for nonplanar wings. A winglet canted towards the root will not increase root bending moment at all, but would of course increase wing weight.

Root bending moment is given by the integral of the lift distribution times the moment arm from the root across the semispan (see the top of Fig. 8.8).

$$M_{br} = \int_0^{b/2} y L' dy \quad (8.6)$$

As you may recall from statics, we can always replace a distributed load with an equivalent point load, at an appropriate location. We will call the nondimensional distance along the semispan η (see the bottom half of Fig. 8.8). Using the point load representation the root bending moment is given by:

$$M_{br} = \frac{\eta b L}{4} \quad (8.7)$$

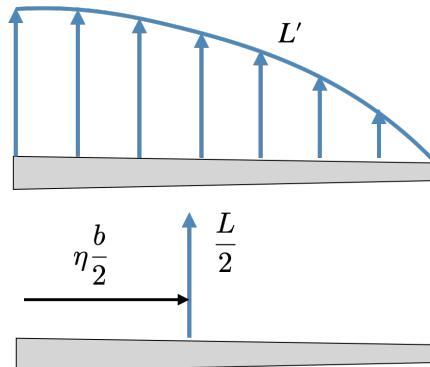


Figure 8.8: The distributed load on a wing can always be represented equivalently as a point load at some nondimensional distance η away from the root.

The position η defines the location of the lift centroid, or the center of lift. While we won't perform the integral here, we can show that if we use an

elliptic lift distribution for L' then the analytic solution for η is:

$$\eta_{elliptic} = \frac{4}{3\pi} \quad (8.8)$$

This means that the lift centroid is at about 42.4% of the semispan for an elliptic lift distribution. If the lift centroid is further out, then the bending moment increases.

For a given lift centroid position, there are an infinite number of possible lift distributions. Of those infinite options, one has the lowest induced drag. R. T. Jones derived what the best inviscid span efficiency is for a given lift centroid position [1]:

$$e_{inv,max} = \frac{1}{\frac{9}{2}\pi^2\eta^2 - 12\pi\eta + 9} \quad (8.9)$$

This function is plotted in Fig. 8.9. Again, it should be emphasized that this equation does not predict what the inviscid span efficiency is for a given lift centroid position, it predicts what the best possible inviscid span efficiency is for a given lift centroid position.

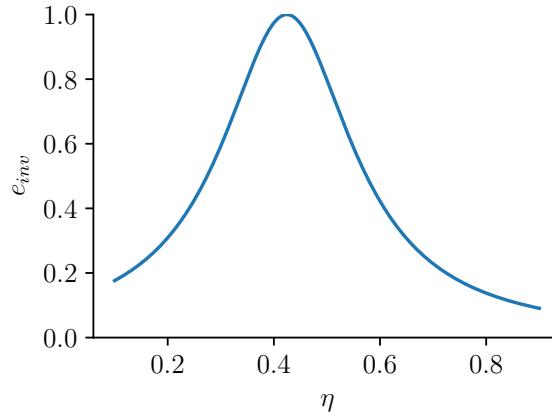


Figure 8.9: The highest possible inviscid span efficiency, for a given position of the lift centroid η .

We can think of this figure as directly representing a tradeoff between drag and weight. The peak of this curve is at $\eta = 4/(3\pi)$ and corresponds to $e_{inv} = 1$: an elliptic lift distribution. If we decrease η then we decrease the bending moments in the wing, and thus the bending weight of the wing.

However, we can see that we have to sacrifice some inviscid span efficiency to make this happen, or in other words the induced drag will increase. For a given wing weight, with fixed area (so that parasitic drag is roughly constant), the optimal aerostructural lift distribution differs from the aerodynamic optimum as seen in Fig. 8.10. As we learned in the Chapter 3 the minimum induced drag solution for a fixed span planar wing is an elliptic wing. However, if we consider structures we can maintain the same wing weight but lower drag by shifting the load distribution inboard while simultaneously increasing the span. The resulting drop in inviscid span efficiency is more than made up by the increase in span. The details on what the optimal span increase and lift distribution is vary considerably depending on the wing taper, critical structural altitude, stall behavior, etc [2]. However, the general principle is that the aircraft performance is usually improved by shifting the load inboard slightly to allow for a higher span (or a lighter wing for a fixed span).

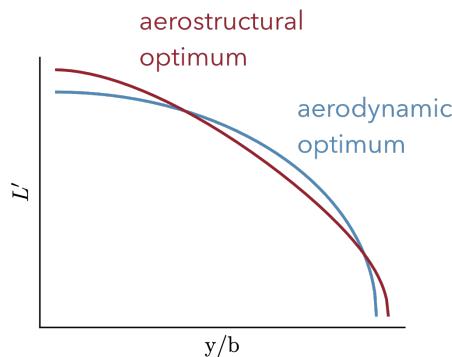


Figure 8.10: Comparison between aerodynamically optimum lift distribution and aerostructural optimum.

8.3 Flight Envelope Preliminaries

There are various ways to look at an aircraft's flight envelope. We will look at two. The first is referred to by various names including a placard speed diagram, a doghouse plot, or just a flight envelope. Sometimes these names refer to other types of diagrams. The plot we are interested in is max/min true airspeed as a function of altitude. The second type of plot is called a $V-n$ diagram. This plot shows the max/min load factors as a

function of speed. Specifically we will look at the maneuvering envelope. Before diving into the details we need to learn more about load factors and airspeeds. The n in $V \cdot n$ is called the *load factor*. It is a ratio of the lift to the weight of the aircraft:

$$n = \frac{L}{W} \quad (8.10)$$

In steady level flight, the aircraft flies with lift equal to weight, or a load factor of one ($n = 1$). Typically, load factors are expressed in “g’s”. For example, $n = 1$ is a load factor of 1g, whereas $n = 2$ is a load factor of 2 g’s. This phrasing is used because at $n = 2$, in the frame of the aircraft, objects would appear to fall with an acceleration twice that of gravitational acceleration (g). The same definition is used in turning flight (Section 7.6), e.g., a 3g turn.²

The V in $V \cdot n$ refers to air speed. There are actually several different air speeds that are used in connection with aircraft.

- *True airspeed (TAS)*: the speed of the aircraft relative to the air. This is what we usually think of as V_∞ .
- *Ground speed*: the speed of the aircraft relative to the ground. If there is no wind then ground speed is equal to true airspeed.
- *Equivalent airspeed (EAS)*: an equivalent airspeed at sea level that has the same dynamic pressure the aircraft is currently flying at. We can relate TAS and EAS with the following expression:

$$\frac{1}{2} \rho (TAS)^2 = \frac{1}{2} \rho_{sl} (EAS)^2 \quad (8.11)$$

where ρ_{sl} is the air density at sea level. Solving for *EAS* gives:

$$EAS = TAS \sqrt{\frac{\rho}{\rho_{sl}}} \quad (8.12)$$

The benefit of EAS is that for structural loading, we don’t care about speed per se, but rather about the dynamic pressure. If the flight handbook specified limits in terms of true airspeed then it would need different speed limits for every altitude (or really for every density). Alternatively, if we use equivalent airspeed to specify the limits than those speeds apply no matter what altitude we fly at.

²From Eq. (7.58) we can see that in turning flight $n = 1/\cos \phi$.

In an aircraft, airspeed is generally not measured directly anyway. Instead dynamic pressure is measured using a [pitot tube](#). From a dynamic pressure (q) measurement one can compute the equivalent airspeed:

$$EAS = \sqrt{\frac{2q}{\rho_{sl}}} \quad (8.13)$$

- Indicated airspeed: This speed (and the next one) are less important for our discussion, but are mentioned here to be complete. A pitot tube measures static pressure and total pressure and from those measurements one can compute dynamic pressure. In an introductory fluids class you learned that dynamic pressure is simply the difference between those two pressures:

$$q = p_T - p \quad (8.14)$$

However, that is not true at high speeds where compressibility is important. For compressible flow the dynamic pressure can be computed from those two pressure measurements, but the formula is more complex:

$$q = \frac{\gamma}{\gamma - 1} p \left[\left(\frac{p_T}{p} \right)^{(\gamma-1)/\gamma} - 1 \right] \quad (8.15)$$

An airspeed indicator reports the speed as estimated using the difference in total and static pressure (and thus ignores compressibility). The reason for this is that a pressure difference can be measured with a purely mechanical device. The resulting speed from an airspeed indicator, is called the indicated airspeed. With modern air data computers, an airspeed indicator is less relevant, though may be used as a backup system.

- Calibrated airspeed. Also less relevant to our discussion, but this is indicated airspeed with instrumentation and position corrections applied. For example, at higher angles of attack the pitot tube is not in ideal alignment and will have additional error.

To summarize the different flight speeds: indicated airspeed uses a simple difference in total pressure and static pressure. Indicated airspeed

becomes calibrated airspeed when corrections to instrumentation and position errors are applied. Calibrated airspeed becomes equivalent airspeed, when compressibility is accounted for. Equivalent airspeed becomes true airspeed when the local air density is used. True airspeed becomes ground speed when the local wind is included.

8.4 Speed-Altitude Flight Envelope

In this section and the next we will use the equations and limits set for transport category aircraft FAR Part 25.333, though the general principles are applicable to other categories. We will only highlight the primary considerations, the FARs provide many more requirements and details needed for certification.

One speed that is defined in the FARs is V_C or M_C , the design cruising speed or Mach number. It is intended to be the maximum cruising speed, and is typically about 6% higher than a normal cruise speed because the aircraft doesn't usually cruise at max continuous power [3]. Note that there are other requirements in the FARs related to adequate separation from gust speeds.

Another specified speed in the FARs is V_D or M_D , the design dive speed or Mach number. If no other information is supplied it must be 1.25 times V_C . However, the FARs allow for demonstration of a lower dive speed by flying at 7.5° below the nominal flight path for 20 seconds then pulling up at $1.5g$. The speed increase from this type of maneuver is usually lower than the maximum required, as is typically close to 1.15 times V_C [3].

The flight envelope diagram shows true air speed on the x -axis and altitude on the y axis (Fig. 8.11). The maximum speed is limited by V_D or M_D and structural limitations. For commercial transports, typically a design dive Mach number is specified because of compressibility drag limitations. In the isothermal portion of the atmosphere that translates to constant true air speed. However, at lower altitudes the temperature increases towards the ground. Thus, the speed of sound increases, and for a given Mach number, the true airspeed increases (blue curve in Fig. 8.11). As the altitude decreases, this true airspeed would continue to increase, and thus the dynamic pressure and resulting structural loads. However, there is no need to fly as fast at such low altitudes, and so the maximum dynamic pressure is constrained. This point is called the "knee" of the

curve.

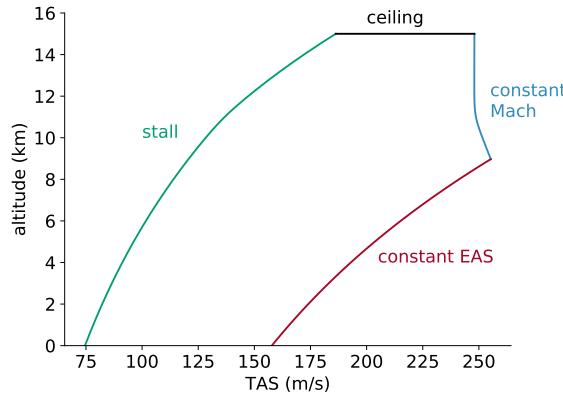


Figure 8.11: An example flight envelope showing the range of true air speeds allowed as a function of altitude.

From this point downward the speed is constrained at constant dynamic pressure, or in other words constant equivalent airspeed. We can use Eq. (8.11) to convert from EAS to TAS. As the density increases towards the ground the resulting TAS decreases so as to maintain constant dynamic pressure (red curve in Fig. 8.11).

The lower limit on speed is set by stall, and this we can compute simply from its definition (Eq. (7.61)). As altitude is increased the density decreases and so the stall speed increases. This is shown in the green curve. Generally some stall margin is added in, and the curve is adjusted at lower altitudes for flaps/slats. Finally the operational ceiling (Section 7.4) sets a limit on the maximum altitude that can be flown (the black curve). For aircraft that would fly at higher altitudes, like supersonic transports, sometimes additional “knees” are added to the curve otherwise the allowable speeds at sea level would be excessive.

8.5 V-n Diagrams

A notional $V\text{-}n$ diagram is shown in Fig. 8.12. We will break down how each part is created, but for now we start with an overview. On the x -axis is the flight speed in equivalent airspeed. Recall, that EAS is more useful than true airspeed for this type of diagram because it is a measure of

dynamic pressure independent of altitude (always based on sea level density). On the y -axis is the load factor. The region between the blue curves defines our flight envelope. We can safely maneuver at any speed/load factor within that region. The flight envelope is limited by stall, maneuver limits, and gust limits.

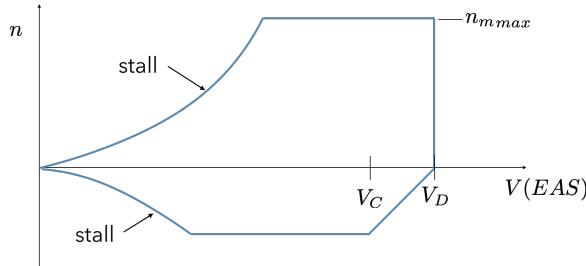


Figure 8.12: A notional V - n diagram.

Let's begin with stall. We are familiar with stall from previous chapters. However, the type of stall we computed before is called the $1g$ stall (level flight), whereas now we compute stall for various load factors. Recall that the load factor is L/W . The maximum load factor that can be achieved before stall is given by:

$$n_{max} = \frac{L_{max}}{W} = \frac{C_{Lmax} \frac{1}{2} \rho_{sl} (EAS)^2 S_{ref}}{W} \quad (8.16)$$

Clearly, we cannot fly at a load factor higher than stall will allow (for a given speed), even if structural damage was not an issue. The equation above defines a quadratic curve as a function of EAS , and is plotted in Fig. 8.13 for both positive and negative angles of attack. The relationship is actually more complex as C_{Lmax} is a function of Mach number and Reynolds number, but for this discussion we will treat it as a constant with flight speed. The other important point to remember is that C_{Lmax} is not the same for positive and negative angles of attack, hence the curves are not symmetric about the x -axis.

The next limiting factor is maneuvers. As we saw in Fig. 7.4, maneuvers like turning are going to increase our load factor. The FARs limit the maximum allowable maneuver load factor:

$$n_{m,max} = 2.1 + \frac{24,000}{W + 10,000} \quad (8.17)$$

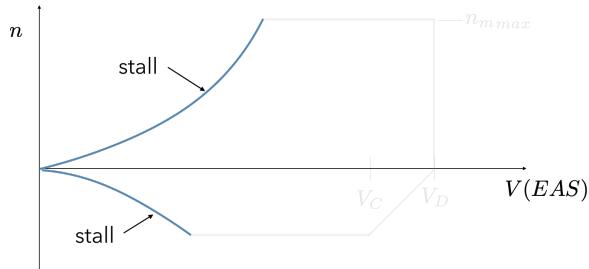


Figure 8.13: Stall limitations prevent higher load factors at low speeds.

where W is the maximum weight of the aircraft in lbs. The additional instruction in the FAR is: "except that n may not be less than 2.5 and need not be greater than 3.8". Most commercial aircraft are large enough such that their maximum maneuver limit load is 2.5. This limit load defines an upper limit that intersects with the stall curve (Fig. 8.14). A factor of safety of 1.5 is also applied in testing for ultimate loading, though not shown on the diagram.

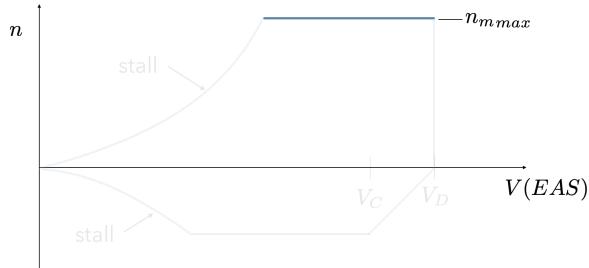


Figure 8.14: The regulations specify a maximum maneuver limit load.

Next, we cannot exceed the dive speed discussed in the previous section, for the given altitude. Because we are plotting this in EAS, this limit applies for the entire red part of the curve shown in Fig. 8.11, which is the most critical. The blue part the curve in Fig. 8.11, would have a lower EAS. We can see that V_D limits the diagram on the right side as shown in Fig. 8.12.

Finally, for negative maneuver limit loads the FARs state that the lower limit may not be less than -1 at speeds up to V_C and that it must vary linearly with speed from the value at V_C to zero at V_D (Fig. 8.15).

The other structural consideration is gusts. A vertical gust can cause a significant change in angle of attack and subsequently on load factors.

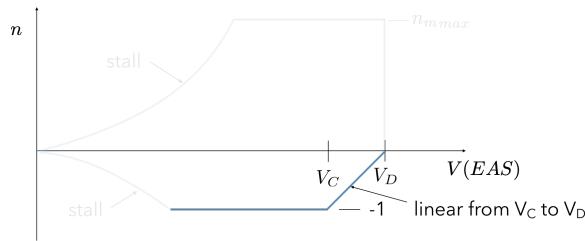


Figure 8.15: The negative maneuver limit load is capped at -1 , and has a linear ramp from V_C to V_D .

However, the determining of critical gust loads is more complex as it requires dynamic analysis and will not be covered here. More details are in the FARs and in corresponding [advisory circulars](#).

Once the complete V - n diagram is completed, we can identify the critical points as shown in Fig. 8.16. The aircraft design must be shown to be safe at these critical points (including the 1.5 factor of safety). The four corners tend to stress all parts of the spar cap in tension and compression (in different amounts) as shown in Fig. 8.17.

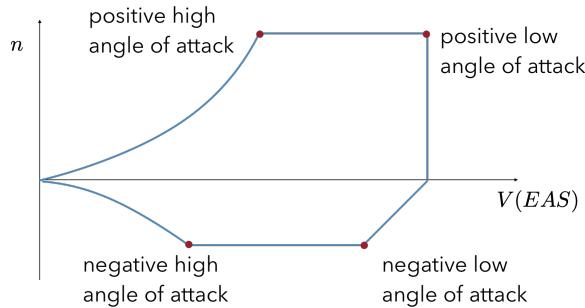


Figure 8.16: For structural testing/simulation we need to verify safety at the extreme points (including safety factors).

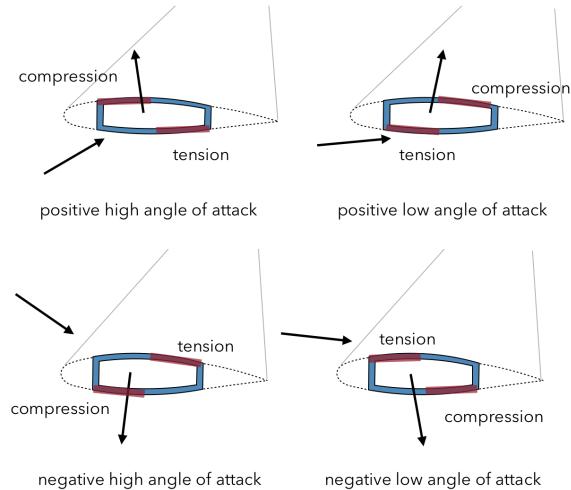


Figure 8.17: The extreme points of the V - n diagram stress different parts of the spar cap in tension and compression.

Bibliography

- [1] Jones, R. T., "The Spanwise Distribution of Lift for Minimum Induced Drag of Wings Having A Given Lift and A Given Bending Moment," Technical note 2249, National Advisory Committee For Aeronautics, Dec 1950.
- [2] Ning, A. and Kroo, I., "Multidisciplinary Considerations in the Design of Wings and Wing Tip Devices," *Journal of Aircraft*, Vol. 47, No. 2, Mar 2010, pp. 534–543.
- [3] Kroo, I., *Aircraft Design: Synthesis and Analysis*, Stanford University, 2006.

CHAPTER 9

Rocket Fundamentals

We now turn our attention to rockets. Rockets are certainly different from aircraft, but we will find that many of the principles we have learned will transfer over with minor adaptations. In this chapter we discuss some high-level principles and definitions related to rocket performance.

9.1 Types

Various types of rockets exist including chemical, nuclear, solar, electric, etc. Chemical rockets are the most common and are the only ones we will discuss in this book. The three primary types of chemical rockets include: *solid propellant*, *liquid propellant*, and *hybrid propellant*. The propellant of a rocket consists of *fuel* and an *oxidizer*.

A solid propellant rocket (Fig. 9.1), also called a solid rocket motor, uses a solid mixture of fuel and oxidizer called the *grain*. The main advantages of a solid propellant rocket are no moving parts enabling a simpler design, the use of propellants that are usually stable enabling ease of storage and handling, a compact design, and generally reduced costs. Their primary disadvantage is that the burning rate must be determined beforehand through its design (discussed below), and so it cannot be throttled. It also can't be shut off once ignited. Solid rockets also tend to have a lower specific impulse (a measure of efficiency discussed later).

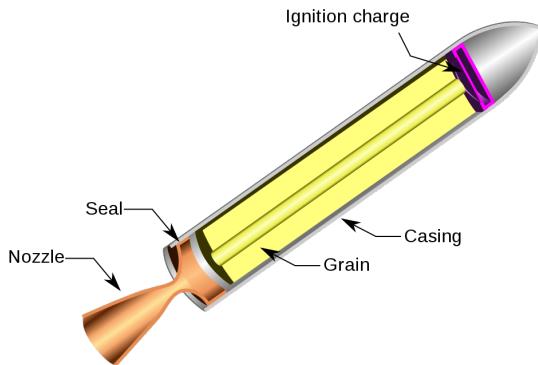


Figure 9.1: Depiction of a solid rocket. Figure by Pbroks13, Wikimedia Commons, CC BY-SA 3.0.

The grain can have a variety of shapes in order to create the desired thrust profile. A few simple shapes are shown in Fig. 9.2. The mass flow rate, and thus the thrust, is proportional to the burning surface area. So, for example, the tube shape at the top of the figure increases in surface area as it is burned making the thrust increase with time. This is called a *progressive burn*. The star shape maintains a relatively constant surface area, and thus constant thrust, in what is called a *neutral burn*. The last example has a double anchor shape with decreasing surface area, and thus decreasing thrust, in what is called a *regressive burn*.

These examples are simplistic—large solid rockets utilize much more complicated shapes. For example, the space shuttle solid rocket boosters used a combination of an 11-point star shape in the forward end and double truncated cone shapes in the aft end. The thrust profile for these boosters is more complex and is depicted in Fig. 9.3.

The highest stresses on the rocket occur when the dynamic pressure is highest, a point referred to as *max-q*. As the rocket increases in altitude the speed increases, but the density decreases creating a dynamic pressure curve that starts out low at the ground, peaks, then decreases again at high altitudes. For the space shuttle max-q occurs at about 10 km altitude.

Besides structural limitations, the drag is also highest at max-q and so it is more efficient to wait to increase thrust until after this point has past. We can now understand the reasoning behind the thrust profile in Fig. 9.3. It is desirable to have high thrust at ignition, a reduction in thrust during the high dynamic pressures, then an increase in thrust after max-q.

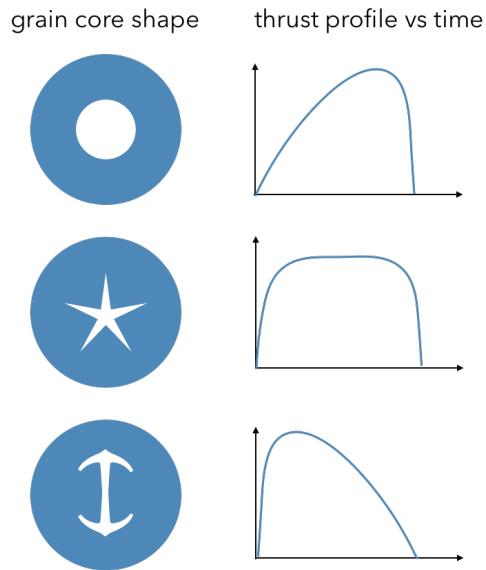


Figure 9.2: A few simple shapes for solid rocket grain cores and notional depictions of their corresponding thrust profiles over time.

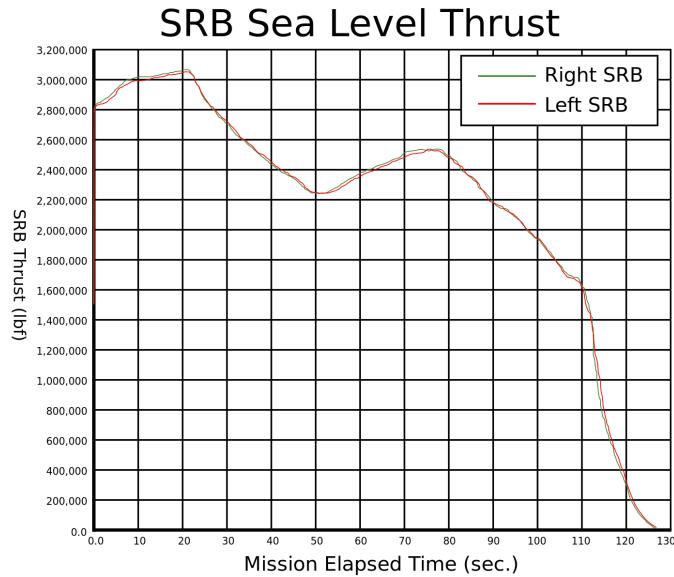


Figure 9.3: Thrust profile for the space shuttle solid rocket boosters. Figure from [NASA](#), Wikimedia Commons, public domain.

A liquid propellant rocket (Fig. 9.4), also called a liquid rocket engine, uses pressurized liquid propellants. The advantages/disadvantages of a liquid propellant rocket are basically just the opposites of a solid rocket. A liquid rocket can be throttled as the flow rates of the fuel and oxidizer can be adjusted or even stopped completely. However, the complexity (and associated cost) is much higher as they often require a turbopump, high pressures, and sometimes must be able to store cryogenic propellant. Liquid propellants tend to be more difficult to handle and store as they may be extremely toxic, highly reactive, and/or require cryogenic storage. Liquids are also subject to cavitation and **pogo**, which is an unstable feedback loop between thrust and flow rates.



Figure 9.4: Depiction of a liquid rocket. Figure by [Nwbeeson](#), Wikimedia Commons, CC BY-SA 3.0.

Hybrid rockets (Fig. 9.5) use a solid fuel with a liquid oxidizer (reverse hybrids are also possible, which are a liquid fuel with a solid oxidizer). The goal is to try to get the advantages of both solid rockets and liquid rockets. A hybrid can be throttled and shut down, and is somewhat simpler than a liquid rocket as less piping and valves are required. The main disadvantage is the mixing of a solid and liquid at burn time is less reliable and is thus generally less efficient. Perhaps the best known example of a hybrid rocket is that used on SpaceShipOne.

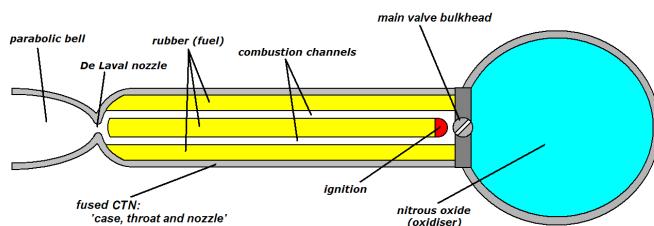


Figure 9.5: Depiction of the hybrid rocket motor used on SpaceShipOne. Figure by [Jack](#), Wikimedia Commons, public domain.

9.2 Thrust and Efficiency (Specific Impulse)

Most rockets produce thrust in essentially the same manner: a pressurized fuel/oxidizer mixture is combusted, and then thermodynamically expanded through a nozzle to produce high exit velocities (Fig. 9.6).

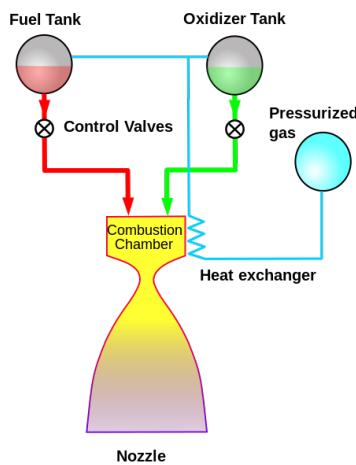


Figure 9.6: A notional depiction of a chemical rocket. Figure from [Duk](#), Wikimedia Commons, CC BY-SA 3.0

Previously, we introduced the thrust equation (Eq. (6.20)), which is repeated here:

$$T = \dot{m}(V_e - V_\infty) + \dot{m}_f V_e + (p_e - p_\infty)A_e \quad (9.1)$$

Most rockets are not designed for air to pass through as a working fluid so the first term is zero. For simplicity the last two terms are often combined leading to the definition of an *effective exit velocity* C .

$$T = \dot{m}_f C \quad (9.2)$$

where

$$C = V_e + \frac{(p_e - p_\infty)A_e}{\dot{m}_f} \quad (9.3)$$

Typically, the $\dot{m}_f V_e$ term in the thrust equation is much larger than the pressure term so that the effective exit velocity is not far off from the actual exit velocity.

Notice that because the first term is the thrust equation is zero, that the thrust of a rocket is independent of flight speed. It is not, however, independent of altitude. As we saw in Chapter 1, as altitude is increased the atmospheric pressure decreases. This means that thrust increases with altitude, with a typical increase of about 10–30% [1].

Recall also the definition of specific fuel consumption:

$$sfc = \frac{\dot{m}_f g}{T} \quad (9.4)$$

Rockets instead use the inverse of this expression and call it the *specific impulse* (actually this is simplified assuming constant thrust and constant mass flow rate as we will see shortly):

$$I_{sp} = \frac{T}{\dot{m}_f g_{sl}} \quad (9.5)$$

We can derive this by recalling the definition of an impulse (using thrust as the force below):

$$I = \int T dt \quad (9.6)$$

The specific impulse then is impulse per unit weight of propellant.

$$I_{sp} = \frac{\int T dt}{g_{sl} \int \dot{m}_f dt} \quad (9.7)$$

We can see that for constant thrust and constant mass flow rate, the above expression reduces to Eq. (9.8). We will use the simplified expression in this book.

Although the acceleration of gravity varies, the definition of I_{sp} uses a constant reference value for g , the value at sea level, so that the specific impulse can be specified independent of altitude. The units of specific impulse are typically given in seconds, although it is not a measure of time (it is thrust per weight flow rate). A higher value of I_{sp} is better. If we use the definition for thrust above (Eq. (9.2)) then we can relate specific impulse and effective exhaust velocity:

$$I_{sp} = \frac{C}{g_{sl}} \quad (9.8)$$

Thus, we can see that either I_{sp} or C can be used as a measure of the rocket's propulsive efficiency.

Usually specific impulse and thrust are inversely correlated for similar reasons to the efficiency/thrust tradeoffs we studied in Chapter 6. A launch vehicle with a chemical engine can produce a few million Newtons of thrust and will have a specific impulse of a few hundred seconds. Whereas a deep space probe with an [ion thruster](#) will produce only a tenth of a Newton of thrust, but with a specific impulse in the thousands of seconds.

9.3 Tsiolkovsky Rocket Equation

Consider the representation of a rocket shown in Fig. 9.7. Fuel is expelled out the back of the rocket in order to accelerate the rocket. We can derive an equation for the rocket acceleration using Newton's second law:

$$F = \frac{dp}{dt} \quad (9.9)$$

We will examine the change in momentum of the system, as well as the external forces.

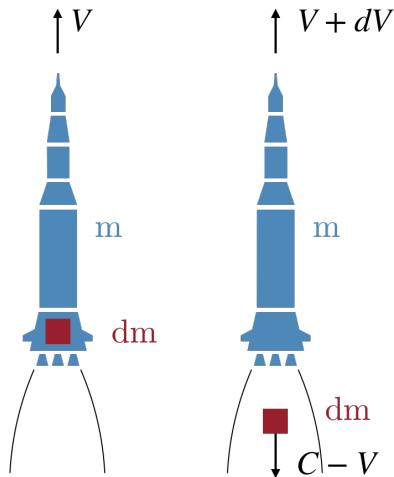


Figure 9.7: Depiction of a rocket where fuel is expelled to accelerate the rocket.

The initial momentum of the system is:

$$p_i = (m + dm)V \quad (9.10)$$

The final momentum of the system is:

$$p_f = m(V + dV) - dm(C - V) \quad (9.11)$$

Thus, the change in momentum per change in time is:

$$\frac{dp}{dt} = \frac{p_f - p_i}{dt} = m \frac{dV}{dt} - C \frac{dm}{dt} \quad (9.12)$$

The external forces acting on the rocket are primarily gravity and aerodynamic drag. Putting these together using Newton's law gives:

$$m \frac{dV}{dt} - C \frac{dm}{dt} = -mg - D \quad (9.13)$$

For simplicity, in the following derivation we will neglect drag. Drag varies through the rocket flight and so its impact would generally need to be included numerically. This is usually a smaller term, so for this first-order analysis we will neglect it:

$$m \frac{dV}{dt} - C \frac{dm}{dt} = -mg \quad (9.14)$$

$$dV - C \frac{dm}{m} = -g dt \quad (9.15)$$

$$\int dV - \int C \frac{dm}{m} = - \int g dt \quad (9.16)$$

We now integrate during the the burn time of the rocket, approximating C and g as constant during the time interval (neither is really constant so using an average value would be more appropriate):

$$\Delta V = C \ln \left(\frac{m_{initial}}{m_{final}} \right) - gt_b \quad (9.17)$$

where t_b is the burn time. We can also express this equation in terms of a fuel fraction to more closely parallel how we expressed structural efficiency of an aircraft:

$$\Delta V = C \ln \left(\frac{1}{1 - m_p/m_{initial}} \right) - gt_b \quad (9.18)$$

where m_p is the propellant mass. In rockets a common nomenclature is to use the *mass ratio* (*MR*), which is defined as the initial mass divided by the

final mass¹. Using the mass ratio, the rocket equation becomes:

$$\Delta V = C \ln(MR) - gt_b \quad (9.19)$$

If the rocket is flying at an angle θ with respect to vertical, then only a component of gravity acts to slow down the rocket, and so the equation is modified as:

$$\boxed{\Delta V = C \ln(MR) - gt_b \cos \theta} \quad (9.20)$$

These are all various forms of what is known as Tsiolkovsky's rocket equation.

This equation is kind of like the Breguet range equation for aircraft performance as it combines various subsystem metrics into a high level metric. The change in velocity ΔV is an important parameter for rockets and is a function of propulsive efficiency (C), structural efficiency ($m_p/m_{initial}$) and aerodynamic efficiency (drag is important but we neglected it in this simplified formula because it requires knowing how drag changes in time).

This equation gives insight into the magnitude of the challenge of designing a rocket that can escape earth's gravity. In order to get into low earth orbit, accounting for gravity and drag, a ΔV of about 9 km/s is required. Some of the most energetic rocket propellants have an effective exhaust velocity (C) of 3–4 km/s. Using the rocket equation and solving for $m_p/m_{initial}$ reveals that in order to get into the low earth orbit the rocket needs to be about 90% propellant! Orbital launch systems are typically closer to 85% propellant as they make use of staging (discussed in this next section), though this is still an amazingly high number.

Compare that to a commercial aircraft that has approximately a 45% fuel fraction, or a car that has a fuel fraction of a few percent. This means the structure of the rocket must be extremely efficient, and after the propellant and structure there isn't much room left for payload. Indeed, the *payload fraction* (mass of payload / initial mass) of the Saturn V was only 4% and that of the space shuttle was about 1.5%.

In Section A.3 this equation is integrated once more to estimate the change in altitude:

$$\Delta h = \left[V_0 + C \left(1 - \frac{\ln(MR)}{MR - 1} \right) - \frac{1}{2} g t_b \cos \theta \right] t_b \cos \theta \quad (9.21)$$

¹In some texts, the mass ratio is defined as the inverse of this definition, so be careful!

9.4 Staging

The purpose of most rockets is to deliver a payload. However, to get that payload to its destination a lot of propellant and structure is required. The primary mass components of a rocket are the *structural mass* (m_s), propellant mass (m_p), and the payload mass ($m_{payload}$).

As propellant is burned some of the structure is no longer necessary and it is desirable to shed this structure to lighten the rocket. This is accomplished with *staging*. A multi-stage rocket is depicted in Fig. 9.8.

Staging increases the altitudes that can be reached, in fact a single stage rocket cannot reach orbit with current technology. Each stage might use different types of motors optimized for the conditions it will be used in.

The downside of staging is added complexity, and some inefficiencies associated with lifting extra motor weight for stages that will be used later.

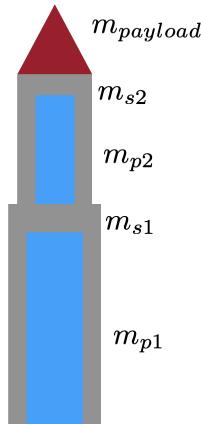


Figure 9.8: Depiction of a two stage rocket broken down into propellant mass (m_p), structural mass (m_s), and payload mass ($m_{payload}$)

Using the example two-stage rocket in Fig. 9.8, the initial mass is given by:

$$m_{initial} = m_{p1} + m_{s1} + m_{p2} + m_{s2} + m_{payload} \quad (9.22)$$

The mass ratio for stage one (initial over final mass) is:

$$MR_1 = \frac{m_{initial}}{m_{initial} - m_{p1}} \quad (9.23)$$

Thus, the velocity change from stage one (ignoring losses from gravity for

this example) is found by using Tsiolkovsky rocket equation:

$$\Delta V_1 = g_{sl} I_{sp} \ln(MR_1) \quad (9.24)$$

For the second stage the mass ratio is (recalling the the structure from stage one will be discarded):

$$MR_2 = \frac{m_{initial} - m_{p1} - m_{s1}}{m_{initial} - m_{p1} - m_{s1} - m_{p2}} \quad (9.25)$$

The corresponding velocity change from stage two is:

$$\Delta V_2 = g_{sl} I_{sp} \ln(MR_2) \quad (9.26)$$

The total increase in velocity across both stages is simply their sum:

$$\Delta V = \Delta V_1 + \Delta V_2 \quad (9.27)$$

We can perform a similar analysis to that just described, but repeated for a range of stages. It turns out that the optimal scenario for a staged rocket, is to design it so that each stage produces the same change in velocity (i.e., $\Delta V_1 = \Delta V_2 = \dots$). If we assume optimal staging, and that each stage has the same propellant fraction, we can compute an analytic expression for the payload fraction that can be delivered [2].

$$\text{payload fraction} = \left[1 - \frac{1}{\zeta} + \frac{1}{\zeta} \exp^{\frac{-\Delta V}{n I_{sp} g_{sl}}} \right]^n \quad (9.28)$$

where n is the number of stages, and ζ is the propellant fraction per stage (identical for all stages).

As an example, let's consider a rocket where each stage is 90% propellant (optimistic), that the specific impulse of each stage is 350 seconds (towards the high end for a chemical rocket), and that the required ΔV is 9 km/s (low earth orbit). The results are shown in Fig. 9.9. Notice that one stage cannot get to orbit, that the additional benefit decreases rapidly with each additional stage, and that even in the best scenario no more than about 5% of the rocket can be payload. The small additional benefits relative to the increased complexity of adding stages, explains why the use of more than three stages in a rocket is rare.

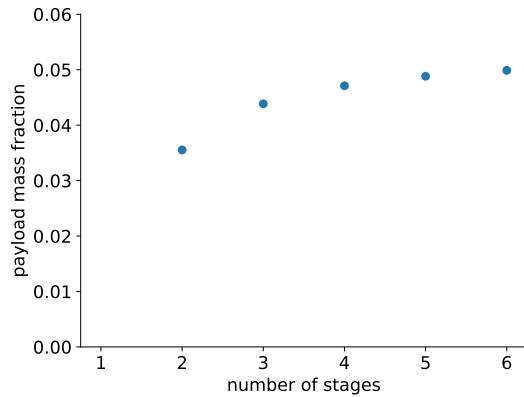


Figure 9.9: Maximum payload that can be carried to orbit as a function of the number of stages assuming optimal staging, 90% propellant per stage, and an I_{sp} of 350 seconds.

9.5 Stability

Roll, pitch, and yaw are defined the same as for an aircraft. In fact, not much new needs to be said about rocket stability as the principle is the same as for an aircraft. Namely that the center of gravity must be forward of the aerodynamic center (and for a rocket, forward means above).

However, for a rocket the angle of attack is essentially always zero.

Because the angle of attack doesn't really change, there is no real need to distinguish between the aerodynamic center and the center of pressure like there is for an airplane. Thus, when discussing rockets the stability criteria is typically expressed as requiring the center of gravity to be above the center of pressure (Fig. 9.10).

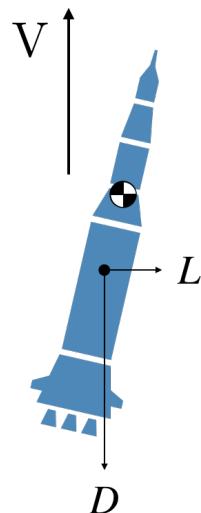


Figure 9.10: A statically stable rocket has its center of gravity above the center of pressure.

Bibliography

- [1] Sutton, G. P. and Biblarz, O., *Rocket Propulsion Elements*, John Wiley & Sons, Nov 2016.
- [2] Cantwell, B. J., "AA 283: Aircraft and Rocket Propulsion," Course Reader, Stanford University, Jan 2007.

CHAPTER 10

Rocket Propulsion

It *is* rocket science! Rocket science isn't a separate field per se, but is primarily a combination of thermodynamics, compressible flow, and chemistry. We will start with a review of some thermodynamic principles. A comprehensive discussion of compressible flow is beyond our scope, but we will learn some key principles and use some of the results. For those interested, and that have a background in compressible flow, derivations of the key equations are provided in appendices. Chemistry is beyond our scope, and fortunately is not necessary for initial performance analysis.

10.1 Thermodynamics Review

We begin by reviewing some terms and concepts from thermodynamics. These include equations of state, the first and second laws of thermodynamics, specific heats, and isentropic relations.

10.1.1 Properties and Equations of State

We need to add a few additional fluid properties to those we have already been using (p : pressure, ρ : density, T : temperature).

v : is the specific volume (volume per unit mass). It is not a new quantity, just an alternative to density: $v = 1/\rho$.

e : is the specific internal energy (internal energy per unit mass)

h : is specific enthalpy (enthalpy per unit mass) and is given by:

$$h = e + pv \quad (10.1)$$

s : is specific entropy (entropy per unit mass)

The above quantities (including pressure, density, and temperature) are a function of state. This means that these properties are related and if any two of them are specified, all the others can be determined. For example:

$$e = f(p, T) \quad (10.2)$$

or

$$h = f(e, v) \quad (10.3)$$

and so on. These relationships are called *equations of state*.

I suspect this is part of the reason why compressible flow (rocket science) is perceived as highly complicated. A given formula can be expressed in a large variety of ways depending on the needs of the current analysis. There are publications dedicated to tabulating some of these variations (for example, this [NASA publication](#)). Don't be scared off by the apparent complexity. There is no need to learn all the variations, rather if one understands the core relationships, the important equations can be derived and rearranged to the form needed.

The most well-known equation of state is the ideal gas law, which assumes that intermolecular forces are negligible:

$$p = \rho RT \quad (10.4)$$

where R is the specific gas constant. The specific gas constant is related to the universal gas constant R_u by:

$$R = \frac{R_u}{M_w} \quad (10.5)$$

where M_w is the molecular weight of the particular gas (or gas mixture). For aircraft we were only concerned with air so we just used the specific gas constant, but for rockets the molecular weight will vary with the

propellants so we need to start from the universal gas constant. The value for the universal gas constant is:

$$R_u = 8.314 \frac{\text{J}}{\text{K mol}} \quad (10.6)$$

Its value was empirically determined, but it can also be derived from [statistical mechanics](#).

10.1.2 Laws of Thermodynamics

The first law of thermodynamics is:

$$de = \delta q + \delta w \quad (10.7)$$

In words it says that for a closed system the change in internal energy is equal to the sum of the heat added (δq) and the work done on the system (δw). Heat and work use a inexact differential (δ rather than d) because they depend on the path and not just the state. For a fluid, the work done on a control volume is from pressure:

$$\delta w = -pdv \quad (10.8)$$

Substituting into the first law gives:

$$de = \delta q - pdv \quad (10.9)$$

The second law of thermodynamics states that for a closed system the entropy does not decrease in time:

$$ds \geq \frac{\delta q}{T} \quad (10.10)$$

where the equality holds only for an ideal reversible process. If we substitute this expression into the first law of thermodynamics (Eq. (10.9)) we have Gibbs equation:

$$de + pdv - Tds \leq 0 \quad (10.11)$$

10.1.3 Specific Heats

Specific heat is the amount of heat needed to raise the temperature of a unit mass of a fluid by a unit degree:

$$C = \frac{\delta q}{dT} \quad (10.12)$$

There are two specific heats. The specific heat at constant volume is

$$C_v = \left(\frac{\delta q}{dT} \right)_v \quad (10.13)$$

and the specific heat at constant pressure is:

$$C_p = \left(\frac{\delta q}{dT} \right)_p \quad (10.14)$$

We can simplify these into more useful expressions using state variables. Recall the first law of thermodynamics, using the expression for work on a fluid:

$$de = \delta q - pdv \quad (10.15)$$

At a constant volume we can see that the last term vanishes and thus:

$$\left(\frac{\delta q}{dT} \right)_v = \left(\frac{\partial e}{\partial T} \right)_v \quad (10.16)$$

Substituting this into the definition for C_v gives:

$$C_v = \left(\frac{\partial e}{\partial T} \right)_v \quad (10.17)$$

Similarly, we can use the definition of enthalpy:

$$h = e + pv \quad (10.18)$$

differentiate:

$$dh = de + pdv + vdp \quad (10.19)$$

then substitute into the first law of thermodynamics (Eq. (10.9)):

$$dh - vdp = \delta q \quad (10.20)$$

When we take partial derivatives at constant pressure the second term vanishes so:

$$\left(\frac{\delta q}{\delta T}\right)_p = \left(\frac{\partial h}{\partial T}\right)_p \quad (10.21)$$

Thus, the specific heat at constant pressure is given by:

$$C_p = \left(\frac{\partial h}{\partial T}\right)_p \quad (10.22)$$

The ratio of specific heats is:

$$\gamma = \frac{C_p}{C_v} \quad (10.23)$$

We've already been using this quantity though its value is no longer 1.4 for most propellants.

Further simplification can be made for *thermally perfect* gases. A thermally perfect gas has specific heats that are functions of temperature only. One example of a thermally perfect gas is an ideal gas. It can be shown that for an ideal gas, the internal energy is a function of temperature only.¹

$$e = e(T) \quad (10.24)$$

By extension, the enthalpy is only a function of temperature:

$$h = e + pv \quad (10.25)$$

$$= e + RT \quad (10.26)$$

$$= e(T) + RT \quad (10.27)$$

$$\Rightarrow h = h(T) \quad (10.28)$$

If the internal energy and enthalpy are only functions of temperature then the partial derivatives in the specific heat formulas become ordinary derivatives.

$$C_v(T) = \frac{de}{dT} \quad (10.29)$$

$$C_p(T) = \frac{dh}{dT} \quad (10.30)$$

¹This is also true for a mixture of ideal gases, as long as they are not chemically reacting. Actually it is true for any equation of the state of the form $v = f(p/T)$, with an ideal gas being the important practical case.

One further level of simplification is a *calorically perfect* gas. A calorically perfect gas has specific heats that are constant (i.e., independent of temperature). In that case, the above expression can easily be integrated yielding:

$$e = C_v T \quad (10.31)$$

$$h = C_p T \quad (10.32)$$

For reference, air behaves like a calorically perfect gas until about 1000 K, and after about 2500 K it no longer behaves like a thermally perfect gas and starts to chemically react. For detailed rocket design a chemically reacting flow model should be used, but for conceptual design we can develop reasonable initial models using calorically perfect gas assumptions.

From the definition of enthalpy, the ideal gas law, and the definition of γ , we can show that for a calorically perfect gas:

$$h = e + pv \quad (10.33)$$

$$C_p T = C_v T + RT \quad (10.34)$$

$$C_p - C_v = R \quad (10.35)$$

$$1 - \frac{1}{\gamma} = \frac{R}{C_p} \quad (10.36)$$

$$C_p = \frac{\gamma R}{\gamma - 1} \quad (10.37)$$

$$C_v = \frac{R}{\gamma - 1} \quad (10.38)$$

10.1.4 Isentropic Relationships

An *isentropic* flow is one that is both *adiabatic* and reversible. Adiabatic means there is no heat or mass transfer (so only work can change the internal energy). Isentropic means constant entropy. It is an idealized state, though often a good approximation for the flow in a rocket nozzle after combustion.

We start with Gibbs equation (Eq. (10.11)) and assume isentropic flow ($ds = 0$). Note also that the equality holds because isentropic flow is reversible:

$$\frac{de}{T} + \frac{Pdv}{T} = 0 \quad (10.39)$$

Using the definition of C_v for a calorically perfect gas (Eq. (10.31)) and the ideal gas law:

$$C_v \frac{dT}{T} + R \frac{dv}{v} = 0 \quad (10.40)$$

Integrating both sides from state 1 to state 2 gives:

$$C_v \ln \frac{T_2}{T_1} + R \ln \frac{v_2}{v_1} = 0 \quad (10.41)$$

If we express C_v in terms of R (Eq. (10.38)) we have:

$$\frac{1}{\gamma - 1} \ln \frac{T_2}{T_1} + \ln \frac{v_2}{v_1} = 0 \quad (10.42)$$

$$\ln \left(\frac{T_2}{T_1} \right)^{1/(\gamma-1)} + \ln \frac{v_2}{v_1} = 0 \quad (10.43)$$

$$\ln \left(\frac{T_2}{T_1} \right)^{1/(\gamma-1)} = -\ln \frac{v_2}{v_1} \quad (10.44)$$

$$\ln \left(\frac{T_2}{T_1} \right)^{1/(\gamma-1)} = \ln \frac{\rho_1}{\rho_2} \quad (10.45)$$

$$\left(\frac{T_2}{T_1} \right)^{1/(\gamma-1)} = \frac{\rho_1}{\rho_2} \quad (10.46)$$

$$(10.47)$$

We can derive a similar expression for pressure if we replace the internal energy with enthalpy in Gibbs equation. The end result is the isentropic relationships:

$$\frac{p_2}{p_1} = \left(\frac{\rho_2}{\rho_1} \right)^\gamma = \left(\frac{T_2}{T_1} \right)^{\gamma/(\gamma-1)}$$

$$(10.48)$$

10.2 Compressible Flow Primer

Compressible flow is a deep subject spanning a course or two of its own. While we can't hope to cover much, we will introduce some fundamental principles required to understand rocket behavior.

10.2.1 Energy Equation

The energy equation for a fluid flow in differential form is given by:

$$\rho \frac{\partial h_T}{\partial t} + \rho \vec{V} \cdot \nabla h_T = \frac{\partial p}{\partial t} + \rho \dot{q} \quad (10.49)$$

where h_T is called the total enthalpy (kind of like total pressure):

$$h_T = h + \frac{V^2}{2} \quad (10.50)$$

If the flow is steady then the time derivative terms are zero, and if adiabatic then $\dot{q} = 0$. This leaves:

$$\vec{V} \cdot \nabla h_T = 0 \quad (10.51)$$

which means that the total enthalpy is constant along streamlines. For a rocket the flow through the nozzle all comes from one reservoir so all the streamlines have the same total enthalpy, and thus total enthalpy is constant everywhere in the flow:

$$h_1 + \frac{V_1^2}{2} = h_2 + \frac{V_2^2} \quad (10.52)$$

If heat is added between point 1 and 2 the equation can be modified as:

$$h_1 + \frac{V_1^2}{2} + q = h_2 + \frac{V_2^2} \quad (10.53)$$

10.2.2 Area Velocity Relationship

In all of our nozzle analyses we will make use of the quasi-1D assumption illustrated in Fig. 10.1. The x direction follows the centerline of the nozzle. Each cross section is circular, and, in general, fluid properties (such as density) vary throughout the cross section as a function of y and z . However, if we use area-averaged properties:

$$\bar{\rho}(x) = \frac{1}{A(x)} \int \rho(x, y, z) dy dz \quad (10.54)$$

then we can treat the flow as one dimensional (only varying in the x direction). We will drop the overbars for convenience with the understanding that all fluid properties are area averaged.

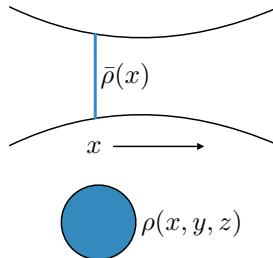


Figure 10.1: Fluid properties vary across each cross section, but if we use area averaged properties we can treat the flow as one-dimensional. This is called a quasi-1D assumption.

For a 1-dimensional flow, in addition to an energy balance that was discussed in the previous section, we must also apply a mass and momentum balance. A mass balance yields:

$$\rho V A = \text{const} \quad (10.55)$$

or equivalently

$$\frac{d(\rho V A)}{dx} = 0 \quad (10.56)$$

An inviscid momentum balance yields (inviscid is a good approximation outside of a very thin boundary layer, which could be thought of as a slight reduction in the outer wall dimensions):

$$\rho V \frac{dV}{dx} + \frac{dp}{dx} = 0 \quad (10.57)$$

Combining the mass and momentum balance, and using the definition for the speed of sound give the following equation:

$$(M^2 - 1) \frac{1}{V} \frac{dV}{dx} = \frac{1}{A} \frac{dA}{dx} \quad (10.58)$$

This is an illuminating equation that will give us some insight into the behavior of compressible flows. As an example, let's consider a diverging nozzle (Fig. 10.2). For a diverging nozzle dA/dx is positive, and A is always positive, so the entire right hand side of the equation is positive. The velocity V is also positive. There are three possibilities for the Mach number:

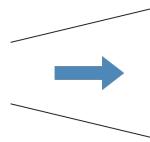


Figure 10.2: A diverging nozzle

- Subsonic ($M < 1$): In this case the term in parenthesis is negative and so dV/dx must also be negative. This is the behavior that is familiar from an introductory fluids class: as the nozzle diverges the fluid must slow down.
- Supersonic ($M > 1$): Now the term in parenthesis is positive and so dV/dx must also be positive. This gives the perhaps surprising result that a supersonic flow must accelerate through a diverging nozzle.
- Sonic ($M = 1$): Now the term in parenthesis is zero. This means that $dA/dx = 0$, which implies that the area is a maximum or a minimum in the nozzle. While either is possible mathematically, physically only the minimum area solution is possible. The implications of this will be discussed in more detail shortly.

The opposite behavior occurs for a converging nozzle ($dA/dx < 0$). For subsonic flow the behavior is the familiar one: acceleration in a converging nozzle, and deceleration in a diverging nozzle. However, for supersonic flow the behavior is opposite: deceleration in a converging nozzle, and acceleration in a diverging nozzle. While that may seem strange, it is just a consequence of a mass and momentum balance. It is fortunate as this behavior is what makes a nozzle work for propulsion.

Consider the converging-diverging nozzle shown in Fig. 10.3. An initially subsonic flow will accelerate through the converging portion. If the speed is high enough then the flow will become sonic at the throat and in the diverging portion the flow will be supersonic and thus will continue to accelerate through the diverging portion. Thus, the fluid can be continuously accelerated through the nozzle and this is why rocket nozzles are shaped the way they are.

Recall that we found that $M = 1$ can only occur at $dA/dx = 0$, or in other words where the area is a minimum. Such a point is called the *throat* of a nozzle. It is important to note that the presence of a throat does not mean the flow is necessarily sonic. It means that if there is a sonic point, it must

occur at the throat. A throat with a Mach number of 1 is said to be *choked*. The reason is that, as discussed in Section 3.5, information in a supersonic flow cannot travel upstream. So regardless of how the pressure decreases downstream, the flow upstream of the throat cannot change and is said to be *frozen*. The mass flow rate through the nozzle cannot increase any further.

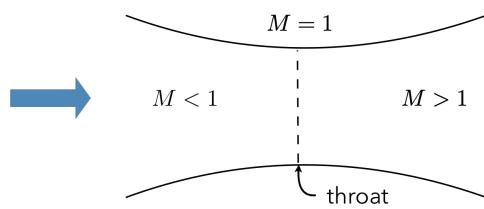


Figure 10.3: A converging-diverging nozzle with sufficient pressure ratio to achieve supersonic flow.

Finally, we need to discuss the exit conditions of a nozzle. For subsonic flow the exit pressure must equal the ambient pressure. For supersonic flow this is no longer true. Again, in a supersonic flow information cannot travel upstream, so the flow inside a nozzle with supersonic exit speeds cannot know about the ambient pressure. In other words the exit pressure is determined by the nozzle conditions and not the exterior atmosphere. For a supersonic nozzle (i.e., most rockets) there are three exit possibilities (Fig. 10.4). The exit pressure can be less than, equal to, or greater than the atmospheric pressure. If the exit pressure is greater than atmospheric then the nozzle is said to be *underexpanded*. The plume will expand quickly once it exits the nozzle. Once in space, or high altitudes, where the atmospheric pressure is essentially zero, the rocket will always be underexpanded. This behavior can be seen in the exhaust plume of the Saturn V rocket shown in Fig. 10.5.

If the exit pressure is equal to atmospheric pressure, then the rocket is *ideally expanded*. This is the most efficient mode of operation for the rocket. As this condition will generally only occur for one altitude, the rocket is usually designed so that this occurs just after max-q where high thrust can be used most efficiently. The last case occurs when the exit pressure is less than atmospheric, and is said to be *overexpanded*. In this case the plume will contract to compress the gas. This is typical behavior at low altitudes where the atmospheric pressure is high (see [here](#) for example).

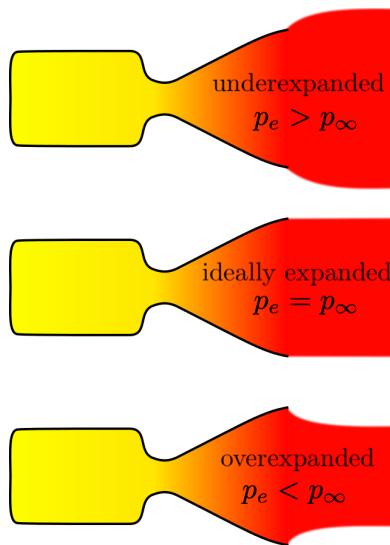


Figure 10.4: Three types of exit conditions for a nozzle with supersonic Mach numbers at the exit. Figure modified from one by [Hohum](#), Wikimedia Commons, public domain.

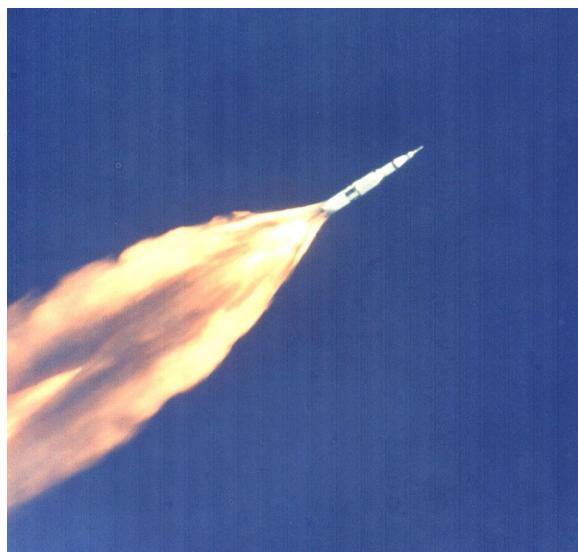


Figure 10.5: Exhaust plume of Saturn V rocket in flight. Picture from [NASA](#), Wikimedia Commons, public domain.

10.3 Nozzle Sizing

In the following discussion we will use the nozzle shown in Fig. 10.6. The subscript c refers to after combustion, t to the throat, and e to the exit.

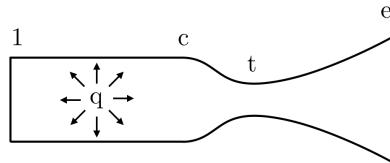


Figure 10.6: A notional nozzle depicted station 1 upstream of combustion, station c after combustion, station t at the throat, and station e at the exit.

The analyses of this chapter use what is called an *ideal rocket*. Some of the assumptions of an ideal rocket include: adiabatic outside of the combustion chamber, isentropic outside of the combustion chamber, calorically perfect ideal gases, chemical equilibrium, and no shock waves. These assumptions are more reasonable than they may seem at first, are accepted practice for initial design, and predict actual measured performance within about 1–6% [1]. For detailed design, correction factors or thermochemical analyses are used.

The analyses of this section require a background in compressible flow. As that background is not expected at this stage, all of the derivations are placed in Section A.4 and only the results are discussed in this chapter. In the following equations we will also make the assumption that the Mach number in the combustion chamber is negligibly small so that total quantities are equal to static quantities. This error is small (quantified in the appendix), and is a typical assumption for initial rocket design. The assumption is not necessary if the concept of total pressure and total temperature for a compressible flow is understood, but the level of error doesn't justify further diversion.

10.3.1 Exit Velocity

As derived in Section A.4.1, the exit velocity of a rocket can be estimated as:

$$V_e = \sqrt{\frac{2\gamma}{\gamma-1} \frac{R_u}{M_w} T_c \left[1 - \left(\frac{p_e}{p_c} \right)^{(\gamma-1)/\gamma} \right]} \quad (10.59)$$

where T_c is the temperature in the combustion chamber and p_e/p_c is the pressure ratio from combustion to exit. We see that for high exit velocity it is desirable to have a high combustion temperature and a low molecular weight for the propellants. The pressure ratio and the ratio of specific heats (γ) play a less significant role. Solid rockets have a lower exit velocity as they have both a relatively high molecular weight and a lower combustion temperature. They partially make up for that deficiency with a higher mass flow rate (recall that thrust is the product of mass flow rate and effective exit velocity). However, overall solid rockets tend to be less efficient.

10.3.2 Throat Size

The throat must be large enough for the required mass flow rate. As derived in Section A.4.2 the smallest throat area to sustain a given propellant mass flow rate (\dot{m}_f) is given by:

$$A_t = \frac{\dot{m}_f}{p_c} \sqrt{\frac{R_u T_c}{M_w \gamma}} \left(\frac{\gamma + 1}{2} \right)^{\frac{\gamma+1}{\gamma-1}} \quad (10.60)$$

where p_c, T_c is the pressure and temperature in the combustion chamber.

10.3.3 Combustion Chamber Sizing

Next, we need to determine how large the combustion chamber should be relative to the throat (Fig. 10.7). In Section A.4.3 we derive the total pressure loss that occurs in the combustion chamber as a function of the area ratio between the combustion chamber and the throat. This function is plotted in Fig. 10.8.

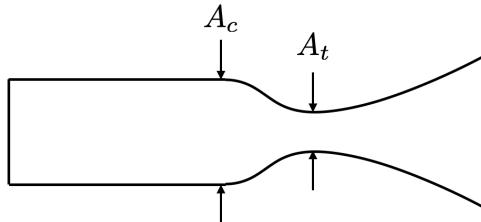


Figure 10.7: Depiction of cross-sectional area of combustion chamber (A_c) and cross-sectional area of throat (A_t).

Ideally the total pressure ratio is one, which means there is no total pressure loss. A smaller total pressure ratio means there is more entropy, decreased efficiency, and less thrust. We can see that as the combustion area (relative to the throat area) becomes larger than the total pressure ratio approaches one. As that area ratio becomes larger then the flow can accelerate to Mach one at the throat from a slower initial Mach number in the combustion chamber. A lower Mach number in the combustion chamber will experience less total pressure losses. However, the tradeoff is that making the combustion chamber larger adds more weight and drag. From the figure we can see that there are diminishing returns, and that the results are relatively insensitive to the propellant (γ of 1.1 and 1.6 covers a reasonable range of propellants). Based on these tradeoffs, most rockets settle on an area ratio of about three ($A_c/A_t = 3$).

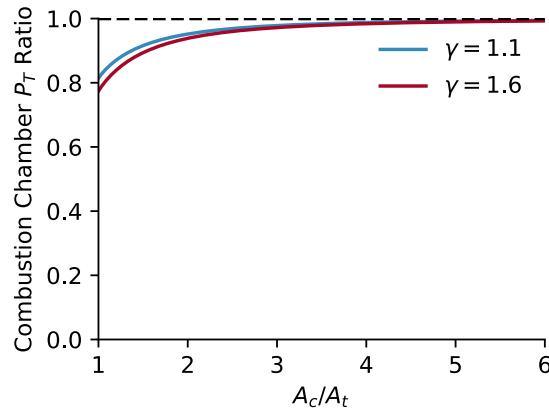


Figure 10.8: Total pressure ratio across combustion chamber for different propellant mixtures as a function of the combustion chamber cross-sectional area relative to the throat area.

10.3.4 Exit Area Sizing

The remaining cross-sectional area to determine is the exit area. Determining the ratio of the exit area relative to the throat area is one of the most critical parameters for rocket nozzle design. It is called the *nozzle expansion ratio* and is important enough that it has its own symbol:

$$\epsilon = \frac{A_e}{A_t} \quad (10.61)$$

In Section A.4.4 we derive how the thrust coefficient of the rocket varies as a function of nozzle expansion ratio. This relationship is plotted in Fig. 10.9, where the thrust is normalized by total pressure in the combustion chamber and the area of the throat. The figure shows that a propellant mixture with low γ is desirable, and that large expansion ratios are necessary for high thrust. As a reference, the space shuttle main engines have a large expansion ratio of 77 (Fig. 10.10).

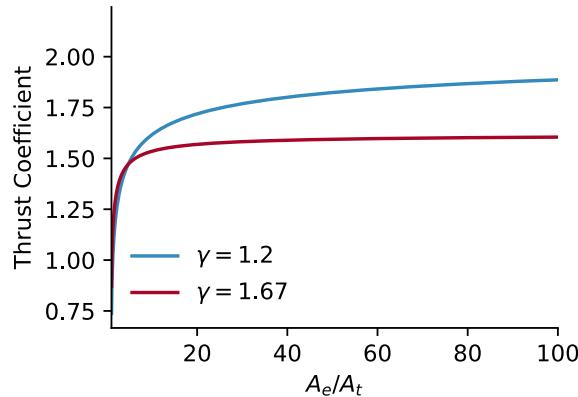


Figure 10.9: Thrust coefficient of a rocket as a function of the nozzle expansion ratio for two different propellants.

As discussed in the last section, ideal performance occurs when the exit pressure equals the atmospheric pressure $p_e = p_\infty$. If we determine the design altitude where we want to optimize performance (generally just after max-q) then we can determine the pressure ratio p_c/p_e required assuming ideal expansion at our design altitude ($p_e = p_\infty$). Then, as derived in Section A.4.4 the Mach number at the exit can be determined from this pressure ratio by:

$$M_e = \sqrt{\left[\left(\frac{p_c}{p_e} \right)^{\frac{\gamma-1}{\gamma}} - 1 \right] \frac{2}{\gamma-1}} \quad (10.62)$$

The requisite expansion ratio to achieve that exit Mach number is then given by:

$$\epsilon = \frac{A_e}{A_t} = \frac{1}{M_e} \left[\frac{2}{\gamma+1} \left(1 + \frac{\gamma-1}{2} M_e^2 \right) \right]^{\frac{(\gamma+1)}{2(\gamma-1)}} \quad (10.63)$$

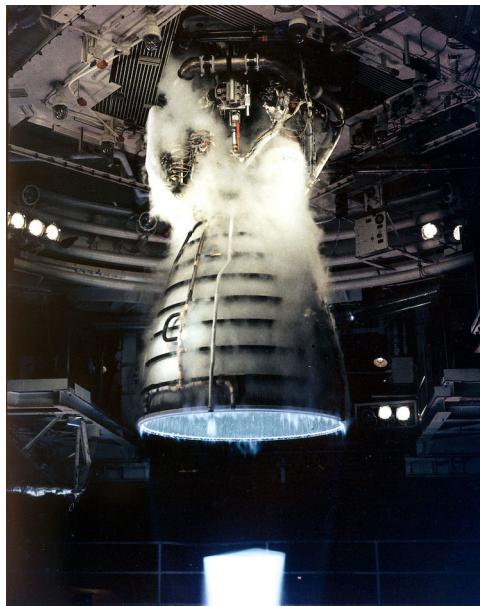


Figure 10.10: Space shuttle main engine. Picture from [NASA](#), Wikimedia Commons, public domain.

We have now determined all the requisite cross-sectional areas of the nozzle, and now need to determine the lengths/shapes of the nozzle.

10.4 Lengths

A notional nozzle with typical convergence and divergence angles is shown in Fig. 10.11. The acceleration of the subsonic flow through the converging section is relatively insensitive to the length or nozzle shape. Thus, these parameters are less important for initial design, and a typical angle of 60° will be assumed.

The performance of the nozzle is much more sensitive to the shape of the diverging section. Ideally, the exit angle (15° in Fig. 10.11) should be small so that streamlines are relatively straight and flow separation does not occur. However, the smaller the exit angle then the longer the nozzle has to be to reach the desired exit area and the more weight that is added. The divergence of the nozzle also causes a loss in the exit velocity and thus in the thrust. For a conical nozzle, if the streamlines were to remain straight then the exit area would be the red projected area shown in

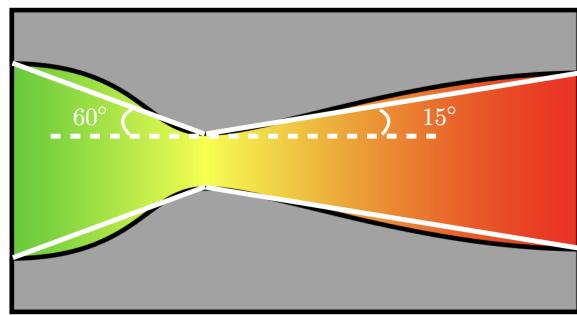


Figure 10.11: Typical equivalent angles for the converging and diverging portions of a rocket nozzle. Figure adapted from [HorsePunchKid](#), Wikimedia Commons, CC BY-SA 3.0.

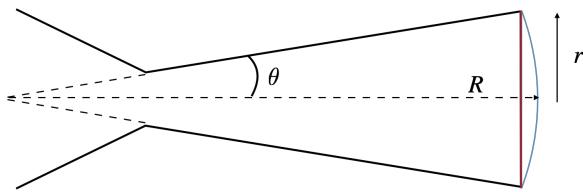


Figure 10.12: Depiction of the increase in area normal to the streamlines caused by flow divergence.

Fig. 10.12, which is:

$$A_{straight} = \pi r^2 \quad (10.64)$$

However, because the streamlines must diverge to follow the conic shape, the actual exit area is the spherical shape shown in blue. Its area can be found from integrating in spherical coordinates as:

$$A_{spherical} = 2\pi R^2(1 - \cos \theta) \quad (10.65)$$

Because this area is larger, for a given mass flow rate, the exit velocity is less than the ideal case and thus the thrust is reduced. We can calculate exactly how much the thrust is reduced by computing the area ratio between these two scenarios.

$$\lambda = \frac{A_{straight}}{A_{spherical}} = \frac{\pi r^2}{2\pi R^2(1 - \cos \theta)} \quad (10.66)$$

From the figure we see that $\sin \theta = r/R$. Thus:

$$\lambda = \frac{\sin^2 \theta}{2(1 - \cos \theta)} \quad (10.67)$$

If we use the trig identity $\sin^2 \theta + \cos^2 \theta = 1$ and simplify we have the *divergence loss factor*:

$$\lambda = \frac{1 + \cos \theta}{2} \quad (10.68)$$

where θ is the exit angle. This factor is plotted in Fig. 10.13. We see that an exit angle of 15° corresponds to about a 2% loss in thrust.² However, this is not the primary limiting factor. After 20° flow separation becomes increasingly likely. Typical, nozzles have exit angles between approximately 12° – 18° [1].

For a conical nozzle, if we know the angles shown in Fig. 10.11 and the areas of the ends, then from geometry we can compute the requisite lengths. The length of the converging section is:

$$L_{con} = \frac{r_c - r_t}{\tan \theta_{con}} \quad (10.69)$$

and the length of the diverging section is:

$$L_{div} = \frac{r_e - r_t}{\tan \theta_{div}} \quad (10.70)$$

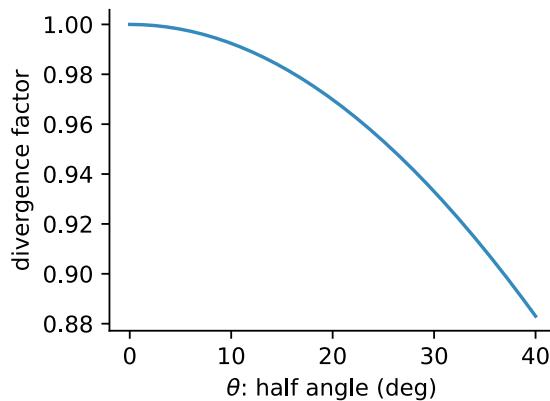


Figure 10.13: Plot of the reduction in the exit velocity as a function of the spreading angle in the diverging section of a nozzle.

Conical nozzles are simple, but they are not the most efficient. The most commonly used shape for rocket nozzles, is the bell (or Rao) nozzle. This shape is optimized to minimize total pressure losses and direct the flow so that streamlines are nearly parallel at the exit. The length of a bell nozzle is typically defined relative to a conical nozzle with a 15° exit angle. For example, an 80% bell nozzle means that it has 80% of the length of a 15° conical nozzle with the same area ratios (Fig. 10.14). Length in this case refers to the distance from the throat to the exit.

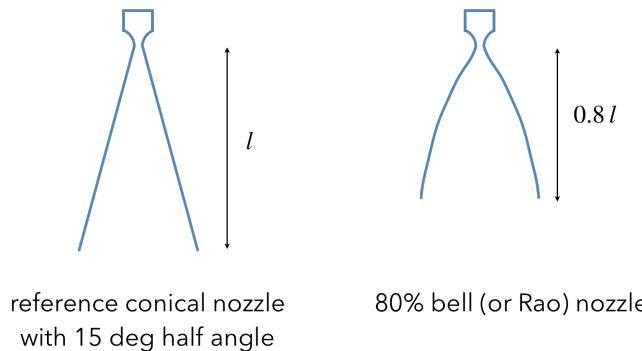


Figure 10.14: The length of a bell nozzle is referenced to a 15° conical nozzle, with the same area ratios.

²Actually this term should only be multiplied against the actual exit velocity (first term in Eq. (9.1)), rather than the total thrust, but this is a small difference.

The length of the combustion chamber is more variable and is generally based on past designs. However, as a estimate one can use the characteristic mixing length defined as:

$$L^* = \frac{V_c}{A_t} \quad (10.71)$$

which is the ratio of the volume of the combustion chamber to the area of the throat. A typical value for L^* is 0.8–3.0 m [1]. However, this provides only a rough guideline to sizing as proper combustion chamber sizing depends on the pressure, propellant, etc.

Bibliography

- [1] Sutton, G. P. and Biblarz, O., *Rocket Propulsion Elements*, John Wiley & Sons, Nov 2016.

CHAPTER 11

Orbital Mechanics

Orbital mechanics is the study of the motion of spacecraft or rockets in space. Aerodynamics has little to no role here—gravity is the main force at work. The study of orbital mechanics usually encompasses a full semester. In this chapter we will only hit a few highlights to give a flavor of the topic.

11.1 Gravity

Newton's equation for the gravitational force between two objects is given by:

$$F = \frac{GMm}{r^2} \quad (11.1)$$

where r is the distance between the objects, M and m are the masses of the two objects, and

$$G = 6.674 \times 10^{-11} \frac{\text{N m}^2}{\text{kg}^2} \quad (11.2)$$

is the universal gravitation constant. We use the notation of capital M and lowercase m, because for our application one mass is usually large (e.g., earth, or another celestial body), and the other mass is comparatively small (e.g., a spacecraft).

In vector form the equation is expressed as:

$$\vec{F} = -\frac{GMm}{r^2} \hat{r} \quad (11.3)$$

which shows that the gravitational force acts along the line joining the center of gravity of the two objects (\hat{r}). The negative sign is necessary because the force is attractive (opposite the direction of \hat{r}). Newtonian gravitation, and other aspects of classical physics, have been supplanted by Einstein's general theory of relativity, however the extra accuracy of general relativity is negligible for our applications and so we will stick with the classical theory of gravity.

Gravity is a conservative force, or in other words it is path independent. This means that only the relative positions between two objects matter. In contrast, friction is path dependent and so is a nonconservative force. For a conservative force we can define an associated potential function and a potential energy. Potential functions are useful because they reduce a three-dimensional vector function (e.g., force of gravity) into a one-dimensional scalar function (e.g., gravitational potential energy). The potential energy that a body of mass m experiences in a gravitational field created by M is:

$$U = -\frac{GMm}{r} \quad (11.4)$$

Notice that the potential energy is negative and approaches zero as you move infinitely far away from the object ($r \rightarrow \infty$). The form of potential energy gives rise to the concept of gravity wells. As r becomes small the potential energy becomes highly negative like falling into a well (until you reach the surface of the object). A gravity well is depicted in Fig. 11.1. This concept can help us understand why it takes a lot of energy to leave the earth and reach to the moon, but not nearly as much energy on the return trip. For example, Fig. 11.2 shows the potential energy of a 1,000 kg spacecraft as it approaches the earth and the moon. Because the earth is over 80 times more massive than the moon it has a much deeper gravity well.

For our applications usually one of the bodies is earth and so it is often convenient to combine G and M into one number where M_e is the mass of the earth:

$$\mu_e = GM_e = 3.986 \times 10^{14} \frac{\text{m}^3}{\text{s}^2} \quad (11.5)$$

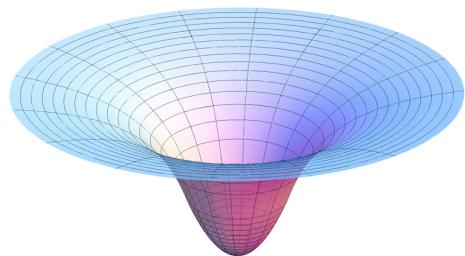


Figure 11.1: Gravitational potential well around a uniformly dense spherical body. Figure from [AllenMcC](#), Wikimedia Commons, CC BY-SA 3.0.

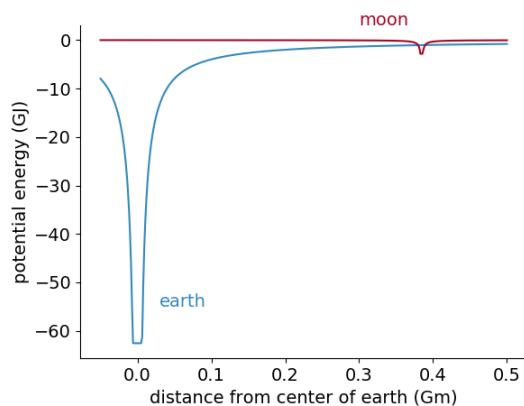


Figure 11.2: Potential energy of a 1,000 kg spacecraft as it approaches earth and the moon. The energy is truncated once the radius of the body is reached.

This parameter is called the geocentric gravitational constant. The effective mean radius of the earth is

$$R_e = 6,371 \text{ km} \quad (11.6)$$

Thus, an object on the surface of the earth experiences a force:

$$F = \frac{GM_e m}{R_e^2} \quad (11.7)$$

For applications near the surface of the earth we typically lump all of the constant values into the acceleration of gravity g :

$$F = mg \quad (11.8)$$

where

$$g_{sl} = \frac{GM_e}{R_e^2} \quad (11.9)$$

Note, that using the values specified above will give a value that is close to but not exactly the same as the standard gravity value: $g = 9.80665 \text{ m/s}^2$.

There are various factors that cause the discrepancy, but one is that the earth is not perfectly spherical and so the “radius” of the earth is not exactly well defined.

Often we need the acceleration of gravity at other locations besides at sea level. We can always go back to Newton’s law of gravity to compute the acceleration, but it is often convenient to perform the computation relative to the acceleration of gravity at sea level. Using the same form as Eq. (11.9) we can compute g for any altitude:

$$g = \frac{GM_e}{(R_e + h)^2} \quad (11.10)$$

where for convenience we have defined the radius of interest relative to the radius of the earth. If we combine this equation with Eq. (11.9) we have:

$$g = g_{sl} \frac{R_e^2}{(R_e + h)^2} = g_{sl} \frac{1}{(1 + h/R_e)^2} \quad (11.11)$$

where the altitude h is zero starting from the radius of the earth. Note that even at low earth orbit (about 2,042 km above the surface of the earth) the acceleration of gravity is still significant (about 57% of the acceleration of gravity at sea level).

The concept of potential energy is convenient in answering many basic orbital mechanics questions. For example, what is the escape velocity from the surface of the earth? The *escape velocity* is the minimum speed needed for a projectile object (no additional thrust after launch) to completely escape from the gravitational influence of some body. From conservation of energy we can equate an initial state where both potential and kinetic energy exists, to a final state infinitely far away. If infinitely far away the potential energy is zero, and if we are after the minimum speed to get there then the final kinetic energy is zero:

$$\frac{1}{2}mV_{\text{escape}}^2 - \frac{GM_e m}{R_e} = 0 \quad (11.12)$$

We can now solve for the escape velocity (using values for the earth):

$$V_{\text{escape}} = \sqrt{\frac{2GM_e}{R_e}} = 11,186 \text{ m/s or } 25,020 \text{ mph!} \quad (11.13)$$

Actually, the rotation of the earth (465 m/s at the equator) can reduce the required speed a little if the launch is to the east in the direction of earth's rotation. This is why most rocket launch facilities are located close to the equator where the surface rotation speed is largest, and in a location where ocean is to the east. Escape velocity really corresponds to a projectile (the speed required to launch an unpowered object into space), and so doesn't exactly represent the speed a rocket must obtain at launch. However, the same energy balance can be used to estimate the ΔV required to obtain a desired orbit, like the 9 km/s to low earth orbit that we used in Chapter 9. Using the force of gravity we can answer the question: what velocity is necessary to maintain a circular orbit? When in a circular orbit the force of gravity is a centripetal force:

$$\frac{GMm}{r^2} = \frac{mV_{\text{orbit}}^2}{r} \quad (11.14)$$

Solving for the velocity required to maintain a circular orbit:

$$V_{\text{orbit}} = \sqrt{\frac{GM}{r}} \quad (11.15)$$

Notice that the orbit velocity is exactly a factor of $1/\sqrt{2}$ less than (or about 71% of) the escape velocity for a given radius.

11.2 Orbits

We are primarily interested in the motion of one object around another (e.g., a spacecraft around the earth). If the only force acting on these objects is gravity, then the scenario is called the *two body problem*. For two-body problems we can derive the orbit equation that describes all possible trajectories.

The orbit equation can be derived from Newton's laws of motion but the vector math is fairly complex and it is easier (mathematically) to derive from an energy perspective using Lagrange's equation. Lagrange's equation may be unfamiliar to you, but it is commonly used in analyzing complex dynamic systems. Lagrange's equation defines the Lagrangian function \mathcal{L} as:

$$\mathcal{L} = T - U \quad (11.16)$$

where T is the kinetic energy and U is the potential energy. Lagrange's equation is:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{x}_i} \right) - \frac{\partial \mathcal{L}}{\partial x_i} = 0 \text{ for } i = 1, \dots, n \quad (11.17)$$

where x_i are **generalized coordinate** directions.

In the following we will take the perspective of finding the motion of mass m relative to mass M (which is usually what we want when M is large).

We already know that the potential energy of mass m is:

$$U = -\frac{GMm}{r} \quad (11.18)$$

we now just need its kinetic energy. In cartesian coordinates the kinetic energy of mass m is (assuming motion in a plane, which is the case for the two-body problem):

$$T = \frac{1}{2}m(\dot{x}^2 + \dot{y}^2) \quad (11.19)$$

Because we have used r in the previous equation it will be convenient to define the kinetic energy in polar coordinates instead:

$$T = \frac{1}{2}m(\dot{r}^2 + (r\dot{\theta})^2) \quad (11.20)$$

Now we can form the Lagrangian:

$$\mathcal{L} = \frac{1}{2}m(\dot{r}^2 + (r\dot{\theta})^2) + \frac{GMm}{r} \quad (11.21)$$

We have two coordinates: r and θ so we need to apply Lagrange's equation twice. First, we take derivatives with respect to θ :

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{\theta}} \right) - \frac{\partial \mathcal{L}}{\partial \theta} = 0 \quad (11.22)$$

$$\frac{d}{dt} (mr^2\dot{\theta}) = 0 \quad (11.23)$$

This equation simply says that the angular momentum of the object ($mr^2\dot{\theta}$) is constant, which we already know to be true as there is no external torque applied to the system.

The second equation uses Lagrange's equation with our other coordinate r :

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{r}} \right) - \frac{\partial \mathcal{L}}{\partial r} = 0 \quad (11.24)$$

$$\frac{d}{dt} (m\dot{r}) - mr\dot{\theta}^2 + \frac{GMm}{r^2} = 0 \quad (11.25)$$

$$m\ddot{r} - mr\dot{\theta}^2 + \frac{GMm}{r^2} = 0 \quad (11.26)$$

$$\ddot{r} - r\dot{\theta}^2 + \frac{GM}{r^2} = 0 \quad (11.27)$$

$$(11.28)$$

Notice that the mass of the spacecraft (m) cancelled out. This is a differential equation, with derivatives of both r and θ , that can be used to describe the motion of the spacecraft. However, we can simplify this expression further by recalling from Eq. (11.23) that the angular momentum is constant. As is convention, we define the the angular momentum per unit mass:

$$h = r^2\dot{\theta} \quad (11.29)$$

The differential equation then simplifies to one only in terms of r :

$$\ddot{r} - \frac{h^2}{r^3} + \frac{\mu}{r^2} = 0 \quad (11.30)$$

where h , and $\mu = GM$ and both constants.

This differential equation has an analytic solution:

$$r = \frac{h^2/\mu}{1 + e \cos(\theta - \theta_0)}$$

(11.31)

where θ_0 is a phase angle, and e is the eccentricity which can be expressed as [1]:

$$e = \sqrt{1 + 2 \left(\frac{h}{\mu} \right)^2 \frac{(T - U)}{m}} \quad (11.32)$$

It turns out the orbit equation $r(\theta)$ is exactly of the same form as a [conic section in polar coordinates](#). What this means is that there are only a few different possible trajectories for the spacecraft:

- If $T < U$ the orbit shape is an ellipse
- If $T = U$ the orbit shape is a parabola
- If $T > U$ the orbit shape is an hyperbola

These three scenarios are visualized in Fig. 11.3. The parabola is just a boundary case separating bounded orbits (ellipse) from unbounded orbits (hyperbola). A circular orbit is the limiting case of an ellipse where the eccentricity (e) is zero. Note that eccentricities cannot be negative.

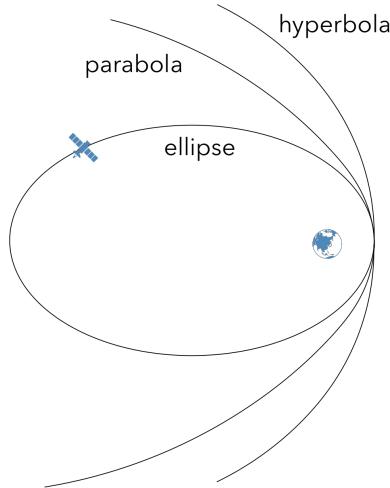


Figure 11.3: The possible orbits for a two body problem.

The unknown constants h and e and θ_0 can be determined from the burnout conditions of the spacecraft. After burnout, the spacecraft is unpowered and follows the path described by the orbit equation, at least until there is another powered trajectory adjustment. Consider the scenario depicted in Fig. 11.4. At burnout we need to know the radius

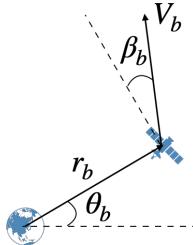


Figure 11.4: Variables used to determine trajectory based on conditions at burnout (end of powered phase).

from the center of the earth (or whatever object we are orbiting), the latitude (θ_b), the speed (V_b), and any angle offset from the local horizon (β_b). From these parameters we can then compute the angular momentum (which depends only on the perpendicular component of velocity because it is a cross product):

$$h = r^2 \dot{\theta} = r V_\theta = r_b V_b \cos(\beta_b) \quad (11.33)$$

the eccentricity (using Eq. (11.32)), and the phase angle by solving the orbit equation for θ_0 :

$$\theta_0 = \theta_b - \cos^{-1} \left[\left(\frac{h^2}{\mu r_b} - 1 \right) \frac{1}{e} \right] \quad (11.34)$$

Figure 11.5 shows an example elliptical orbit around the earth, with the parameters used defined in the caption.

Elliptical orbits (with circular as a special case) are the most common so we will dive into elliptical orbits a bit deeper. The point of closest approach during the orbit is called the *periapsis* (or *perigee* if the body being orbited is earth), and the point on the orbit furthest away from M is called the *apoapsis* (or *apogee* if orbiting earth). We have already showed that the angular momentum is constant so that the quantity rV_θ is constant throughout the orbit. This means that the orbiting object is traveling fastest at periapsis, and slowest at apoapsis.

More generally, we can compute the velocity at any point in the orbit though it takes a bit more work to derive. We know that the total energy of the orbit is conserved (written per unit mass below, since mass is constant):

$$\frac{V^2}{2} - \frac{\mu}{r} = \epsilon \quad (11.35)$$

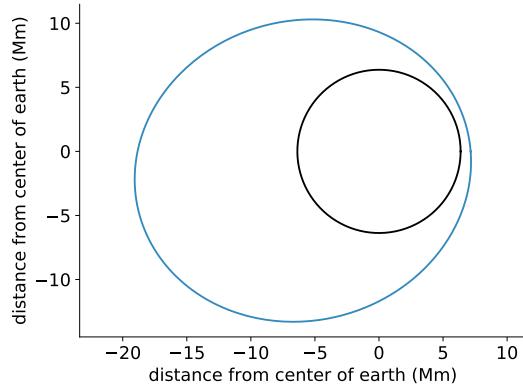


Figure 11.5: An example orbit around the earth (in black) for an object with a burnout speed of 9 km/s, and an altitude of 800 km above the earth, a latitude of 30° and an inclination angle (β) of 5°.

where e is some unknown constant. In order to find a value for e we just need to evaluate the energy at some location where the energy state is known. To do this, we will first evaluate the orbit equation at periapsis (r_p) and apoapsis (r_a), where we will choose the phase angle to be zero (it can be chosen arbitrarily) with periapsis at 0° and apoapsis at 180°. The result is:

$$r_p = \frac{h^2/\mu}{1+e} \quad (11.36)$$

$$r_a = \frac{h^2/\mu}{1-e} \quad (11.37)$$

Evaluating the energy equation at r_p gives (after simplification):

$$\epsilon = -\frac{\mu^2}{2h^2}(1-e^2) \quad (11.38)$$

This expression applies to any orbit (not just elliptic) because we evaluated at periapsis and all of the orbits have a point of closest approach.

In the case of the ellipse we can simplify this using the geometry of an ellipse. If we call the semimajor axis a then by definition:

$$2a = r_p + r_a \quad (11.39)$$

and using the definitions in Eqs. (11.36) and (11.37) and simplifying gives:

$$a = \frac{h^2}{\mu(1 - e^2)} \quad (11.40)$$

Next, we substitute this value into the total energy (Eq. (11.38)) to give:

$$\epsilon = -\frac{\mu}{2a} \quad (11.41)$$

which shows that the total energy of an ellipse is always negative as expected. Plugging this into Eq. (11.35) gives:

$$\frac{V^2}{2} - \frac{\mu}{r} = -\frac{\mu}{2a} \quad (11.42)$$

Finally, we can solve this for V to compute the speed at any point in the elliptical orbit as a function of r :

$$V = \sqrt{\mu \left(\frac{2}{r} - \frac{1}{a} \right)}$$

(11.43)

Figure 11.6 illustrates the parameters that appear in this equation.

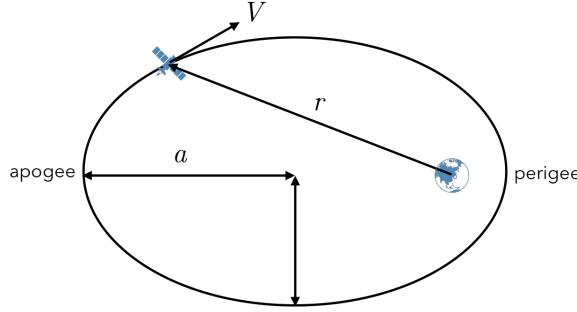


Figure 11.6: Depiction of the semimajor axis and velocity along the orbit as a function of radius.

For the case of a circle $a = r$ and the expression simplifies to $V = \sqrt{\mu/r}$, which we found already in Eq. (11.15). As an aside: for a parabolic orbit we know that the kinetic energy equals the potential energy so we can solve for the velocity in a parabolic orbit as a function of r giving $V = \sqrt{2\mu/r}$. This is the escape velocity (Eq. (11.13)). Thus, another interpretation of the escape velocity is the speed needed to reach a parabolic trajectory.

11.3 Hohmann Transfer

We often would like to move the spacecraft from one orbit to another. For example, Fig. 11.7 depicts the starting orbit of a spacecraft, called the *parking orbit*, and a desired final orbit called the *target orbit*.

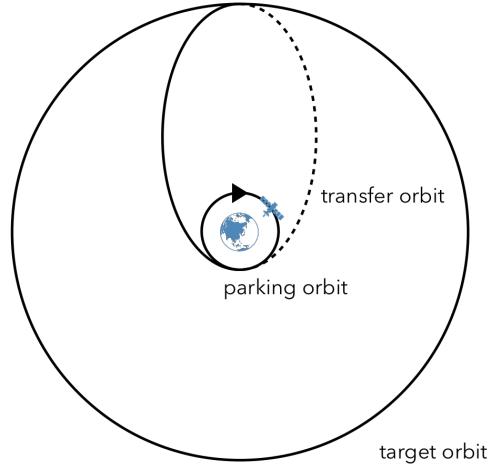


Figure 11.7: Depiction of the Hohmann transfer from a circular parking orbit to a circular target orbit via an elliptical transfer orbit.

There are various types of orbit transfers, but the simplest is the Hohmann transfer. The *Hohmann transfer* is the minimum energy transfer between two coplanar circular orbits with a common center point (it can be generalized to coplanar elliptical orbits). The transfer occurs along an elliptical orbit as depicted in the figure. An impulsive increase in speed transfers into the elliptical transfer orbit, and a second impulsive increase in speed transfers into the target orbit. It may seem counterintuitive that the spacecraft must speed up twice, because we know that the speed required to maintain the larger circular orbit is less than that needed to maintain the smaller circular orbit. The reason you need to speed up on both transfers is that the spacecraft velocity decreases as it moves from periapsis to apoapsis on the far side of the transfer ellipse.

We can compute the velocity required to maintain the circular parking orbit from Eq. (11.15). We also can compute the velocity needed to enter the elliptical transfer orbit at periapsis from Eq. (11.43). The difference in velocities gives the required ΔV and from the rocket equation discussed in Chapter 9 we can then find the necessary amount of propellant to

complete the transfer. Similar calculations are then repeated for moving into the target orbit. If performing the opposite maneuver, the ΔV would be the same in both cases, except that the rockets would be fired opposite to the direction of motion to slow the spacecraft down.

Hohmann transfers can also be used for interplanetary trajectories where the sun, rather than the earth, is the central point of the orbit. In our solar system the orbital planes of all planets are not exactly coplanar, but they are fairly close. Another added difficulty of an interplanetary transfer is that it is not enough just to get to the target orbit, we need to arrive at the right moment so we can rendezvous with the planet along that orbit.

Thus, we have to wait for the right window of time to launch the spacecraft. For example, a Hohmann transfer from earth to Mars is only possible every 2.13 years [2]. Once the spacecraft reached Mars (in 258.8 days), it would have to wait at least 453.8 days to start the return journey leading to a total minimum mission time of 2.66 years. Each return window that was missed would add another 2.13 years to the mission [2].

Bibliography

- [1] Anderson, J. D., *Introduction to Flight*, McGraw-Hill Professional, 2005.
- [2] Curtis, H. D., *Orbital Mechanics: For Engineering Students*, Elsevier, Jul 2015.

APPENDIX A

Supplementary Material

A.1 Skin Friction Coefficient for a Flat Plate with Transition

In Chapter 3 we discussed how to estimate the skin friction coefficient for both a laminar and turbulent flat plate boundary layer. However, most of the time an airfoil is not fully laminar or fully turbulent but rather transitions somewhere along the airfoil. Accounting for this in a rational way is the subject of this section.

In the following derivation we will utilize the figure below. We start with a laminar region that extends up to the transition location x_t . The turbulent boundary layer begins from there and continues to L . The turbulent boundary layer begins with some thickness, but we can imagine tracing it back to figure out where the turbulent boundary layer would have needed to start in order to develop the same thickness at x_t . We will call this the effective distance x_e , and it is defined as a distance upstream of the transition location. The three regions that we will perform separate calculations on are labeled as 1, 2, and 3.

Region 1 is the easiest. We simply use the laminar skin friction coefficient acting over a distance x_t .

$$C_{f1} = \frac{1.328}{\sqrt{Re_{xt}}} \quad (\text{A.1})$$

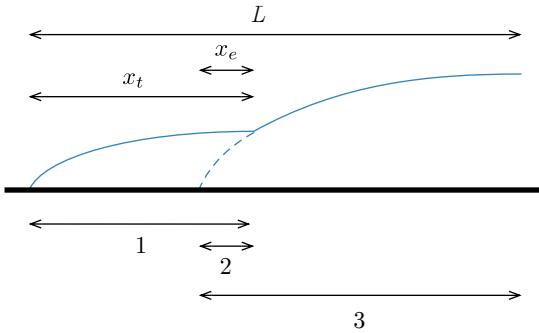


Figure A.1:

Region 3 is not as straightforward as we can't just start the turbulent calculation from the transition location. Our formulas are based on starting from the beginning of the boundary layer. Thus, we first need to calculate the effective distance x_e .

The size of the boundary layer can be measured by its momentum thickness θ (you will learn more about this if in an aerodynamics course). The Blasius solution provides a formula for this thickness, which we will evaluate over a distance x_t

$$\theta_l = \frac{0.664x_t}{\sqrt{Re_{xt}}} \quad (\text{A.2})$$

This is the thickness of the boundary layer at the transition location. To work backwards in the turbulent boundary layer we use a similar formula, but instead use one for a turbulent boundary layer. The Schlichting empirical formula for momentum thickness of a turbulent boundary layer evaluate at a distance x_e is [1]:

$$\theta_t = \frac{0.036x_e}{Re_{xe}^{0.2}} \quad (\text{A.3})$$

These two momentum thicknesses must be equal. Since we can already

compute θ_l we can now solve for x_e :

$$\theta_t = \frac{0.036x_e}{Re_{xe}^{0.2}} \quad (\text{A.4})$$

$$\theta_l = \frac{0.036x_e}{Re_{xe}^{0.2}} \quad (\text{A.5})$$

$$\theta_l = \frac{0.036x_e}{\left(\frac{\rho V x_e}{\mu}\right)^{0.2}} \quad (\text{A.6})$$

$$x_e^{0.8} = \frac{\theta_l \left(\frac{\rho V}{\mu}\right)^{0.2}}{0.036} \quad (\text{A.7})$$

$$x_e^{0.8} = \frac{\theta_l \left(\frac{\rho V L}{\mu}\right)^{0.2}}{0.036 c^{0.2}} \quad (\text{A.8})$$

$$(x_e/c)^{0.8} = \frac{(\theta_l/c) Re_L^{0.2}}{0.036} \quad (\text{A.9})$$

$$\frac{x_e}{c} = \left(\frac{(\theta_l/c) Re_L^{0.2}}{0.036} \right)^{1.25} \quad (\text{A.10})$$

Now that we know the effective distance we can compute the skin friction coefficient for Region 3 using the Schlichting formula. The distance that this acts over is $L - x_t + x_e$. We will call this x_f (for fictitious since it includes the dashed fictitious portion that is not actually turbulent).

$$C_{f3} = \frac{0.074}{Re_{xf}^{0.2}} \quad (\text{A.11})$$

We now need to deal with the overlapping Region 2, which we have double counted. We just calculated the skin friction coefficient for all of Region 3 but we need to subtract off Region 2. This is easy enough as it is just a turbulent boundary layer acting over a distance x_e .

$$C_{f2} = \frac{0.074}{Re_{xe}^{0.2}} \quad (\text{A.12})$$

We are now ready to put everything together. However, skin friction coefficients can't just be added as they are normalized over different areas. It is the forces that we can add. Recall that

$$c_f = \frac{\tau}{q_\infty} \quad (\text{A.13})$$

Thus the drag force acting over region j is

$$D_j = C_{fj} q_\infty x_j (1) \quad (\text{A.14})$$

where x_j is the distance that the skin friction acts over (with unit width into the page). Once we add all the drag forces we would like to normalize by the length L so that we have one final skin friction coefficient.

$$C_f = \frac{\tau}{q_\infty} \quad (\text{A.15})$$

$$= \frac{\sum_j D_j / A}{q_\infty} \quad (\text{A.16})$$

$$= \frac{\sum_j D_j}{c q_\infty} \quad (\text{A.17})$$

$$= \frac{\sum_j C_{fj} q_\infty x_j}{c q_\infty} \quad (\text{A.18})$$

$$= \sum_j C_{fj} \frac{x_j}{c} \quad (\text{A.19})$$

$$(\text{A.20})$$

Thus, rather than sum the skin friction coefficients we need to sum the skin friction coefficients times the area that act over. This is exactly the same situation when adding drag coefficients that use different areas for normalization.

The total skin friction coefficient is then

$$C_f = C_{f1} \left(\frac{x_t}{c} \right) + C_{f3} \left(\frac{x_f}{c} \right) - C_{f2} \left(\frac{x_e}{c} \right) \quad (\text{A.21})$$

The last thing to note is that if we are given the Reynolds number for one length scale

$$Re_c = \frac{\rho V c}{\mu} \quad (\text{A.22})$$

we can easily calculate it at another length scale as follows:

$$Re_{xt} = \frac{\rho V x_t}{\mu} = Re_c \frac{x_t}{c} \quad (\text{A.23})$$

Summary:

The final algorithm is as follows (which we do nondimensionally):

Given input:

$$\frac{x_t}{c} \text{ and } Re_c \quad (\text{A.24})$$

Reynolds number and skin friction coefficient over Region 1:

$$Re_{xt} = Re_c \frac{x_t}{c} \quad (\text{A.25})$$

$$C_{f1} = \frac{1.328}{\sqrt{Re_{xt}}} \quad (\text{A.26})$$

Effective distance:

$$\frac{\theta}{c} = \frac{0.664(x_t/c)}{\sqrt{Re_{xt}}} \quad (\text{A.27})$$

$$x_e/c = \left(\frac{(\theta/c)Re_c^{0.2}}{0.036} \right)^{1.25} \quad (\text{A.28})$$

Reynolds number and skin friction coefficient over Region 2:

$$Re_{xe} = Re_c(x_e/c) \quad (\text{A.29})$$

$$C_{f2} = \frac{0.074}{Re_{xe}^{0.2}} \quad (\text{A.30})$$

Reynolds number and skin friction coefficient over Region 3:

$$\frac{x_f}{c} = 1 - \frac{x_t}{c} + \frac{x_e}{c} \quad (\text{A.31})$$

$$Re_{xf} = Re_c \frac{x_f}{c} \quad (\text{A.32})$$

$$C_{f3} = \frac{0.074}{Re_{xf}^{0.2}} \quad (\text{A.33})$$

Total skin friction coefficient

$$C_f = C_{f1} \left(\frac{x_t}{c} \right) + C_{f3} \left(\frac{x_f}{c} \right) - C_{f2} \left(\frac{x_e}{c} \right) \quad (\text{A.34})$$

A.2 Supersonic Wave Drag Equations

There is a lift-dependent and a volume-dependent component of wave drag.

Lift-dependent R.T. Jones derived the following expression for combined vortex and wave drag due to lift of a minimum-drag ellipse with a given aspect ratio [?].

$$C_D = \frac{C_L^2}{\pi AR} \sqrt{1 + (M^2 - 1) \left(\frac{\pi AR}{4} \right)^2} \quad (\text{A.35})$$

We want just the wave drag so we will subtract off the vortex drag $C_L^2/(\pi AR)$ (we assume $e_{inv} = 1$ because that is consistent with the above minimum-drag expression, which has an elliptic distribution of lift in all directions).

$$C_{Dc,l} = \frac{C_L^2}{\pi AR} \left[\sqrt{1 + (M^2 - 1) \left(\frac{\pi AR}{4} \right)^2} - 1 \right] \quad (\text{A.36})$$

Supersonic wings are generally longer than they are wide. A better fit to experimental data is observed if we use an equivalent ellipse based on length and area rather than span and area (aspect ratio). The area of an ellipse of length l and width b is

$$S_{ellipse} = \frac{\pi bl}{4} \quad (\text{A.37})$$

We equate the gross area of the wing S_g with that of an equivalent ellipse and then solve for the span in terms of gross wing area and length:

$$b = \frac{4S_g}{\pi l} \quad (\text{A.38})$$

Thus, we can express aspect ratio in terms of length and area, recalling that aspect ratio is defined using the reference area:

$$AR = \frac{b^2}{S_{ref}} \quad (\text{A.39})$$

$$= \frac{b^2 S_g}{S_{ref} S_g} \quad (\text{A.40})$$

$$= \frac{16S_g^2}{\pi^2 l^2 S_g} \frac{S_g}{S_{ref}} \quad (\text{A.41})$$

$$= \frac{16S_g}{\pi^2 l^2} \frac{S_g}{S_{ref}} \quad (\text{A.42})$$

$$(\text{A.43})$$

Substituting into the the wave drag expression

$$C_{Dc,l} = \frac{C_L^2 \pi l^2 S_{ref}}{16S_g^2} \left[\sqrt{1 + (M^2 - 1) \left(\frac{4S_g^2}{\pi l^2 S_{ref}} \right)^2} - 1 \right] \quad (\text{A.44})$$

For convenience let's define the following terms:

$$\beta = \sqrt{M^2 - 1} \quad (\text{A.45})$$

(the only reason why we use a sqrt is because this is a common term used in supersonic aerodynamic analysis and this is how it is defined).

$$r = \frac{\pi l^2}{4S_g} \quad (\text{A.46})$$

$$S_r = \frac{S_g}{S_{ref}} \quad (\text{A.47})$$

then the expression becomes

$$C_{Dc,l} = C_L^2 \frac{r}{4S_r} \left[\sqrt{1 + \left(\frac{\beta S_r}{r} \right)^2} - 1 \right] \quad (\text{A.48})$$

Volume Dependent For an elliptic wing with biconvex airfoil sections R.T. Jones derived the following expression for supersonic volume-dependent wave drag [9]:

$$C_{Dc,v} = \frac{t^2}{a^2} \frac{(\beta^2 + 2(a/e)^2)}{(\beta^2 + (a/e)^2)^{1.5}} \quad (\text{A.49})$$

where a is the ellipse semiaxis in the flight direction, e is the orthogonal semiaxis, and t is the maximum thickness in the center of the wing.

In terms of our typically used wing quantities $a = c/2$ and $e = b/2$ (where c is the maximum chord at the center of the wing). Making the substitutions yields:

$$C_{Dc,v} = 4 \left(\frac{t}{c} \right)^2 \frac{(\beta^2 + 2(c/b)^2)}{(\beta^2 + (c/b)^2)^{1.5}} \quad (\text{A.50})$$

however we again would like to eliminate span from the expression in favor of the length parameter (c in this case).

We use the previous expression relating span to gross area and length (c), but interchange with the variable l we used previously

$$C_{Dc,v} = 4 \left(\frac{t}{c} \right)^2 \frac{(\beta^2 + 2(l^2\pi/4S_g)^2)}{(\beta^2 + (l^2\pi/4S_g)^2)^{1.5}} \quad (\text{A.51})$$

Using the new variable definitions simplifies the expression.

$$C_{Dc,v} = 4 \left(\frac{t}{c} \right)^2 \frac{(\beta^2 + 2r^2)}{(\beta^2 + r^2)^{1.5}} \quad (\text{A.52})$$

Finally, in order to add this drag to the other we need to use the same reference area. This volume-dependent drag is defined relative to the gross area. Thus, we must multiply by S_g/S_{ref} so that it is normalized relative to the same reference area.

$$C_{Dc,v} = 4 \left(\frac{t}{c} \right)^2 S_r \frac{(\beta^2 + 2r^2)}{(\beta^2 + r^2)^{1.5}} \quad (\text{A.53})$$

Fuselage Volume Wave Drag The Sears-Haack minimum drag solution for a body of revolution with fixed length and diameter is

$$C_{Dc,v} = 4\pi^2 \left(\frac{r}{l} \right)^2 \quad (\text{A.54})$$

We make the substitution $r = d/2$ and because this drag is based on frontal area, we need to multiply by the frontal area and divide by S_{ref} so that all of our drag coefficients use the same reference area:

$$C_{Dc,v} = 4\pi^2 \frac{d^2}{4l^2} \frac{\pi d^2}{4S_{ref}} \quad (\text{A.55})$$

or after simplifying:

$$C_{Dc,v} = \frac{\pi^3}{4} \frac{d^2}{(l/d)^2 S_{ref}} \quad (\text{A.56})$$

A.3 Rocket Altitude

We start with Tsiolkovsky's rocket equation:

$$\Delta V = C \ln(MR) - gt_b \cos \theta \quad (\text{A.57})$$

which gives the change in velocity over some time period. We would like to integrate this equation once more to find a final altitude.

The final altitude will give given by:

$$h = \int_0^{t_b} V \cos \theta dt \quad (\text{A.58})$$

$$= \int_0^{t_b} (V_0 + \Delta V) \cos \theta dt \quad (\text{A.59})$$

$$= \int_0^{t_b} (V_0 + C \ln(MR) - gt \cos \theta) \cos \theta dt \quad (\text{A.60})$$

$$(\text{A.61})$$

A numerical integration would be best, but for the purposes of creating a simple closed-form expression we will assume that C , g , and θ are constant during the integration interval. The mass ratio is never constant so we must express in terms of time t

$$MR = \frac{m_0}{m_0 - \dot{m}t} \quad (\text{A.62})$$

and we will assume that \dot{m} is constant. The integral then becomes:

$$h = \int_0^{t_b} (V_0 + C \ln\left(\frac{m_0}{m_0 - \dot{m}t}\right) - gt \cos \theta) \cos \theta dt \quad (\text{A.63})$$

$$= \int_0^{t_b} V_0 + C \ln m_0 - C \ln(m_0 - \dot{m}t) - gt \cos \theta \cos \theta dt \quad (\text{A.64})$$

$$= V_0 \cos \theta t_b + C \ln m_0 \cos \theta t_b - g \frac{t_b^2}{2} \cos^2 \theta \quad (\text{A.65})$$

$$- C \cos \theta \left[\left(t - \frac{m_0}{\dot{m}} \right) \ln(m_0 - \dot{m}t) - t \right]_0^{t_b} \quad (\text{A.66})$$

After a bunch of algebraic manipulations the expression simplifies to:

$$h = \left[V_0 + C \left(1 - \frac{\ln MR}{MR - 1} \right) - \frac{1}{2} g t_b \cos \theta \right] t_b \cos \theta \quad (\text{A.67})$$

A.4 Rocket Nozzle Performance

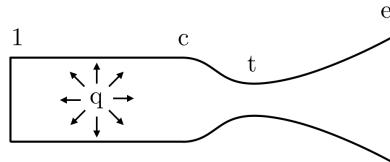


Figure A.2: A notional nozzle depicted station 1 upstream of combustion, station c after combustion, station t at the throat, and station e at the exit.

We use the same nozzle nomenclature depicted previously, but shown again in Fig. A.2. The derivations in this section assume familiarity with compressible flow. The only equation that may not be immediately recognizable is the way the mass balance is defined. The equation for mass flow rate is unambiguous, but when applied in a mass balance there are various ways it can be expressed. As a mass balance is perhaps the most important equation for nozzle analysis this will be derived first.

The mass flow rate is:

$$\dot{m} = \rho V A \quad (\text{A.68})$$

However, for a compressible flow Mach numbers are more fundamental than velocity. If we use the ideal gas law ($p = \rho RT$), express velocity in terms of the Mach number and the speed of sound we have:

$$\dot{m} = \frac{p}{RT} M \sqrt{\gamma R T A} \quad (\text{A.69})$$

$$= p M A \sqrt{\frac{\gamma}{RT}} \quad (\text{A.70})$$

$$(A.71)$$

Additionally, total pressures and temperatures are much more useful quantities in a nozzle than static pressures and temperatures. The reason is that for an adiabatic flow total temperature is constant, and for an isentropic flow total pressure is constant (whereas temperature and pressure are not constant). Using the definitions for total pressure and total temperature gives:

$$\dot{m} = \frac{p_t A}{\sqrt{T_t}} \sqrt{\frac{\gamma}{R}} M \left(1 + \frac{\gamma - 1}{2} M^2\right)^{-\frac{(\gamma+1)}{2(\gamma-1)}} \quad (\text{A.72})$$

This expression is becoming complex, but we will make it easier to use shortly. First, we define a new function that depends only on the Mach number for a given fluid:

$$f(M) = M \left[\frac{2}{\gamma + 1} \left(1 + \frac{\gamma - 1}{2} M^2 \right) \right]^{\frac{-(\gamma+1)}{2(\gamma-1)}} \quad (\text{A.73})$$

This expression is actually the ratio of the cross-sectional area of a (potentially fictitious) location in the nozzle where the flow is sonic, relative to the local cross-sectional area (A^*/A in compressible flow nomenclature). This is really beside the point, we could make up any function we wanted, but this ends up being a convenient choice. If we substitute this choice for $f(M)$ into the mass flow rate equation we can express it as:

$$\dot{m} = \frac{p_T A}{\sqrt{T_T}} \sqrt{\frac{\gamma}{R}} \left(\frac{2}{\gamma + 1} \right)^{\frac{\gamma+1}{2(\gamma-1)}} f(M) \quad (\text{A.74})$$

Usually we don't care about the mass flow rate per se (although we will use it once below). More generally, because the mass flow rate is constant through the nozzle, we wish to equate the mass flow rate at two points in the nozzle. Because the fluid doesn't change, γ and R are constant and so the mass balance simplifies to the following important equation:

$$\frac{p_{T1} A_1 f(M_1)}{\sqrt{T_{T1}}} = \frac{p_{T2} A_2 f(M_2)}{\sqrt{T_{T2}}} \quad (\text{A.75})$$

This equation is rather simple to remember, though most of the complexity is hidden in the function $f(M)$. This function is defined above Eq. (A.73) and is plotted in Fig. A.3. There is no need to memorize the function, but the key points to remember are that $f(1) = 1$ (which ends up being convenient as we often use the throat as one of our evaluation points), and that the value of the function is always between 0 and 1. This equation will either be used to compute $f(M)$ for a given Mach number, or to solve for the Mach number that corresponds to a given value of $f(M)$. The latter can be solved with a root-finding method or interpolating with a table.

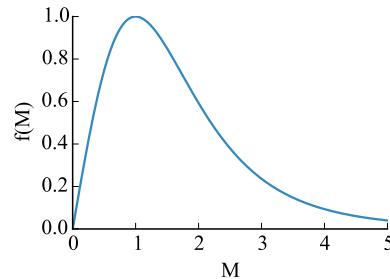


Figure A.3: Plot of the function $f(M)$ used in the mass balance.

A.4.1 Exit Velocity

Applying conservation of energy from after combustion to the exit we have:

$$h_c + \frac{V_c^2}{2} = h_e + \frac{V_e^2} \quad (\text{A.76})$$

Solving for the exit velocity:

$$V_e = \sqrt{2(h_c - h_e) + V_c^2} \quad (\text{A.77})$$

$$= \sqrt{2C_p(T_c - T_e) + V_c^2} \quad (\text{A.78})$$

$$= \sqrt{2C_p T_c \left[1 - \frac{T_e}{T_c} \right] + V_c^2} \quad (\text{A.79})$$

$$(\text{A.80})$$

Using the isentropic relationships between combustion to exit:

$$V_e = \sqrt{2C_p T_c \left[1 - \left(\frac{p_e}{p_c} \right)^{(\gamma-1)/\gamma} \right] + V_c^2} \quad (\text{A.81})$$

We can express C_p in terms of R .

$$V_e = \sqrt{2 \frac{\gamma R}{\gamma - 1} T_c \left[1 - \left(\frac{p_e}{p_c} \right)^{(\gamma-1)/\gamma} \right] + V_c^2} \quad (\text{A.82})$$

We express R in terms of the universal gas constant. The last term is small compared to the first term and so we neglect it (which can be seen by

expanding the term in terms of the Mach number). Physically, this means that the velocity in the combustion chamber is relatively small.

$$V_e = \sqrt{\frac{2\gamma}{\gamma-1} \frac{R_u}{M_w} T_c \left[1 - \left(\frac{p_e}{p_c} \right)^{(\gamma-1)/\gamma} \right]} \quad (\text{A.83})$$

A.4.2 Throat Size

The nozzle throat must be large enough for the requisite mass flow rate. The general formula for the mass flow rate for a compressible flow was given previously in Eq. (A.72). If we apply this equation at the throat, assuming that it is choked we have:

$$\dot{m} = \frac{p_{Tt} A_t}{\sqrt{T_t}} \sqrt{\frac{\gamma}{R}} \left(1 + \frac{\gamma-1}{2} \right)^{-\frac{(\gamma+1)}{2(\gamma-1)}} \quad (\text{A.84})$$

We can now solve for the required throat area given the mass flow rate:

$$A_t = \frac{\dot{m}}{p_{Tt}} \sqrt{\frac{RT_{Tt}}{\gamma}} \left(\frac{\gamma+1}{2} \right)^{\frac{\gamma+1}{2(\gamma-1)}} \quad (\text{A.85})$$

The flow is isentropic from after combustion to the throat so $T_{Tt} = T_{Tc}$ and $p_{Tt} = p_{Tc}$. Furthermore, like before, we assume that the velocity right after combustion is negligibly small so that $T_{Tc} = T_c$ and $p_{Tc} = p_c$. Making these substitutions and simplifying yields:

$$A_t = \frac{\dot{m}}{p_c} \sqrt{\frac{R_u T_c}{M_w \gamma}} \left(\frac{\gamma+1}{2} \right)^{\frac{\gamma+1}{\gamma-1}} \quad (\text{A.86})$$

This is the smallest throat area that can sustain a given mass flow rate.

A.4.3 Combustion Chamber Sizing

We now determine an appropriate size for the combustion chamber area relative to the throat area. Using a mass balance from c to t the flow is isentropic and adiabatic so the mass balance simplified to:

$$A_c f(M_c) = A_t f(M_t) \quad (\text{A.87})$$

The throat will be choked so $f(M_t) = 1$ and we can solve for the Mach number at the end of the combustion chamber for a given area ratio:

$$f(M_c) = \frac{A_t}{A_c} \quad (\text{A.88})$$

We can now compute the total pressure loss across the combustion chamber from 1 to c . There is heat addition between these two stations so we need to use the Rayleigh line equations:

$$\frac{p_{Tc}}{p_{T1}} = \frac{1 + \gamma M_1^2}{1 + \gamma M_c^2} \left[\frac{1 + \frac{\gamma-1}{2} M_c^2}{1 + \frac{\gamma-1}{2} M_1^2} \right]^{\frac{\gamma}{\gamma-1}} \quad (\text{A.89})$$

The speed at station 1 is essentially zero so the expression simplifies to:

$$\frac{p_{Tc}}{p_{T1}} = \frac{1}{1 + \gamma M_c^2} \left[1 + \frac{\gamma - 1}{2} M_c^2 \right]^{\frac{\gamma}{\gamma-1}} \quad (\text{A.90})$$

This equation allows us to quantify the total pressure loss in the combustion chamber as a function of the area ratio between the throat and the combustion chamber.

This equation can also justify our assumption that the Mach number in the combustion chamber is low. As discussed in Chapter 10 because of the total pressure losses a typical area ratio is $A_c/A_t = 3$. From the above mass balance this corresponds to a Mach number in the combustion chamber of approximately $M_c = 0.2$. This means that the total temperature differs from the static temperature by less than a percent, and total pressure and static pressure differ by less than a few percent. For students with a compressible flow background it would be better just to use the total pressure/temperature values as it really isn't much harder. However, the small error is justifiable for the audience of this text as it allows students without a compressible flow background to use familiar static quantities.

A.4.4 Exit Area Sizing

We start with a mass balance from the throat to the exit. It is isentropic and adiabatic between these stations, and the Mach number at the throat is one, so the area ratio is just a function of the exit Mach number:

$$A_t = A_e f(M_e) \quad (\text{A.91})$$

$$\Rightarrow \epsilon = \frac{A_e}{A_t} = \frac{1}{f(M_e)} \quad (\text{A.92})$$

Now we just need to determine the requisite exit Mach number for a desired pressure ratio. From the definition of total pressure:

$$\frac{p_{Te}}{p_e} = \left(1 + \frac{\gamma - 1}{2} M_e^2\right)^{\frac{\gamma}{\gamma-1}} \quad (\text{A.93})$$

We can solve this for M_e :

$$M_e = \sqrt{\left[\left(\frac{p_{Te}}{p_e}\right)^{\frac{\gamma-1}{\gamma}} - 1\right] \frac{2}{\gamma-1}} \quad (\text{A.94})$$

Because the flow is isentropic from c to e then $p_{Te} = p_{Tc}$. Furthermore, we can make the same approximation as before, assuming the Mach number at c is low so that $p_{Tc} \approx p_c$. With that approximation we have:

$$M_e = \sqrt{\left[\left(\frac{p_c}{p_e}\right)^{\frac{\gamma-1}{\gamma}} - 1\right] \frac{2}{\gamma-1}} \quad (\text{A.95})$$

To see the effectiveness of large expansion ratios let us determine the performance in terms of thrust coefficient. Recall that the thrust is given by:

$$T = \dot{m}_f V_e + (p_e - p_\infty) A_e \quad (\text{A.96})$$

We will normalize the thrust by the total pressure in the combustion chamber $p_{Tc} \approx p_c$ and the area of the throat.

$$C_T = \frac{\dot{m}_f V_e + (p_e - p_\infty) A_e}{p_c A_t} \quad (\text{A.97})$$

Thus mass flow rate, using properties at the exit is, $\dot{m}_f = \rho_e V_e A_e$, giving a thrust coefficient of:

$$C_T = \frac{\rho_e V_e^2 A_e + (p_e - p_\infty) A_e}{p_c A_t} \quad (\text{A.98})$$

Factoring out some common terms:

$$C_T = \frac{A_e}{A_t} \frac{p_e}{p_c} \left(\frac{\rho_e V_e^2}{p_e} + 1 - \frac{p_\infty}{p_e} \right) \quad (\text{A.99})$$

The speed of sound is:

$$a = \sqrt{\frac{\gamma p}{\rho}} \quad (\text{A.100})$$

Making the substitution for p_e/ρ_e gives:

$$C_T = \frac{A_e}{A_t} \frac{p_e}{p_c} \left(\frac{\gamma V_e^2}{a_e^2} + 1 - \frac{p_\infty}{p_e} \right) \quad (\text{A.101})$$

If we assume ideal expansion then $p_\infty = p_e$ and the equation simplifies to:

$$C_T = \frac{A_e}{A_t} \frac{p_e}{p_c} \gamma M_e^2 \quad (\text{ideal expansion}) \quad (\text{A.102})$$