Contents

1 Prerequisites........................................ 5
   A few high-school formulas.......................... 5
   Polynomials....................................... 5
   Exponentiation and logarithm.......................... 6
   Geometric series.................................. 6
   Trigonometric formulas............................. 6
   Solving equations................................... 7
   Differentiation.................................... 8
   Basic limits...................................... 9
   Integration........................................ 10
   Geometry........................................... 12
   Matrices............................................ 13
   Complex numbers.................................... 14

I ORDINARY DIFFERENTIAL EQUATIONS............ 15
   General review of differential equations........... 17

2 First-Order Differential Equations............... 19
   'Trivial' equation.................................. 19
   Separable equation.................................. 20
   Scale-independent equation.......................... 21
   Linear equation...................................... 22
   Bernoulli equation.................................. 24
   Exact equation....................................... 25
   More 'exotic' equations.............................. 27
   Applications........................................ 29

3 Second Order Differential Equations............... 33
   Reducible to first order............................. 33
   Linear equation...................................... 34
   With constant coefficients........................... 37
   Cauchy equation..................................... 41

4 Third and Higher-Order Linear ODEs............... 43
   Polynomial roots.................................... 43
   Constant-coefficient equations...................... 47
5 Sets of Linear, First-Order, Constant-Coefficient ODEs 49
Matrix Algebra ........................................ 49
Set (system) of differential equations .................. 54
Non-homogeneous case ................................ 59

6 Power-Series Solution 65
The main idea ........................................... 65
Sturm-Liouville eigenvalue problem .................... 68
Method of Frobenius ................................... 71
A few Special functions of Mathematical Physics .... 76

II VECTOR ANALYSIS 83

7 Functions in Three Dimensions – Differentiation 85
3-D Geometry (overview) ................................ 85
Fields ..................................................... 94
Optional: Curvilinear coordinates ..................... 96

8 Functions in 3-D – Integration 99
Line Integrals ............................................ 99
Double integrals ....................................... 102
Surfaces in 3-D .......................................... 106
Surface integrals ....................................... 107
’Volume’ integrals ..................................... 110
Review exercises ....................................... 113

III COMPLEX ANALYSIS 119

9 Complex Functions – Differentiation 121
Preliminaries ............................................. 121
Introducing complex functions ......................... 122
Chapter summary ........................................ 126

10 Complex Functions – Integration 127
Integrating analytic functions ......................... 128
Contour integration ..................................... 129
Applications ............................................ 132
Chapter 1  PREREQUISITES

A few high-school formulas

\[(a + b)^2 = a^2 + 2ab + b^2\]
\[(a + b)^3 = a^3 + 3a^2b + 3ab^2 + b^3\]
\[\vdots\]

\[
\begin{array}{cccc}
1 \\
1 & 1 \\
1 & 2 & 1 \\
\end{array}
\]

the coefficients follow from PASCAL’S TRIANGLE: 1 3 3 1 , the expansion is

\[
\begin{array}{cccc}
1 & 4 & 6 & 4 & 1 \\
1 & 5 & 10 & 10 & 5 & 1 \\
\end{array}
\]

....... called BINOMIAL.

Also:

\[a^2 - b^2 = (a - b)(a + b)\]
\[a^3 - b^3 = (a - b)(a^2 + ab + b^2)\]
\[\vdots\]

(do you know how to continue)?

Understand the basic rules of algebra: addition and multiplication, individually, are COMMUTATIVE and ASSOCIATIVE, when combined they follow the DISTRIBUTIVE law

\[(a + b + c)(d + e + f + h) = ad + ae + af + ah + bd + be + bf + bh + cd + ce + cf + ch\]

(each term from the left set of parentheses with each term on the right).

Polynomials

their DEGREE, the notion of individual COEFFICIENTS, basic operations including SYNTHETIC DIVISION, e.g.

\[
\begin{align*}
(x^3 - 3x^2 + 2x - 4) \div (x - 2) & = x^2 - x \quad \text{[QUOTIENT]} \\
x^3 - 2x^2 & \quad \text{(subtract)} \\
-x^2 + 2x - 4 & \\
-x^2 + 2x & \quad \text{(subtract)} \\
-4 & \quad \text{[REMAINDER]}
\end{align*}
\]

which implies that \((x^3 - 3x^2 + 2x - 4) = (x^2 - x)(x - 2) - 4\). The quotient’s degree equals the degree of the dividend (the original polynomial) minus the degree of the divisor. The remainder’s degree is always less than the degree of the divisor. When the remainder is zero, we have found two FACTORS of the dividend.
Exponentiation and logarithm

Rules of exponentiation:

\[ a^A \cdot a^B = a^{A+B} \]
\[ (a^A)^B = a^{AB} \]

Also note that

\[ (a^A)^B \neq a^{(AB)} \]

The solution to

\[ a^x = A \]

is:

\[ x = \log_a(A) \]

(the inverse function to exponentiation). When \( a = e \) (= 2.7183...), this is written as

\[ x = \ln(A) \]

and called **natural logarithm**.

Its basic rules are:

\[ \ln(A \cdot B) = \ln(A) + \ln(B) \]
\[ \ln(A^B) = B \cdot \ln(A) \]
\[ \log_a(A) = \frac{\ln(A)}{\ln(a)} \]

Geometric series

First **infinite**:

\[ 1 + a + a^2 + a^3 + a^4 + ... = \frac{1}{1-a} \]

when \( |a| < 1 \) (understand the issue of series convergence)

and then **finite** (truncated):

\[ 1 + a + a^2 + a^3 + ... + a^N = \frac{1 - a^{N+1}}{1-a} \]

valid for all \( a \neq 1 \), (we don’t need \( a = 1 \), why?).

Trigonometric formulas

such as, for example

\[ (\sin a)^2 + (\cos a)^2 \equiv 1 \]

and

\[ \sin(\alpha + \beta) = \sin \alpha \cos \beta + \sin \beta \cos \alpha \]
\[ \cos(\alpha + \beta) = \cos \alpha \cos \beta - \sin \alpha \sin \beta \]

Our angles are always in **radians**.

**Inverse** trigonometric functions: \( \arcsin(x) \) – don’t use the \( \sin^{-1}(x) \) notation!
Solving equations

**Single Unknown**

**Linear**: \(2x = 7\) (quite trivial).

**Quadratic**: \(x^2 - 5x + 6 = 0\) \(\Rightarrow\) \(x_{1,2} = \frac{5 \pm \sqrt{(-5)^2 - 4 \cdot 1 \cdot 6}}{2} = \frac{5 \pm \sqrt{25 - 24}}{2} = \frac{5 \pm 1}{2}\)

This further implies that \(x^2 - 5x + 6 = (x - 3)(x - 2)\).

**Cubic** and beyond: will be discussed when needed.

Can we solve other (non-polynomial) equations? Only when the left hand side involves a composition of functions which we know how to invert individually, for example

\[
\ln \left( \sin \left( \frac{1}{1 + x^2} \right) \right) = -3
\]

should pose no difficulty.

On the other hand, a simpler looking

\[
\sin(x) = \frac{x}{2}
\]

can be solved only numerically, usually by **Newton’s technique**, which works as follows:

To solve an equation of the \(f(x) = 0\) type, we start with an initial value \(x_0\) which should be reasonably close to a root of the equation (found graphically), and then follow the tangent straight line from \([x_0, f(x_0)]\) till we cross the \(x\)-axis at \(x_1\). This is repeated until the consecutive \(x\)-values no longer change.

In summary:

\[
x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}
\]

To solve \(\sin(x) - \frac{x}{2} = 0\) (our previous example) we first look at its graph

![Graph](image)

which indicates that there is a root close to \(x = 1.9\). Choosing this as our \(x_0\) we get: \(x_1 = 1.9 - \frac{\sin(1.9) - 0.95}{\cos(1.9) - 0.5} = 1.895506\), \(x_2 = x_1 - \frac{\sin(x_1) - \frac{x_1}{2}}{\cos(x_1) - \frac{1}{2}} = 1.895494\), after which the values no longer change. Thus \(x = 1.895494\) is a solution (in this case, not the only one) to the original equation.
More Than One Unknown

We all know how to solve linear sets (systems) of equations: \(2 \times 2\) (for sure, e.g. \(2x - 3y = 4 \iff [\text{add } 3 \times \text{Eq.}2 \text{ to Eq.}1] 11x = -11 \Rightarrow x = -1 \text{ and } y = -2\), \(3 \times 3\) (still rather routinely, I hope), \(4 \times 4\) (gets tedious).

How about any other (nonlinear) case? We can try eliminating one unknown from one equation and substituting into the other equation, but this will work (in the \(2 \times 2\) case) only if we are very lucky. We have to admit that we don’t know how to solve most of these equations (and, in this course, we will have to live with that).

Differentiation

Interpretation: Slope of a tangent straight line.

Using three basic rules (PRODUCT, QUOTIENT and CHAIN) one can differentiate just about any function, repeatedly if necessary (i.e. finding the second, third, ... derivative), for example:

\[
\frac{d}{dx} \left( x \sin x^2 \right) = \sin x^2 + 2x^2 \cos x^2
\]

[Note that \(\sin x^2 \equiv \sin(x^2)\) and not \((\sin x)^2\) – I will always be careful with my notation, using parentheses whenever an ambiguity may arise].

The main formulas are

\[
\frac{d}{dx} (x^\alpha) = \alpha x^{\alpha-1}
\]

and

\[
\frac{d}{dx} (e^{\beta x}) = \beta e^{\beta x}
\]

The product rule can be extended to the second derivative:

\[
(f \cdot g)'' = f'' \cdot g + 2f' \cdot g' + f \cdot g''
\]

the third:

\[
(f \cdot g)''' = f''' \cdot g + 3f'' \cdot g' + 3f' \cdot g'' + f \cdot g'''
\]

e tc. (Pascal’s triangle again).

Partial Derivatives

Even when the function is bivariate (of two variables), or multivariate (of several variables), we always differentiate with respect to only one of the variables at a time (keeping the others constant). Thus, no new rules are needed, the only new thing is a more elaborate notation, e.g.

\[
\frac{\partial^3 f(x, y)}{\partial x^2 \partial y}
\]

(A function of a single variable is called univariate).
Taylor (Maclaurin) Expansion

\[ f(x) = f(0) + xf'(0) + \frac{x^2}{2}f''(0) + \frac{x^3}{3!}f'''(0) + \frac{x^4}{4!}f^{(4)}(0) + \ldots \]

Alternate notation for \( f^{(4)}(0) \) is \( f^{(4)}(0) \).

Remember the expansions of at least these functions:

\[ e^x = 1 + x + \frac{x^2}{2} + \frac{x^3}{3!} + \frac{x^4}{4!} + \ldots \]
\[ \sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots \]
\[ \ln(1 + x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots \quad \text{(no factorials)} \]

and of course

\[ \frac{1}{1-x} = 1 + x + x^2 + x^3 + x^4 + \ldots \]

The bivariate extension of Taylor’s expansion (in a rather symbolic form):

\[ f(x, y) = f(0, 0) + x \frac{\partial f(0, 0)}{\partial x} + y \frac{\partial f(0, 0)}{\partial y} + \frac{1}{2} \left\{ \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right\}^2 f(0, 0) + \frac{1}{3!} \left\{ \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right\}^3 f(0, 0) + \ldots \]

where the partial derivatives are to be applied to \( f(x, y) \) only. Thus

\[ \left\{ \frac{\partial}{\partial x} + y \frac{\partial}{\partial y} \right\}^3 f \equiv x^3 \frac{\partial^3 f}{\partial x^3} + 3x^2y \frac{\partial^3 f}{\partial x^2 \partial y} + 3xy^2 \frac{\partial^3 f}{\partial x \partial y^2} + y^3 \frac{\partial^3 f}{\partial y^3} \]

etc.

Basic limits

Rational expressions, e.g.

\[ \lim_{n \to \infty} \frac{2n^2 + 3}{n^2 - 4n + 1} = 2 \]

Special limit:

\[ \lim_{n \to \infty} \left( 1 + \frac{1}{n} \right)^n = e \]

and also

\[ \lim_{n \to \infty} \left( 1 + \frac{a}{n} \right)^n = e^a \equiv \exp(a) \]

(introducing an alternate notation for \( e^a \).

L’hôpital rule (to deal with the \( \frac{0}{0} \) case):

\[ \lim_{x \to 0} \frac{\sin x}{x} = \lim_{x \to 0} \frac{(\sin x)'}{(x)'} = \frac{\cos(0)}{1} = 1 \]
Integration

**Interpretation:** Area between $x$-axis and $f(x)$ (up is positive, down is negative).

Basic formulas:

$$
\int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1}, \quad \alpha \neq -1
$$

$$
\int e^{\beta x} dx = \frac{e^{\beta x}}{\beta}
$$

$$
\int \frac{dx}{x} = \ln |x|
$$

etc. (use tables).

Useful techniques and tricks:

1. **Substitution** (change of variable), for example

   $$
   \int \frac{dx}{\sqrt{5x-2}} = \frac{2}{5} \int \frac{z dz}{z} = \frac{2}{5} z = \frac{2}{5} \sqrt{5x-2}
   $$

   where $z = \sqrt{5x-2}$, thus $x = \frac{z^2 + 2}{5}$, and $\frac{dx}{dz} = \frac{2}{5} z \Rightarrow dx = \frac{2}{5} z dz$.

2. **By parts**

   $$
   \int f g' dx = f g - \int f' g dx
   $$

   for example

   $$
   \int xe^{-x} dx = \int x(-e^{-x})' dx = -xe^{-x} + \int e^{-x} dx = -(x+1)e^{-x}.
   $$

3. **Partial fractions:** When integrating a rational expression such as, for example

   $$
   \frac{1}{(1+x)^2(1+x^2)}
   $$

   first rewrite it as

   $$
   \frac{a}{1+x} + \frac{b}{(1+x)^2} + \frac{c + dx}{1+x^2}
   $$

   then solve for $a, b, c, d$ based on

   $$
   1 = a(1+x)(1+x^2) + b(1+x^2) + (c + dx)(1+x)^2 \quad (#1)
   $$

   This can be done most efficiently by substituting $x = -1 \Rightarrow 1 = 2b \Rightarrow b = \frac{1}{2}$.

   When this $(b = \frac{1}{2})$ is substituted back into (#1), $(1+x)$ can be factored out, resulting in

   $$
   \frac{1-x}{2} = a(1+x^2) + (c + dx)(1+x) \quad (#2)
   $$

   Then substitute $x = -1$ again, get $a = \frac{1}{2}$, and further simplify (#2) by factoring out yet another $(1+x)$:

   $$
   \frac{-x}{2} = c + dx
   $$

   yielding the values of $c = 0$ and $d = -\frac{1}{2}$. In the new form, the function can be integrated easily. ■
Basic properties:

\[ \int [cf(x) + dg(x)]dx = c \int f(x) \, dx + d \int g(x) \, dx \quad \text{(linear)} \]
and

\[ \int_a^b f(x) \, dx = \int_a^c f(x) \, dx + \int_c^b f(x) \, dx \quad \text{(try picture it)} \]

These properties are shared with (actually inherited from) SUMMATION; we all know that, for example

\[ \sum_{i=1}^{\infty} \left( 3a^i - \frac{1}{i^2} \right) = 3 \sum_{i=1}^{\infty} a^i - \sum_{i=1}^{\infty} \frac{1}{i^2} \]

Have some basic ideas about double (and, eventually, triple) integration, e.g.

\[ \iint_R (x^2y - 3xy^3) \, dA \]

where \( R \) is a REGION OF INTEGRATION (it may be a square, a triangle, a circle and such), and \( dA \) a symbol for an infinitesimal area (usually visualized as a small rectangle within this region).

The simplest case is the so called\ SEPARABLE\ integral, where the function (of \( x \) and \( y \)) to be integrated is a \textit{product} of a function of \( x \) (only) times a function of \( y \) (only), and the region of integration is a \textit{generalized rectangle} (meaning that the limits of integration don’t depend on each other, but any of them may be infinite).

For example:

\[ \int_0^\infty \int_1^3 \sin(x)e^{-y} \, dA = \int_1^3 \sin(x) \, dx \times \int_0^\infty e^{-y} \, dy \]

which is a product of two ordinary integrals.

In general, a double integral must be converted to two consecutive \textit{univariate} integrations, the first in \( x \) and the second in \( y \) (or vice versa – the results must be identical). Notationally, the \( x \) and \( y \) limits, and the \( dx \) and \( dy \) symbols should follow proper ’nesting’ (similar to two sets of parentheses), since that is effectively the logic of performing the integration in general.

More complicated cases may require a \textbf{change of variables}. This may lead to POLAR COORDINATES and requires understanding of the JACOBIAN (details to be discussed when needed).
Geometry

In Two Dimensions

Know that \( y = 1 - x \) is an equation of a straight line, be able to identify its slope and intercept.

The collection of points which lie below this straight line and are in the first quadrant defines the following two-dimensional region:

There are two other ways of describing it:

1. \( 0 \leq y \leq 1 - x \) (conditional \( y \)-range), where \( 0 \leq x \leq 1 \) (marginal \( x \)-range) – visualize this as the triangle being filled with vertical lines, the marginal range describing the \( x \)-scale 'shadow' of the region.

2. \( 0 \leq x \leq 1 - y \) (conditional \( x \)-range), where \( 0 \leq y \leq 1 \) (marginal \( y \)-range) – horizontal lines.

The two descriptions will usually not be this symmetric; try doing the same thing with \( y = 1 - x^2 \) (a branch of a parabola). For the 'horizontal-line' description, you will get: \( 0 \leq x \leq \sqrt{1 - y^2} \), where \( 0 \leq y \leq 1 \).

Regions of this kind are frequently encountered in two-dimensional integrals. The vertical (or horizontal)-line description facilitates constructing the inner and outer limits of the corresponding (consecutive) integration. Note that in this context we don't have to worry about the boundary points being included (e.g. \( 0 \leq x \leq 1 \)) or excluded (\( 0 < x < 1 \)), this makes no difference to the integral's value.

Recognize equation of a circle (e.g. \( x^2 + y^2 - 3x + 4y = 9 \)) be able to identify its center and radius; also that of parabola and hyperbola.

In Three Dimensions

Understand the standard (rectangular) coordinate system and how to use it to display points, lines and planes.

Both (fixed) points and (free) vectors have three components, written as: [2, -1, 4] (sometimes organized in a column). Understand the difference between the two: point describes a single location, vector defines a length (or strength, in physics), direction and orientation.
Would you be able to describe a three-dimensional region using \( x, y \) and \( z \) (consecutive) ranges, similar to (but more complicated than) the 2-D case?

Later on we discuss vector functions (physicists call them 'fields'), which assign to each point in space a vector. We also learn how to deal with curves and surfaces.

Matrices

such as, for example \[
\begin{bmatrix}
4 & -6 & 0 \\
-3 & 2 & 1 \\
9 & -5 & -3
\end{bmatrix}
\] which is a \( 3 \times 3 \) \((n \times m \text{ in general})\) array of numbers (these are called the matrix elements).

Understand the following definitions:

- SQUARE matrix \((n = m)\), UNIT matrix \( \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \), ZERO matrix;
- and operations:
  - matrix TRANSPOSE \( \begin{bmatrix} 4 & -3 & 9 \\ -6 & 2 & -5 \\ 0 & 1 & -3 \end{bmatrix} \),
  - matrix ADDITION (for same-size matrices),
  - MULTIPLICATION \( \begin{bmatrix} 3 & -2 & 0 \\ 5 & -3 & 1 \end{bmatrix} \begin{bmatrix} 2 & -1 \\ 0 & 3 \\ -4 & 5 \end{bmatrix} = \begin{bmatrix} 6 & -9 \\ 6 & -9 \end{bmatrix} \),
  - INVERSE \( \begin{bmatrix} 4 & -6 & 0 \\ -3 & 2 & 1 \\ 9 & -5 & -3 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{7} & \frac{9}{7} & \frac{3}{7} \\ 0 & 3 & 1 \\ \frac{3}{7} & \frac{17}{7} & \frac{5}{7} \end{bmatrix} \),
- and DETERMINANT \( \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \).

For two square matrices, their product is not necessarily commutative, e.g.
\[
\begin{bmatrix} 2 & 0 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} 3 & 5 \\ 1 & -2 \end{bmatrix} = \begin{bmatrix} 6 & 10 \\ -5 & -23 \end{bmatrix} , \quad \begin{bmatrix} 3 & 5 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} -9 & 20 \\ 8 & -8 \end{bmatrix} .
\]

Notation:

\( \mathbb{I} \) stands for the unit matrix, \( \mathbb{O} \) for the zero matrix, \( A^T \) for transpose, \( A^{-1} \) for inverse, \( |A| \) for the determinant, \( AB \) for multiplication (careful with the order); \( A_{23} \) is the second-row, third-column element of \( A \).

A few basic rules:

\( A\mathbb{I} = \mathbb{I}A = A, \quad AO = OA = \mathbb{O}, \quad AA^{-1} = A^{-1}A = \mathbb{I}, \quad (AB)^T = B^TA^T, \quad \text{and } (AB)^{-1} = B^{-1}A^{-1} \), whenever the dimensions allow the operation.

Let us formally prove the second last equality:

\[
\begin{align*}
\left\{ (AB)^T \right\}_{ij} &= \{ \sum_k A_{ik}B_{kj} \}^T = \sum_k A_{jk}B_{ki} = \sum_k B_{ki}A_{jk} = \sum_k \{ B^T \}_{ik} \{ A^T \}_{kj} \\
&= \{ B^TA \}^T_{ij} \quad \text{for each } i \text{ and } j.
\end{align*}
\]
Complex numbers
Understand the basic algebra of adding and subtracting (trivial), multiplying $(3 - 2i)(4 + i) = 12 + 2 - 8i + 3i = 14 - 5i$ (distributive law plus $i^2 = -1$) and dividing:

$$\frac{3 - 2i}{4 + i} = \frac{(3 - 2i)(4 - i)}{(4 + i)(4 - i)} = \frac{10 - 11i}{17} = \frac{10}{17} - \frac{11}{17}i$$

**Notation:**
Complex conjugate of a number $z = x + yi$ is $\overline{z} = x - yi$ (change the sign of $i$).

**Magnitude** (absolute value): $|z| = \sqrt{x^2 + y^2}$.
**Real** and **imaginary** parts are $\text{Re}(z) = x$ and $\text{Im}(z) = y$, respectively.

**Polar representation**
of $z = x + yi = re^{i\theta} = r\cos(\theta) + i r\sin(\theta)$, where $r = |z|$ and $\theta = \arctan_2(y, x)$ which is called the argument of $z$ and is chosen so that $\theta \in [0, 2\pi)$.

Here, we have used

$$e^{i\theta} = \cos(\theta) + i\sin(\theta)$$

which follows from Maclaurin expansion (try the proof).

When multiplying two complex numbers in polar representation, their magnitudes multiply as well, but their 'arguments' only add. When raising a complex number to an integer $n$, the resulting argument (angle) is simply $n\theta$.

This is quite useful when taking a large integer power of a complex number, for example: $(3 + 2i)^{31} = (9 + 4)^{\frac{31}{2}} \left[ \cos \left( 31 \arctan \left( \frac{2}{3} \right) \right) + i \sin \left( 31 \arctan \left( \frac{2}{3} \right) \right) \right] = 1.5005 \times 10^{17} - 1.0745 \times 10^{17}i$ (give at least four significant digits in all your answers).
Part I

ORDINARY DIFFERENTIAL EQUATIONS
General review of differential equations

► (Single) Ordinary Differential Equation ◄

There is only one independent variable, usually called \( x \), and one dependent variable (a function of \( x \)), usually called \( y(x) \). The equation involves \( x, y(x), y'(x), \) and possibly higher derivatives of \( y(x) \), for example: \( y''(x) = \frac{1 + y(x) \cdot y'(x)}{1 + x^2} \). We usually leave out the argument of \( y \) to simplify the notation, thus:

\[
y'' = \frac{1 + yy'}{1 + x^2}
\]

The highest derivative is the order of the equation (this one is of the second order).

Our task is to find the (unknown) function \( y(x) \) which meets the equation. The solution is normally a family of functions, with as many extra parameters as the order of the equation (these are usually called \( C_1, C_2, C_3, \ldots \)).

In the next chapter, we study first-order ODE. Then, we move on to higher-order ODE, but, for these, we restrict our attention almost entirely to the linear [in \( y \) and its derivatives, e.g. \( y'' + \sin x \cdot y' - (1 + x^2) \cdot y = e^{-x} \)] case with constant coefficients [example: \( y'' - 2y' + 3y = e^{2x} \)]. When the right hand side of such an equation is zero, the equation is called homogenous.

► A Set (or System) of ODEs ◄

Here, we have several dependent (unknown) functions \( y_1, y_2, y_3, \ldots \) of still a single independent variable \( x \) (sometimes called \( t \)). We will study only a special case of these systems, when the equations (and we need as many equations as there are unknown functions) are of the first order, linear in \( y_1, y_2, y_3, \ldots \) and its derivative, and having constant coefficients.

Example:

\[
\begin{align*}
\dot{y}_1 &= 2y_1 - 3y_2 + t^2 \\
\dot{y}_2 &= y_1 + 5y_2 + \sin t
\end{align*}
\]

(when the independent variable is \( t, \dot{y}_1, \dot{y}_2, \ldots \) is a common notation for the first derivatives).

► Partial Differential Equations (PDE) ◄

They usually have a single dependent variable and several independent variables. The derivatives are then automatically of the partial type. It is unlikely that we will have time to discuss even a brief introduction to these. But you will study them in Physics, where you also encounter systems of PDEs (several dependent and independent variables), to be solved by all sorts of ingenious techniques. They are normally tackled on an individual basis only.
Chapter 2  FIRST-ORDER DIFFERENTIAL EQUATIONS

The most general form of such an equation (assuming we can solve it for $y'$, which is usually the case) is

$$y' = f(x, y)$$

where $f(x, y)$ is any expression involving both $x$ and $y$. We know how to solve this equation (analytically) in only a few special cases (to be discussed shortly).

Graphically, the situation is easier: using an $x$–$y$ set of coordinates, we can draw (at as many points as possible) the slope $\equiv f(x, y)$ of a solution passing through the corresponding $(x, y)$ point and then, by attempting to connect these, visualize the family of all solutions. But this is not very accurate nor practical.

Usually, there is a whole family of solutions which covers the whole $x$–$y$ plane, by curves which don’t intersect. This means that exactly one solution passes through each point. Or, equivalently: given an extra condition imposed on the solution, namely $y(x_0) = y_0$ where $x_0$ and $y_0$ are two specific numbers (the so called initial condition), this singles out a unique solution. But sometimes it happens that no solution can be found for some initial conditions, and more than one (even infinitely many) solutions for others.

We will look at some of these issues in more detail later on, let us now go over the special cases of the first-order ODE which we know how to solve:

'Trivial' equation
This case is so simple to solve that it is usually not even mentioned as such, but we want to be systematic. The equation has the form:

$$y' = f(x)$$

i.e. $y'$ is expressed as a function of $x$ only.

It is quite obvious that the general solution is

$$y(x) = \int f(x)dx + C$$

(graphically, it is the same curve slid vertically up and down). Note that even in this simplest case we cannot always find an analytical solution (we don’t know how to integrate all functions).

EXAMPLE: $y' = \sin(x)$.

Solution: $y(x) = -\cos(x) + C$. 

Separable equation

Its general form is:

\[ y' = h(x) \cdot g(y) \]

(a product of a function of \( x \) times a function of \( y \)).

Solution: Writing \( y' \) as \( \frac{dy}{dx} = h(x) \cdot g(y) \) we can achieve the actual separation (of \( x \) from \( y \)), thus:

\[ \frac{dy}{g(y)} = h(x)dx \]

where the left and the right hand sides can be individually integrated (in terms of \( y \) and \( x \), respectively), a constant \( C \) added (to the right hand side only, why is this sufficient?), and the resulting equation solved for \( y \) whenever possible (to get the so called explicit solution). If the equation cannot be solved for \( y \), we leave it in the so called implicit form.

EXAMPLES:

1. \( y' = x \cdot y \Rightarrow \frac{dy}{y} = x \, dx \Rightarrow \ln |y| = \frac{x^2}{2} + \tilde{C} \Rightarrow y = \pm e^{\frac{x^2}{2}} \cdot e^{\tilde{C}} \equiv Ce^{\frac{x^2}{2}} \) (by definition of a new \( C \), which may be positive or negative). Let us plot these with \( C = -3, -2, -1, 0, 1, 2, 3 \) to visualize the whole family of solutions:

   (essentially the same curve, expanded or compressed along the \( y \)-direction).

2. \( 9yy' + 4x = 0 \Rightarrow 9y \, dy = -4x \, dx \Rightarrow 9\frac{y^2}{2} = -4\frac{x^2}{2} + \tilde{C} \Rightarrow y^2 + \frac{4}{9}x^2 = C \) (by intelligently redefining the constant). We will leave the answer in its explicit form, which clearly shows that the family of solutions are ellipses centered on the origin, with the vertical versus horizontal diameter in the 2 : 3 ratio.

3. \( y' = -2xy \Rightarrow \frac{dy}{y} = -2x \, dx \Rightarrow \ln |y| = -x^2 + \tilde{C} \Rightarrow y = Ce^{-x^2} \) (analogous to Example 1).

4. \( (1 + x^2)y' + 1 + y^2 = 0 \), with \( y(0) = 1 \) (initial value problem).

   Solution: \( \frac{dy}{1 + y^2} = -\frac{dx}{1 + x^2} \Rightarrow \arctan(y) = -\arctan(x) + \tilde{C} \equiv \arctan(C) - \arctan(x) \Rightarrow y = \tan(\arctan(C) - \arctan(x)) = \frac{C-x}{1+C \cdot x} \) [recall the \( \tan(\alpha - \beta) \) formula]. To find \( C \) we solve \( 1 = \frac{C-0}{1+C \cdot 0} \Rightarrow C = 1 \).

   Answer: \( y(x) = \frac{1-x}{1+x} \).

   Check: \( (1 + x^2)\frac{d}{dx} \left( \frac{1-x}{1+x} \right) + 1 + \left( \frac{1-x}{1+x} \right)^2 = 0 \checkmark \).
Scale-independent equation looks as follows:
\[ y' = g \left( \frac{y}{x} \right) \]
(note that the right hand side does not change when replacing \( x \to ax \) and, simultaneously, \( y \to ay \), since \( a \) will cancel out).

**Solve** by introducing a new dependent variable \( u(x) = \frac{y(x)}{x} \) or, equivalently, \( y(x) = x \cdot u(x) \). This implies \( y' = u + xu' \); substituted into the original equation yields:
\[ xu' = g(u) - u \]
which is separable in \( x \) and \( u \):
\[ \frac{du}{g(u) - u} = \frac{dx}{x} \]
Solve the separable equation for \( u(x) \), and convert to \( y(x) = xu(x) \).

**EXAMPLES:**

1. \( 2xyy' - y^2 + x^2 = 0 \) \( \Rightarrow \) \( y' = \frac{y/x}{2} - \frac{1}{2y/x} \) \( \Rightarrow \) \( xu' = -\frac{u^2 + 1}{2u} \) \( \Rightarrow \) \( 2u \frac{du}{u^2 + 1} = \frac{dx}{x} \) \( \Rightarrow \) \( \ln(1 + u^2) = -\ln|x| + \tilde{C} \) \( \Rightarrow u^2 + 1 = \frac{2\tilde{C}}{x} \) (factor of 2 is introduced for future convenience) \( \Rightarrow \) \( y^2 + x^2 - 2Cx = 0 \) \( \Rightarrow y^2 + (x - \tilde{C})^2 = \tilde{C}^2 \) (let us leave it in the implicit form). This is a family of circles having a center at any point of the \( x \)-axis, and being tangent to the \( y \)-axis.

2. \( x^2y' = y^2 + xy + x^2 \) \( \Rightarrow \) \( y' = \left( \frac{y}{x} \right)^2 + \frac{y}{x} + 1 \) \( \Rightarrow \) \( xu' = u^2 + 1 \) \( \Rightarrow \) \( \frac{du}{1 + u^2} = \frac{dx}{x} \) \( \Rightarrow \) \( \arctan(u) = \ln|x| + \tilde{C} \) \( \Rightarrow u = \tan(\ln|x| + \tilde{C}) \) \( \Rightarrow y = x \cdot \tan(\ln|x| + \tilde{C}) \) ▫

**Modified Scale-Independent**
\[ y' = \frac{y}{x} + g \left( \frac{y}{x} \right) \cdot h(x) \]
The same substitution gives
\[ xu' = g(u) \cdot h(x) \]
which is also separable.

The main point is to be able to recognize that the equation is of this type.

**EXAMPLE:**
\[ y' = \frac{y}{x} + \frac{2x^3 \cos(x^2)}{y} \] \( \Rightarrow \) \( xu' = \frac{2x^2 \cos(x^2)}{u} \) \( \Rightarrow \) \( u \ du = 2x \cos(x^2) \ dx \) \( \Rightarrow \) \( u^2 = \frac{\sin(x^2) + \tilde{C}}{2} \) \( \Rightarrow u = \pm \sqrt{2 \sin(x^2) + \tilde{C}} \) \( \Rightarrow y = \pm x \sqrt{2 \sin(x^2) + \tilde{C}} \) ▫

**Any Other Smart Substitution**
(usually suggested), which makes the equation separable.

**EXAMPLES:**

1. \((2x - 4y + 5)y' + x - 2y + 3 = 0\) [suggestion: introduce: \(v(x) = x - 2y(x)\),
i.e. \(y = \frac{x - v}{2}\) and \(y' = \frac{1 - v'}{2}\)\] \(\Rightarrow\) \((2v + 5)\frac{1 - v'}{2} + v + 3 = 0\ \Rightarrow\ -(v + \frac{5}{2})v' + 2v + \frac{11}{2} = 0\ \Rightarrow\ v\frac{dv}{v + \frac{11}{4}} = 2dx \Rightarrow\ v - \frac{1}{2}\ln|v + \frac{11}{4}| = 2x + C \Rightarrow\ x - 2y - \frac{1}{4}\ln|x - 2y + \frac{11}{4}| = 2x + C.\ We have to leave the solution in the implicit form because we cannot solve for \(y\), except numerically – it would be a painstaking procedure to draw even a simple graph now).

2. \(y' \cos(y) + x \sin(y) = 2x\), seems to suggest \(\sin(y) \equiv v(x)\) as the new dependent variable, since \(v' = y' \cos(y)\) [by chain rule]. The new equation is thus simply: \(v' + xv = 2x\), which is linear (see the next section), and can be solved as such: \(\frac{dv}{v} = -x dx \Rightarrow \ln|v| = -\frac{x^2}{2} + \tilde{c} \Rightarrow v = ce^{-\frac{x^2}{2}}\), substitute: \(c' e^{-\frac{x^2}{2}} - xce^{-\frac{x^2}{2}} + xce^{-\frac{x^2}{2}} = 2x \Rightarrow c' = 2x e^{\frac{x^2}{2}} \Rightarrow c(x) = 2e^{\frac{x^2}{2}} + C \Rightarrow v(x) = 2 + Ce^{\frac{x^2}{2}} \Rightarrow y(x) = \arcsin\left(2 + Ce^{\frac{x^2}{2}}\right)\]

**Linear equation**
has the **form** of:

\[y' + g(x) \cdot y = r(x)\]

[both \(g(x)\) and \(r(x)\) are arbitrary – but specific – functions of \(x\)].

The solution is constructed in two stages, by the so called

► **Variation-of-Parameters Technique**<br>
which works as follows:

1. Solve the **homogeneous equation** \(y' = -g(x) \cdot y\), which is separable, thus:

\[y_h(x) = c \cdot e{-\int g(x)dx}\]

2. Assume that \(c\) itself is a **function** of \(x\), substitute \(c(x) \cdot e{-\int g(x)dx}\) back into the full equation, and solve the resulting [trivial] differential equation for \(c(x)\).

**EXAMPLES:**

1. \(y' + \frac{y}{x} = \frac{\sin x}{x}\)

Solve \(y' + \frac{y}{x} = 0 \Rightarrow \frac{dy}{y} = -\frac{dx}{x} \Rightarrow \ln|y| = -\ln |x| + \tilde{c} \Rightarrow y = \frac{c}{x}\).

Now substitute this to the original equation: \(\frac{c'}{x} - \frac{c}{x^2} + \frac{c}{x^2} = \frac{\sin x}{x} \Rightarrow c' = \sin x \Rightarrow c(x) = -\cos x + C\) (the big \(C\) being a true constant) \(\Rightarrow y(x) =\)
\[- \frac{\cos x}{x} + \frac{C}{x} \]. The solution has always the form of \( y_p(x) + Cy_h(x) \), where \( y_p(x) \) is a **PARTICULAR** solution to the full equation, and \( y_h(x) \) solves the homogeneous equation only.

Let us verify the former: \( \frac{d}{dx} \left( - \frac{\cos x}{x} \right) - \frac{\cos x}{x^2} = \frac{\sin x}{x} \) (check).

2. \( y' - y = e^{2x} \).

First \( y' - y = 0 \Rightarrow \frac{dy}{y} = dx \Rightarrow y = ce^x \).

Substitute: \( c'e^x + ce^x - ce^x = e^{2x} \Rightarrow c' = 2e^{x} \Rightarrow c(x) = e^x + C \Rightarrow y(x) = e^{2x} + Ce^x \).

3. \( xy' + y = 0 \)

Homogeneous: \( \frac{dy}{y} = -\frac{dx}{x} \Rightarrow \ln |y| = -\ln |x| + c \Rightarrow y = \frac{c}{x} \).

Substitute: \( c' - \frac{c}{x} + \frac{c}{x} = -4 \Rightarrow c(x) = -4x + C \Rightarrow y(x) = -4 + \frac{C}{x} \).

4. \( y' + y \cdot \tan(x) = \sin(2x), y(0) = 1 \).

Homogeneous: \( \frac{dy}{y} = -\frac{\sin x \ dx}{\cos x} \Rightarrow \ln |y| = \ln |\cos x| + \tilde{c} \Rightarrow y = c \cdot \cos x \).

Substitute: \( c' \cos x - c \sin x + c \sin x = 2 \sin x \cos x \Rightarrow c' = 2 \sin x \Rightarrow c(x) = -2 \cos x + C \Rightarrow y(x) = -2 \cos^2 x + C \cos x \) [\( \cos^2 x \) is the usual 'shorthand' for \( (\cos x)^2 \)].

To find the value of \( C \), solve: \( 1 = -2 + C \Rightarrow C = 3 \).

The final answer is thus: \( y(x) = -2 \cos^2 x + 3 \cos x \).

To verify: \( \frac{d}{dx} [-2 \cos^2 x + 3 \cos x] + [-2 \cos^2 x + 3 \cos x] \cdot \frac{\sin x}{\cos x} = 2 \cos x \sin x \) (check).

5. \( x^2 y' + 2xy - x + 1 = 0, y(1) = 0 \)

Homogeneous [realize that here \( -x + 1 \) is the non-homogeneous part]: \( \frac{dy}{y} = -2\frac{dx}{x} \Rightarrow \ln |y| = -2 \ln |x| + \tilde{C} \Rightarrow y = \frac{c}{x^2} \).

Substitute: \( c' - \frac{2c}{x^3} + \frac{2c}{x^3} - x + 1 = 0 \Rightarrow c' = x - 1 \Rightarrow c = \frac{x^2}{2} - x + C \Rightarrow y = \frac{1}{2} - \frac{1}{x} + \frac{C}{x^2} \).

To meet the initial-value condition: \( 0 = \frac{1}{2} - 1 + C \Rightarrow C = \frac{1}{2} \).

Final answer: \( y = \frac{(1 - x)^2}{2x^2} \).

Verify: \( x^2 \frac{d}{dx} \left( \frac{(1 - x)^2}{2x^2} \right) + 2x \left( \frac{(1 - x)^2}{2x^2} \right) - x + 1 = 0 \) √.

6. \( y' - \frac{2y}{x} = x^2 \cos(3x) \)

First: \( \frac{dy}{y} = 2\frac{dx}{x} \Rightarrow \ln |y| = 2 \ln |x| + \tilde{c} \Rightarrow y = cx^2 \).
Substitute: \( c'x^2 + 2cx - 2cx = x^2 \cos(3x) \Rightarrow c' = \cos(3x) \Rightarrow c = \frac{\sin(3x)}{3} + C \Rightarrow y = \frac{x^2}{3} \sin(3x) + Cx^2 \)

To verify the particular solution: \( \frac{d}{dx} \left( \frac{x^2}{3} \sin(3x) \right) - \frac{2x}{3} \sin(3x) = x^2 \cos(3x) \checkmark \)

**Bernoulli equation**

\[ y' + f(x) \cdot y = r(x) \cdot y^a \]

where \( a \) is a specific (constant) exponent.

Introducing a new dependent variable \( u = y^{1-a} \), i.e. \( y = u^{1-a} \), one gets:

\[ \frac{1}{1-a} u \cdot \frac{1}{1-a} u' = \frac{1}{1-a} u' \text{ [chain rule]} + f(x) \cdot u^{1-a} = r(x) \cdot u^{1-a}. \]

Multiplying by \((1-a)u^{-\frac{1}{1-a}}\) results in:

\[ u' + (1-a)f(x) \cdot u = (1-a)r(x) \]

which is linear in \( u' \) and \( u \) (i.e., of the previous type), and solved as such.

The answer is then easily converted back to \( y = u^{1-a} \).

**EXAMPLES:**

1. \( y' + xy = \frac{x}{y} \) (Bernoulli, \( a = -1 \), \( f(x) \equiv x \), \( g(x) \equiv x \) \( \Rightarrow u' + 2xu = 2x \) where \( y = u^{\frac{1}{2}} \).

   Solving as linear: \( \frac{du}{u} = -2x \, dx \Rightarrow \ln |u| = -x^2 + c \Rightarrow u = c \cdot e^{-x^2} \)

   Substitute: \( c' e^{-x^2} - 2xce^{-x^2} + 2x e^{-x^2} = 2x \Rightarrow c' = 2x e^{x^2} \Rightarrow c(x) = e^{x^2} + C \Rightarrow u(x) = 1 + C e^{-x^2} \Rightarrow y(x) = \pm \sqrt{1 + C e^{-x^2}} \) (one can easily check that this is a solution with either the + or the − sign).

2. \( 2xy' = 10x^3 y^5 + y \) (terms reshuffled a bit). Bernoulli with \( a = 5 \), \( f(x) = -\frac{1}{2x} \), and \( g(x) = 5x^2 \).

   This implies \( u' + \frac{2}{x} u = -20x^2 \) with \( y = u^{-\frac{1}{4}} \).

   Solving as linear: \( \frac{du}{u} = -2 \frac{dx}{x} \Rightarrow \ln |u| = -2 \ln |x| + c \Rightarrow u = \frac{c}{x^2} \)

   Substituted back into the full equation: \( \frac{d}{dx} \left( \frac{c}{x^2} \right) - 2 \frac{c}{x^2} + 2 \frac{c}{x^2} = -20x^2 \Rightarrow c' = -20x^4 \Rightarrow c(x) = -4x^5 + C \Rightarrow u(x) = -4x^3 + \frac{C}{x^2} \Rightarrow y(x) = \pm \sqrt{\left(-4x^3 + \frac{C}{x^2}\right)^{-\frac{1}{4}}}. \)

3. \( 2xyy' + (x-1)y^2 = x^2e^x \), Bernoulli with \( a = -1 \), \( f(x) = \frac{x-1}{x} \), and \( g(x) = \frac{x^2}{2} e^x \).

   This translates to: \( u' + \frac{x-1}{x} u = x e^x \) with \( y = u^{\frac{1}{2}} \).

   Solving homogeneous part: \( \frac{du}{u} = (\frac{1}{x} - 1) \, dx \Rightarrow \ln |u| = \ln |x| - x + c \Rightarrow u = c x e^{-x} \)

   Substituted: \( c' x e^{-x} + c e^{-x} - c x e^{-x} + (x-1) c e^{-x} = x e^x \Rightarrow c' = c^2 e^x \Rightarrow c(x) = \frac{1}{2} e^{2x} + C \Rightarrow u(x) = \frac{x}{2} e^x + C x e^{-x} \Rightarrow y(x) = \pm \sqrt{\frac{x}{2} e^x + C x e^{-x}} \).
**Exact equation**

First we have to explain the **general idea** behind this type of equation:

Suppose we have a function of $x$ and $y$, $f(x,y)$ say. Then $\frac{\partial f}{\partial x} \, dx + \frac{\partial f}{\partial y} \, dy$ is the so called **total differential** of this function, corresponding to the function’s increase when its arguments change from $(x,y)$ to $(x + dx, y + dy)$. By setting this quantity equal to zero, we are effectively demanding that the function does not change its value, i.e. $f(x,y) = C$ (constant). The last equation is then an [implicit] solution to

$$\frac{\partial f}{\partial x} \, dx + \frac{\partial f}{\partial y} \, dy = 0$$

(the corresponding exact equation).

**EXAMPLE**: Suppose $f(x,y) = x^2 y - 2x$. This means that $(2xy - 2) \, dx + x^2 \, dy = 0$ has a simple solution $x^2 y - 2x = C \Rightarrow y = \frac{2}{x} + C \frac{x}{x^2}$. Note that the differential equation can also be re-written as: $y' = \frac{2}{x^2} \frac{1}{x} - \frac{x}{x^2}$, $x^2 y' + 2xy = 2$, etc. [coincidentally, this equation is also linear; we can thus double-check the answer].

We must now try to reverse the process, since in the actual situation we will be given the differential equation and asked to find the corresponding $f(x,y)$.

There are then two issues to be settled:

1. How do we verify that the equation is exact?
2. Knowing it is, how do we solve it?

To answer the first question, we recall that

$$\frac{\partial^2 f}{\partial x \partial y} \equiv \frac{\partial^2 f}{\partial y \partial x}$$

Thus, $g(x,y) \, dx + h(x,y) \, dy = 0$ is exact if and only if

$$\frac{\partial g}{\partial y} \equiv \frac{\partial h}{\partial x}$$

As to solving the equation, we proceed in three stages:

1. Find $G(x,y) = \int g(x,y) \, dx$ (considering $y$ a constant).

2. Construct $H(y) = h(x,y) - \frac{\partial G}{\partial y}$ [must be a function of $y$ only, as $\frac{\partial h}{\partial x} = \frac{\partial}{\partial x} \frac{\partial G}{\partial y} - \frac{\partial}{\partial y} \frac{\partial G}{\partial y} \equiv 0$].

3. $f(x,y) = G(x,y) + \int H(y) \, dy$

[Proof: $\frac{\partial f}{\partial x} = \frac{\partial G}{\partial x} = g$ and $\frac{\partial f}{\partial y} = \frac{\partial G}{\partial y} + H = h \Box$].

Even though this looks complicated, one must realize that the individual steps are rather trivial, and exact equations are therefore easy to solve.
EXAMPLE: \(2x \sin(3y) \, dx + (3x^2 \cos(3y) + 2y) \, dy = 0\).

Let us first verify that the equation is exact: \(\frac{\partial}{\partial y} 2x \sin(3y) = 6x \cos 3y,\)
\(\frac{\partial}{\partial x} (3x^2 \cos(3y) + 2y) = 6x \cos 3y \) (check)

Solving it: \(G = x^2 \sin(3y), H = 3x^2 \cos(3y) + 2y - 3x^2 \cos(3y) = 2y, f(x, y) = x^2 \sin(3y) + y^2.\)

Answer: \(y^2 + x^2 \sin(3y) = C\) (implicit form)  

 Integrating Factors  

Any first-order ODE (e.g. \(y' = \frac{y}{x}\)) can be expanded in a form which makes it look like an exact equation, thus: \(\frac{dy}{dx} = \frac{y}{x} \Rightarrow y \, dx - x \, dy = 0\). But since \(\frac{\partial y}{\partial y} \neq -\frac{\partial x}{\partial x}\), this equation is not exact.

The good news is that there is always a function of \(x\) and \(y\), say \(F(x, y)\), which can multiply any such equation (a legitimate modification) to make it exact. This function is called an integrating factor.

The bad news is that there is no general procedure for finding \(F(x, y)\) [if there were, we would know how to solve all first-order differential equations – too good to be true].

Yet, there are two special cases when it is possible:
Let us write the differential equation in its 'look-like-exact' form of

\[P(x, y) \, dx + Q(x, y) \, dy = 0\]

where \(\frac{\partial P}{\partial y} \neq \frac{\partial Q}{\partial x}\) (thus the equation is not exact yet). One can find an integrating factor from

1.

\[
\frac{d \ln F}{dx} = \frac{\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}}{Q}
\]

iff the right hand side of this equation is a function of \(x\) only

Proof: \(FP \, dx + FQ \, dy = 0\) is exact when \(\frac{\partial (FP)}{\partial y} = \frac{\partial (FQ)}{\partial x} \Rightarrow F \frac{\partial P}{\partial y} = \frac{dF}{dx} \cdot Q + F \frac{\partial Q}{\partial x}\)
assuming that \(F\) is a function of \(x\) only. Solving for \(\frac{dF}{dx}\) results in \(\frac{\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x}}{Q}\).
When the last expression contains no \(y\), we simply integrate it (with respect to \(x\)) to find \(\ln F\). Otherwise (when \(y\) does not cancel out of the expression), the formula is meaningless. \(\square\)

2. or from

\[
\frac{d \ln F}{dy} = \frac{\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}}{P}
\]

iff the right hand side is a function of \(y\) only.

EXAMPLES:
1. Let us try solving our $y\,dx - x\,dy = 0$. Since $\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} = -\frac{2}{x}$, we have $\ln F = -2 \int \frac{dx}{x} = -2 \ln x$ (no need to bother with a constant) $\Rightarrow F = \frac{1}{x^2}$. Thus $\frac{1}{x^2} \, dx - \frac{1}{x} \, dy = 0$ must be exact (check it). Solving it gives $-\frac{y}{x} = C$, or $y = Cx$.

The original equation is, coincidentally, also separable (sometimes it happens that an equation can be solved in more than one way), so we can easily verify that the answer is correct.

But this is not the end of this example yet! We can also get: $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = -\frac{2}{y}$, which implies that $\ln F = -2 \int \frac{dy}{y} = -2 \ln y \Rightarrow F = \frac{1}{y}$. Is this also an integrating factor?

The answer is yes, there are infinitely many of these; one can multiply an integrating factor (such as $\frac{1}{x}$) by any function of what we know must be a constant ($-\frac{y}{x}$ in our case, i.e. the left hand side of our solution). Since $\frac{1}{y^2} = \frac{1}{x^2} \cdot \left(-\frac{y}{x}\right)^2$, this is also an integrating factor of our equation. One can verify that, using this second integration factor, one still obtains the same simple $y = Cx$ solution.

To formalize our observation: When $g \, dx + h \, dy = 0$ is exact (i.e. $g = \frac{\partial f}{\partial x}$ and $h = \frac{\partial f}{\partial y}$), so is $R(f) \, g \, dx + R(f) \, h \, dy = 0$ where $R$ is any function of $f$.

Proof: $\frac{\partial (Rg)}{\partial y} = \frac{dR}{df} \cdot \frac{\partial f}{\partial y} \cdot g + R \cdot \frac{\partial R}{\partial y} = \frac{dR}{df} \cdot h \cdot g + R \cdot \frac{\partial R}{\partial y}$. Similarly $\frac{\partial (Rh)}{\partial x} = \frac{dR}{df} \cdot \frac{\partial f}{\partial x} \cdot h + R \cdot \frac{\partial R}{\partial x} = \frac{dR}{df} \cdot g \cdot h + R \cdot \frac{\partial R}{\partial x}$. Since $\frac{\partial R}{\partial y} \equiv \frac{\partial R}{\partial x}$, the two expressions are identical. □

2. $(2 \cos y + 4x^2) \, dx = x \sin y \, dy$ [i.e. $Q = -x \sin y$].

Since $\frac{\partial P}{\partial x} - \frac{\partial Q}{\partial y} = \frac{-2 \sin y + \sin y}{x \sin y} = -\frac{x}{x}$ we get $\ln F = \int \frac{1}{x} \, dx = \ln x \Rightarrow F = x$.

$(2x \cos y + 4x^3) \, dx - x^2 \sin y \, dy = 0$

is therefore exact, and can be solved as such: $x^2 \cos y + x^4 = C \Rightarrow y = \arccos \left(\frac{C}{x^2 - x^2}\right)$.

3. $(3xe^y + 2y) \, dx + (x^2e^y + x) \, dy = 0$.

Trying again $\frac{\partial P}{\partial y} - \frac{\partial Q}{\partial x} = \frac{xe^y + 1}{x^2e^y + x} = \frac{1}{x}$, which means that $\ln F = \int \frac{dx}{x} = \ln x \Rightarrow F = x$.

$(3x^2e^y + 2xy) \, dx + (x^3e^y + x^2) \, dy = 0$

is exact. Solving it gives: $x^3e^y + x^2y = C$ [implicit form]. □

More ‘exotic’ equations

and methods of solving them:

There are many other types of first-order ODEs which can be solved by all sorts of ingenious techniques (stressing that our list of ‘solvable’ equations has been far from complete). We will mention only one, for illustration:
**Clairaut equation**:

\[ y = xy' + g(y') \]

where \( g \) is an arbitrary function. The idea is to introduce \( p(x) \equiv y'(x) \) as an unknown function, differentiate the original equation with respect to \( x \), obtaining

\[ p = p + xp' + p'g'(p) \Rightarrow p' \cdot (x + g'(p)) = 0. \]

This implies that either \( p \equiv y' = C \Rightarrow y = xC + g(C) \)

which represents a family of **regular** solutions (all straight lines), or

\[ x = -g'(p) \]

which, when solved for \( p \) and substituted back into \( y = xp + g(p) \) provides the so called **singular** solution (an envelope of the regular family).

**EXAMPLE**:

\[(y')^2 - xy' + y = 0\] (terms reshuffled a bit) is solved by either \( y = Cx - C^2 \),

or \( x = 2p \Rightarrow p = \frac{x}{2} \Rightarrow y = xp - p^2 = \frac{x^2}{4} \) (singular solution). Let us display them graphically:

Note that for an initial condition below or at the parabola two possible solutions exist, above the parabola there is none.

This concludes our discussion of Clairaut equation.

And finally, a **useful trick** worth mentioning: When the (general) equation appears more complicated in terms of \( y \) rather than \( x \) [e.g. \( (2x + y^4)y' = y \)], one can try reversing the role of \( x \) and \( y \) (i.e. considering \( x \) as the **dependent** variable and \( y \) as the **independent** one). All it takes is to replace \( y' \equiv \frac{dy}{dx} \) by \( \frac{dx}{dy} \), for example (using the previous equation): \( \frac{dx}{dy} = 2\frac{x^2}{y} + y^3 \) (after some simplification). The last equation is linear (in \( x \) and \( \frac{dx}{dy} \)) and can be solved as such:

\[ \frac{dx}{y} = 2\frac{dy}{y} \Rightarrow \ln |x| = 2 \ln |y| + \tilde{c} \Rightarrow x(y) = y^2 \cdot c(y). \]

Substituted into the full equation: \( 2yc + y^2 \frac{dc}{dy} = 2\frac{x^2}{y} + y^3 \Rightarrow \frac{dc}{dy} = y \Rightarrow c(y) = \frac{y^2}{2} - C \) [the minus sign is more convenient here] \( \Rightarrow x = \frac{y^4}{4} - Cy^2. \)

This can now be solved for \( y \) in terms of \( x \), to get a solution to the original equation: \( y = \pm \sqrt{C \pm \sqrt{C^2 + 2x}}. \)
Applications

1. Find a curve such that (from each of its points) the distance to the origin is the same as the distance to the intersection of its normal (i.e. perpendicular straight line) with the \(x\)-axis.

Solution: Suppose \(y(x)\) is the equation of the curve (yet unknown). The equation of the normal is
\[
Y - y = -\frac{1}{y'} (X - x)
\]
where \((x, y)\) [fixed] are the points of the curve, and \((X, Y)\) [variable] are the points of the normal [which is a straight line passing through \((x, y)\), with its slope equal minus the reciprocal of the curve’s slope \(y'\)]. This normal intersects the \(x\)-axis at \(Y = 0\) and \(X = yy' + x\). The distance between this and the original \((x, y)\) is \(\sqrt{(yy')^2 + y^2}\), the distance from \((x, y)\) to \((0,0)\) is \(\sqrt{x^2 + y^2}\). These two distances are equal when \(y^2(y')^2 = x^2\), or \(y' = \pm \frac{x}{y}\). This is a separable differential equation easy to solve: \(y^2 \pm x^2 = C\). The curves are either circles centered on \((0,0)\) [yes, that checks, right?], or hyperbolas [with \(y = \pm x\) as special cases].

2. Find a curve whose normals (all) pass through the origin.

Solution (we can guess the answer, but let us do it properly): Into the same equation of the curve’s normal (see above), we substitute 0 for both \(X\) and \(Y\), since the straight line must pass through \((0,0)\). This gives: \(-y dy = x dx \Rightarrow x^2 + y^2 = C\) (circles centered on the origin – we knew that!).

3. A family of curves covering the whole \(x-y\) plane enables one to draw lines perpendicular to these curves. The collection of all such lines is yet another family of curves orthogonal (i.e. perpendicular) to the original family. If we can find the differential equation \(y' = f(x, y)\) having the original family of curves as its solution, we can find the corresponding orthogonal family by solving \(y' = -\frac{1}{f(x,y)}\). The next set of examples relates to this.

(a) The original family is described by \(x^2 + (y - C)^2 = C^2\) with \(C\) arbitrary (i.e. collection of circles tangent to the \(x\)-axis at the origin). To find the corresponding differential equation, we differentiate the original equation with respect to \(x\): \(2x + 2(y - C)y' = 0\), solve for \(y' = \frac{x}{C-y}\), and then eliminate \(C\) by solving the original equation for \(C\), thus: \(x^2 + y^2 - 2Cy = 0 \Rightarrow C = \frac{x^2+y^2}{2y}\), further implying \(y' = \frac{x}{x^2+y^2} - \frac{2y}{2y} = \frac{2y}{x^2+y^2}\). To find the orthogonal family, we solve \(y' = \frac{y^2-x^2}{2xy}\) [scale-independent equation solved earlier]. The answer is: \((x - C)^2 + y^2 = C^2\), i.e. collection of circles tangent to the \(y\)-axis at the origin.
(b) Let the original family be circles centered on the origin (it should be clear what the orthogonal family is, but again, let’s solve it anyhow): \(x^2 + y^2 = C^2\) describes the original family, \(2x + 2yy' = 0\) is the corresponding differential equation (equivalent to \(y' = -\frac{x}{y}\), this time there is no \(C\) to eliminate). The orthogonal family is the solution to \(yy' = \frac{y}{x} \Rightarrow \frac{dy}{y} = \frac{dx}{x} \Rightarrow y =Cx\) (all straight lines passing through the origin).

(c) Let the original family be described by \(y^2 = x + C\) (the \(y^2 = x\) parabola slid horizontally). The corresponding differential equation is \(2yy' = 1\), the ‘orthogonal’ equation: \(y' = -\frac{x}{2y}\). Answer: \(\ln |y| = -2x + \bar{C}\). or \(y = Ce^{-2x}\) (try to visualize the curves).

(d) Finally, let us start with \(y = Cx^2\) (all parabolas tangent to the \(x\)-axis at the origin). Differentiating: \(y' = 2Cx \Rightarrow \) [since \(C = \frac{y}{x}\)] \(y' = 2\frac{y}{x}\). The ‘orthogonal’ equation is \(y' = -\frac{x}{2y} \Rightarrow y^2 + \frac{x^2}{2} = C\) [collection of ellipses centered on the origin, with the \(x\)-diameter being \(\sqrt{2}\) times bigger than the \(y\)-diameter).

4. The position of four ships on the ocean is such that the ships form vertices of a square of length \(L\). At the same instant each ship fires a missile that directs its motion towards the missile on its right. Assuming that the four missiles fly horizontally and with the same constant speed, find the path of each.

Solution: Let us place the origin at the center of the original square. It should be obvious that when we find one of the four paths, the other three can be obtained just by rotating it by 90, 180 and 270 degrees. This is actually true for the missiles’ positions at any instant of time. Thus, if a missile is at \((x, y)\), the one to its right is at \((y, -x)\) \([x, y)\) rotated by 90°). If \(y(x)\) is the resulting path for the first missile, \(Y - y = y' \cdot (X - x)\) is the straight line of its immediate direction. This straight line must pass through \((y, -x)\) [that’s where the other missile is, at the moment]. This means that, when we substitute \(y\) and \(-x\) for \(X\) and \(Y\), respectively, the equation must hold: \(-x - y = y' \cdot (y - x)\). And this is the differential equation to solve (as scale independent): \(xu' + u(\rightarrow y' = \frac{x+u}{x-y}) = \frac{x+u}{x-y} \Rightarrow xu' = \frac{1+u}{1-u} \Rightarrow \frac{1-u}{1+u} du = \frac{dx}{x} \Rightarrow \arctan(u) - \frac{1}{2} \ln(1 + u^2) = \ln |x| + \bar{C} \Rightarrow e^{\frac{\arctan(u)}{\sqrt{1 + u^2}}} = Cx\). This solution becomes a lot easier to understand in polar coordinates \([\theta = \arctan\left(\frac{y}{x}\right)\), and \(r = \sqrt{x^2 + y^2}\), where it looks like this: \(r = \frac{e^\theta}{C}\) (a spiral).

To Physics
1. A cylindrical container of radius $r$.

Solution: First we have to establish the volume $V$ of the remaining water as a function of height. In this case we get simply $V(h(t)) = \pi r^2 h(t)$. Differentiating with respect to $t$ we get: $\frac{dV}{dt} = \pi r^2 \frac{dh}{dt}$. This in turn must be equal to $-a\sqrt{h(t)}$ since the rate at which the water is flowing out must be equal to the rate at which its volume is decreasing. Thus $\pi r^2 \frac{dh}{dt} = -a\sqrt{h}$, where $h \equiv \frac{dh}{dt}$. This is a simple (separable) differential equation for $h(t)$, which we solve by

\[ \frac{dh}{\sqrt{h}} = \frac{-a}{\pi r^2} \, dt \Rightarrow \frac{h^{1/2}}{2} = -\frac{at}{2\pi r^2} + C \Rightarrow \sqrt{h} = -\frac{at}{2\pi r^2} + \sqrt{h_0} \quad [h_0 \text{ is the initial height at time } t = 0], \]

or equivalently $t = \frac{2\pi r^2}{a} \left( \sqrt{h_0} - \sqrt{h} \right)$.

Subsidiary: What percentage of time is spent emptying the last 20% of the container? Solution: $t_1 = \frac{2\pi r^2}{a} \sqrt{h_0}$ is the time to fully empty the container (this follows from our previous solution with $h = 0$). $t_{0.8} = \frac{2\pi r^2}{a} \left( \sqrt{h_0} - \sqrt{0.8h_0} \right)$ is the time it takes to empty the first 80% of the container. The answer:

\[ \frac{t_1 - t_{0.8}}{t_1} = \sqrt{\frac{1}{5}} = 44.72\% . \]

2. A conical container with the top radius (at $h_0$) equal to $r$.

Solution: $V(h(t)) = \frac{1}{3} \pi h \left( \frac{r}{h_0} \right)^2$ [follows from $\int_0^h \mu \left( \frac{r}{h_0} \right)^2 dx$]. Note that one fifth of the full volume corresponds to $h = \left(\frac{1}{5}\right)^{1/3} h_0$ (i.e. 58.48% of the full height!), obtained by solving $\frac{1}{3} \pi \left( \frac{r}{h_0} \right)^2 h = \frac{1}{3} \pi r^2 h_0 / 5$ for $h$. Thus

\[ \frac{2\pi r^2}{3h_0^2} h^2 = -a\sqrt{h} \quad \text{is the (separable) equation to solve, as follows:} \quad \frac{2\pi r^2}{3h_0^2} h^2 \, dh = -\frac{3a h_0^2}{2\pi r^2} \, dt \Rightarrow h^{5/2} = -\frac{15a h_0^2}{4\pi r^2} t + h_0^{5/2} \quad \Rightarrow \quad t = \frac{4\pi r^2}{15a h_0^2} \left( h_0^{5/2} - h^{5/2} \right) . \]

This implies $t_1 = \frac{4\pi r^2 \sqrt{h_0}}{15a}$ and $t_{0.8} = \frac{4\pi r^2 \sqrt{0.8h_0}}{15a} \cdot \left[ 1 - \left(\frac{1}{5}\right)^{5/6} \right] \Rightarrow \frac{t_1 - t_{0.8}}{t_1} = \left(\frac{1}{5}\right)^{5/6} = 26.15\% .

3. A hemisphere of radius $R$ (this is the radius of its top rim, also equal to the water’s full height).

Solution: $V(h(t)) = \frac{1}{3} \pi h^2 (3R - h)$ [follows from $\int_0^h \pi \left[ R^2 - (R - x)^2 \right] dx$].

Making the right hand side equal to $\frac{2}{3} \pi R^3 / 5$ and solving for $h$ gives the height of the 20% (remaining) volume. This amounts to solving $z^3 - 3z^2 + \frac{2}{5} = 0$ (a cubic equation) for $z \equiv \frac{h}{R}$. We will discuss formulas for solving cubic and quartic equations in the next chapter, for the time being we extract the desired root by Newton’s technique: $z_{n+1} = z_n - \frac{z_n^3 - 3z_n^2 + \frac{2}{5}}{3z_n^2 - 6z_n}$ starting with $z_0 = 0.3 \Rightarrow z_1 = 0.402614 \Rightarrow z_2 = 0.391713 \Rightarrow z_3 = z_4 = \ldots = 0.391600 \Rightarrow h = 0.391600 R$. Finish as your assignment.
Chapter 3 SECOND ORDER DIFFERENTIAL EQUATIONS

These are substantially more difficult to solve than first-order ODEs, we will thus concentrate mainly on the simplest case of linear equations with constant coefficients. Only in the following introductory section we look at two special non-linear cases:

Reducible to first order
If, in a second-order equation (does not appear explicitly; only \(x, y\) and \(y''\) do), then \(y' \equiv z(x)\) can be considered the unknown function of the equation. In terms of \(z(x)\), the equation is of the first order only, and can be solved as such. Once we have the explicit expression for \(z(x)\), we need to integrate it with respect to \(x\) to get \(y(x)\).

The final solution will thus have two arbitrary constants, say \(C_1\) and \(C_2\) — this is the case of all second-order equations, in general. With two arbitrary constants we need two conditions to single out a unique solution. These are usually of two distinct types

1. INITIAL CONDITIONS: \(y(x_0) = a\) and \(y'(x_0) = b\) \([x_0\) is quite often 0], specifying a value and a slope of the function at a single point,

or

2. BOUNDARY CONDITIONS: \(y(x_1) = a\) and \(y(x_2) = b\), specifying a value each at two distinct points.

EXAMPLES:

1. \(y'' = y' \Rightarrow z' = z\) [separable] \(\Rightarrow \frac{dz}{z} = dx \Rightarrow \ln |z| = x + C_1 \Rightarrow z = C_1 e^x \Rightarrow y = C_1 e^x + C_2\). Let’s impose the following initial conditions: \(y(0) = 0\) and \(y'(0) = 1\). By substituting into the general solution we get: \(C_1 + C_2 = 0\) and \(C_1 = 1 \Rightarrow C_0 = -1 \Rightarrow y = e^x - 1\) as the final answer.

2. \(x y'' + y' = 0 \Rightarrow x z' + z = 0\) [separable] \(\Rightarrow \frac{dz}{z} = -\frac{dx}{x} \Rightarrow \ln |z| = \ln |x| + C_1 \Rightarrow z = \frac{C_1}{x} \Rightarrow y = C_1 \ln |x| + C_2\). Let us make this into a boundary-value problem:

\(y(1) = 1\) and \(y(3) = 0\) \(\Rightarrow C_2 = 1\) and \(C_1 \ln 3 + C_2 = 0 \Rightarrow C_1 = -\frac{1}{\ln 3} \Rightarrow y = 1 - \frac{\ln |x|}{\ln 3}\).

3. \(x y'' + 2 y' = 0 \Rightarrow x z' + 2z = 0\) [still separable] \(\Rightarrow \frac{dz}{z} = -2 \frac{dx}{x} \Rightarrow z = \frac{C_1}{x^2} \Rightarrow y = \frac{C_1}{x} + C_2\). Sometimes the two extra conditions can be of a more bizarre
type: \( y(2) = \frac{1}{2} \) and requiring that the solution intersects the \( y = x \) straight line at the right angle. Translated into our notation: \( y'(x_0) = -1 \) where \( x_0 \) is a solution to \( y(x) = x \), i.e. \( -\frac{C_1}{x_0^2} = -1 \) with \( \frac{C_1}{x_0} + C_2 = x_0 \). Adding the original \( \frac{C_1}{2} + C_2 = \frac{1}{2} \), we can solve for \( C_2 = 0 \), \( C_1 = 1 \) and \( x_0 = 1 \) (that is where our solution intersects \( y = x \)). The final answer: \( y(x) = \frac{1}{x} \).

The second type (of a second-order equation reducible to first order) has

\[ \text{\textbf{EXAMPLES:}} \]

1. \( y \cdot y' + (y')^2 = 0 \) \( \Rightarrow \) \( y \frac{dz}{dy} z + z^2 = 0 \) [separable] \( \Rightarrow \) \( \frac{dz}{z} = -\frac{dy}{y} \) \( \Rightarrow \) \( \ln |z| = -\ln |y| + \tilde{C}_1 \) \( \Rightarrow \) \( z = \frac{\tilde{C}_1}{y} \) \( \Rightarrow \) \( y' = \frac{\tilde{C}_1}{y} \) [separable again] \( \Rightarrow \) \( y \, dy = \tilde{C}_1 \Rightarrow y^2 = \tilde{C}_1 x + C_2 \).

2. \( y'' + e^{2y}(y')^3 = 0 \) \( \Rightarrow \) \( \frac{dz}{dy} z + e^{2y} z^3 = 0 \) \( \Rightarrow \) \( \frac{dz}{z^2} = -e^{2y} \, dy \) \( \Rightarrow \) \( -\frac{1}{z} = -\frac{1}{2} e^{2y} - C_1 \) \( \Rightarrow \) \( z = \frac{1}{C_1 + \frac{1}{2} e^{2y}} \) \( \Rightarrow \) \( (C_1 + \frac{1}{2} e^{2y}) \, dy = dx \) \( \Rightarrow \) \( C_1 y + \frac{1}{2} e^{2y} = x + C_2 \).

3. \( y'' + \left(1 + \frac{1}{y}\right)(y')^2 = 0 \) \( \Rightarrow \) \( \frac{dz}{dy} z + \left(1 + \frac{1}{y}\right) z^2 = 0 \) \( \Rightarrow \) \( \frac{dz}{z} = -(1 + \frac{1}{y}) \, dy \) \( \Rightarrow \) \( \ln |z| = -\ln |y| - y + \tilde{C}_1 \) \( \Rightarrow \) \( z = \frac{\tilde{C}_1}{y} e^{-y} \) \( \Rightarrow \) \( y e^{y} \, dy = C_1 \, dx \) \( \Rightarrow \) \( (y - 1)e^{y} = C_1 x + C_2 \).

\textbf{Linear equation}

The most general form is

\[ y'' + f(x) y' + g(x) y = r(x) \] (*)

where \( f \), \( g \) and \( r \) are specific functions of \( x \). When \( r \equiv 0 \) the equation is called \textbf{HOMOGENEOUS}.

There is no general technique for solving this equation, but some \textbf{results} relating to it are worth quoting:

1. The general solution must look like this: \( y = C_1 y_1 + C_2 y_2 + y_p \), where \( y_1 \) and \( y_2 \) are linearly independent ‘BASIC’ solutions (if we only knew how to find them!) of the corresponding \textbf{homogeneous} equation, and \( y_p \) is any \textbf{particular} solution to the \textbf{full} equation. None of these are unique (e.g. \( y_1 + y_2 \) and \( y_1 - y_2 \) is yet another basic set, etc.).
2. When one basic solution (say $y_1$) of the homogeneous version of the equation is known, the other can be found by a technique called variation of parameters (V of P): Assume that its solution has the form of $c(x)y_1(x)$, substitute this trial solution into the equation and get a first-order differential equation for $c' \equiv z$.

EXAMPLES:

1. $y'' - 4xy' + (4x^2 - 2)y = 0$ given that $y_1 = \exp(x^2)$ is a solution (verify!). Substituting $y_T(x) = c(x) \cdot \exp(x^2)$ back into the equation [remember: $y'_T = c'y_1 + \{cy'_1\}$ and $y''_T = c''y_1 + 2c'y'_1 + \{cy''_1\}$; also remember that the $c$-proportional terms, i.e. $\{\ldots\}$ and $(4x^2 - 2)cy_1$ must cancel out] yields $c'' \exp(x^2) + 4xc' \exp(x^2) - 4xc' \exp(x^2) = 0 \Rightarrow c'' = 0$. With $z \equiv c'$, the resulting equation is always of the first order: $'z' = 0 \Rightarrow z = C_1 \Rightarrow c(x) = C_1x + C_2$. Substituting back to $y_T$ results in the general solution of the original homogeneous equation: $y = C_1x \exp(x^2) + C_2 \exp(x^2)$. We can thus identify $x \exp(x^2)$ as our $y_2$.

2. $y'' + \frac{2}{x}y' + y = 0$ given that $y_1 = \frac{\sin x}{x}$ is a solution (verify!). $y_T = c(x) \frac{\sin x}{x}$ substituted: $c'' \frac{\sin x}{x} + 2c' \left(\frac{\cos x}{x} - \frac{\sin x}{x^2}\right) + \frac{2}{x} \frac{\sin x}{x} = 0 \Rightarrow c'' \sin x + 2c' \cos x = 0 \Rightarrow z' \sin x + 2z \cos x = 0 \Rightarrow \frac{dz}{x} = -2\frac{\cos x}{\sin x} \Rightarrow \ln |z| = -2 \ln |\sin x| + \tilde{C}_1 \Rightarrow z = \frac{C_1}{\sin x} \Rightarrow c(x) = C_1 \frac{\cos x}{\sin x} + C_2 \Rightarrow y = C_1 \frac{\cos x}{x} + C_2 \frac{\sin x}{x} \quad [y_2 \text{ equals to } \frac{\cos x}{x}]$. ■

When both basic solutions of the homogeneous version are known, a particular solution to the full non-homogeneous equations can be found by using an extended V-of-P idea. This time we have two unknown ‘parameters’ called $u(x)$ and $v(x)$ [in the previous case there was only one, called $c(x)$]. Note that, in this context, ‘parameters’ are actually functions of $x$.

We now derive general formulas for $u(x)$ and $v(x)$:

The objective is to solve Eq. (*) assuming that $y_1$ and $y_2$ (basic solutions of the homogeneous version) are known. We then need to find $y_p$ only, which we take to have a trial form of $u(x) \cdot y_1 + v(x) \cdot y_2$, with $u(x)$ and $v(x)$ yet to be found (the variable parameters). Substituting this into the full equation [note that terms proportional to $u(x)$, and those proportional to $v(x)$, cancel out], we get:

$$u''y_1 + 2u'y'_1 + v''y_2 + 2v'y'_2 + f(x)(u'y_1 + v'y_2) = r(x) \quad \text{(V of P)}$$

This is a single differential equation for two unknown functions ($u$ and $v$), which means we are free to impose yet another arbitrary constraint on $u$ and $v$. This is chosen to simplify the previous equation, thus:

$$u'y_1 + v'y_2 = 0$$

which further implies (after one differentiation) that $u''y_1 + u'y'_1 + v''y_2 + v'y'_2 = 0$.

The original (V of P) equation therefore simplifies to

$$u'y'_1 + v'y'_2 = r(x)$$
The last two (centered) equations can be solved (algebraically, using Cramer’s rule) for
\[
\begin{align*}
\frac{u'}{r} &= \begin{vmatrix} 0 & y_2 \\ 1 & y_2' \\ y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} \\
\frac{v'}{r} &= \begin{vmatrix} y_1 & 0 \\ 1 & y_2 \\ y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}
\end{align*}
\]
and
where the denominator it called the Wronskian of the two basic solutions (these are linearly independent iff their Wronskian is nonzero; one can use this as a check – useful later when dealing with more than two basic solutions). From the last two expressions one can easily find \(u\) and \(v\) by an extra integration (the right hand sides are known functions of \(x\)).

**EXAMPLE:**
\[
y'' - 4xy' + (4x^2 - 2)y = 4x^4 - 3.\]
We have already solved the homogeneous version getting \(y_1 = \exp(x^2)\) and \(y_2 = x \exp(x^2)\). Using the previous two formulas as we get
\[
\begin{align*}
u' &= \begin{vmatrix} \exp(x^2) & 0 \\ 2x \exp(x^2) & 4x^4 - 3 \end{vmatrix} = (3 - 4x^4) x \exp(-x^2) \\
u(x) &= \left(\frac{5}{2} + 4x^2 + 2x^4\right) e(-x^2) + C_1 \text{ and } v' &= \begin{vmatrix} \exp(x^2) & 0 \\ 2x \exp(x^2) & 4x^4 - 3 \end{vmatrix} =
\end{align*}
\]
\[
(4x^4 - 3)e(-x^2) \Rightarrow v(x) = -(3 + 2x^2) x e(-x^2) + C_2 \text{ [the last integration is a bit more tricky, but the result checks]. Simplifying } u y_1 + v y_2 \text{ yields } \left(\frac{5}{2} + x^2\right) + C_1 \exp(x^2) + C_2 x \exp(x^2), \text{ which identifies } \frac{5}{2} + x^2 \text{ as a particular solution of the full equation (this can be verified easily).} \]

Given three specific functions \(y_1, y_2\) and \(y_p\), it is possible to construct a differential equation of type (*) which has \(C_1 y_1 + C_2 y_2 + y_p\) as its general solution (that’s how I set up exam questions).

**EXAMPLE:**
Knowing that \(y = C_1 x^2 + C_2 \ln x + \frac{1}{x}\), we first substitute \(x^2\) and \(\ln x\) for \(y\) in
\[
y'' + f(x)y' + g(x)y = 0 \quad \text{[the homogeneous version]} \quad \text{to get:}
\]
\[
\begin{align*}
2 + 2x \cdot f + x^2 \cdot g &= 0 \\
&= \frac{1}{x^2} + \frac{1}{x} \cdot f + \ln x \cdot g = 0
\end{align*}
\]
and solve, algebraically, for \[
\begin{bmatrix} f \\ g \end{bmatrix} = \left[ \frac{2x}{x} \right]^{-1} \left[ \frac{-2}{\frac{1}{x}} \right] \Rightarrow f = \frac{-2 \ln x - 1}{x(2 \ln x - 1)} \quad \text{and } g = \frac{4}{x(2 \ln x - 1)}. \]
The left hand side of the equation is therefore \(y'' + \)
\[
\frac{-2 \ln x - 1}{x(2 \ln x - 1)} y' + \frac{4}{x^2(2 \ln x - 1)} y \quad \text{[one could multiply the whole equation by } x^2(2 \ln x - 1) \text{ to simplify the answer]. To ensure that } \frac{1}{x} \quad \text{is a particular solution, we substitute it into the left hand side of the last equation (for } y), \text{ yielding } r(x) \quad \text{[} = \frac{3(2 \ln x + 1)}{x^2(2 \ln x - 1)} \text{ in our case]. The final answer is thus:}
\]

\[
x^2(2 \ln x - 1)y'' - x(2 \ln x + 1)y' + 4y = \frac{3}{x}(2 \ln x + 1)
\]

**With constant coefficients**

From now on we will assume that the two 'coefficients' \( f(x) \) and \( g(x) \) are \( x \)-independent constants, and call them \( a \) and \( b \) (to differentiate between the two cases). The equation we want to solve is then

\[
y'' + ay' + by = r(x)
\]

with \( a \) and \( b \) being two specific numbers. We will start with the

\[\blacktriangleright \text{Homogeneous Case} \blacktriangleright \quad [r(x) \equiv 0]\]

All we have to do is to find two linearly independent basic solutions \( y_1 \) and \( y_2 \), and then combine them in the \( c_1y_1 + c_2y_2 \) manner (as we already know from the general case).

To achieve this, we try a solution of the following form:

\[
y_T = e^{\lambda x}
\]

where \( \lambda \) is a number whose value is yet to be determined. Substituting this into \( y'' + ay' + by = 0 \) and dividing by \( e^{\lambda x} \) results in

\[
\lambda^2 + a\lambda + b = 0
\]

which is the so called **characteristic polynomial** for \( \lambda \).

When this (quadratic) equation has two real roots the problem is solved (we have gotten our two basic solutions). What do we do when the two roots are complex, or when only a single root exists? Let us look at these possibilities, one by one.

1. **Two (distinct) real roots.**

   **EXAMPLE:** \( y'' + y' - 2y = 0 \Rightarrow \lambda^2 + \lambda - 2 = 0 \quad \text{[characteristic polynomial]} \)

   \( \Rightarrow \lambda_{1,2} = -\frac{1}{2} \pm \sqrt{\frac{1}{4} + 2} = -\frac{1}{2} \pm \frac{3}{2} = 1 \) and \(-2 \). This implies \( y_1 = e^x \) and \( y_2 = e^{-2x} \), which means that the general solution is \( y = C_1e^x + C_2e^{-2x} \).

2. **Two complex conjugate roots** \( \lambda_{1,2} = p \pm iq \).

   This implies that \( \tilde{y}_1 = e^{px}[\cos(qx) + i \sin(qx)] \) and \( \tilde{y}_2 = e^{px}[\cos(qx) - i \sin(qx)] \) [remember that \( e^{iA} = \cos A + i \sin A \)]. But at this point we are interested in real solutions only, and these are complex. But we can take the following linear combination of the above functions: \( y_1 \equiv \frac{\tilde{y}_1 + \tilde{y}_2}{2} = e^{px} \cos(qx) \) and
\[ y_2 \equiv \frac{y_1 - y_2}{2!} = e^{px} \sin(qx), \] and have a new, equivalent, basis set. The new functions are both real, thus the general solution can be written as

\[ y = e^{px}[C_1 \cos(qx) + C_2 \sin(qx)] \]

One can easily verify that both \( y_1 \) and \( y_2 \) do (individually) meet the original equation.

**EXAMPLE:** \( y'' - 2y' + 10y = 0 \Rightarrow \lambda_{1,2} = 1 \pm \sqrt{1 - 10} = 1 \pm 3i. \) Thus \( y = e^{x}[C_1 \cos(3x) + C_2 \sin(3x)] \) is the general solution.

### 3. One (double) real root.

This can happen only when the original equation has the form of: \( y'' + ay' + \frac{a^2}{4}y = 0 \) (i.e. \( b = \frac{a^2}{4} \)). Solving for \( \lambda \), one gets: \( \lambda_{1,2} = -\frac{a}{2} \pm 0 \) (double root).

This gives us only one basic solution, namely \( y_1 = e^{-\frac{a}{2}x} \); we can find the other by the V-of-P technique. Let us substitute the following trial solution \( c(x) \cdot e^{-\frac{a}{2}x} \) into the equation getting (after we divide by \( e^{-\frac{a}{2}x} \)): \( c'' - ac' + ac = 0 \Rightarrow c'' = 0 \Rightarrow c(x) = C_1 x + C_2. \) The trial solution thus becomes:

\[ y_2 = xe^{-\frac{a}{2}x} \]

as the second basic solution.

**Remember:** For duplicate roots, the second solution can be obtained by multiplying the first basic solution by \( x \).

**EXAMPLE:** \( y'' + 8y' + 16y = 0 \Rightarrow \lambda_{1,2} = -4 \) (both). The general solution is thus \( y = e^{-4x}(C_1 + C_2x). \) Let’s try finishing this as an initial-value problem [lest we forget]: \( y(0) = 1, y'(0) = -3. \) This implies \( C_1 = 1 \) and \(-4C_1 + C_2 = -3 \Rightarrow C_2 = 1. \) The final answer: \( y = (1 + x)e^{-4x}. \)

For a second-order equation, these three possibilities cover the whole story.

**Non-homogeneous Case**

When any such equation has a nonzero right hand side \( r(x) \), there are two possible ways of building a particular solution \( y_p \):

**Using the variation-of-parameters formulas derived earlier for the general case.**

**EXAMPLES:**

1. \( y'' + y = \tan x \Rightarrow \lambda^2 + 1 = 0 \Rightarrow \lambda_{1,2} = \pm i \Rightarrow \sin x \) and \( \cos x \) being the two basic solutions of the homogeneous version. The old formulas give: \( u' = \)

\[
\begin{vmatrix}
0 & \cos x \\
\tan x & -\sin x \\
\sin x & \cos x \\
\cos x & -\sin x
\end{vmatrix} = \sin x \Rightarrow u(x) = -\cos x + C_1 \text{ and } u' = \]

\[
\begin{vmatrix}
\sin x & 0 \\
\cos x & \tan x \\
\sin x & \cos x \\
\cos x & -\sin x
\end{vmatrix} = \]
\[-\frac{\sin^2 x}{\cos x} = \cos x - \frac{1}{\cos x} \Rightarrow v(x) = \sin x - \ln \left(\frac{1 + \sin x}{\cos x}\right) + C_2.\] The final solution is thus \(y = \{-\cos x \sin x + \sin x \cos x\} - \cos x \ln \left(\frac{1 + \sin x}{\cos x}\right) + C_1 \sin x + C_2 \cos x\) [the terms inside the curly brackets cancelling out, which happens frequently in these cases].

2. \(y'' - 4y' + 4y = \frac{e^{2x}}{x}\). Since \(\lambda_{1,2} = 2 \pm 0\) [double root], the basic solutions are \(e^{2x}\) and \(xe^{2x}\). \(u' = \begin{vmatrix} 0 & xe^{2x} \\ \frac{e^{2x}}{x} & \frac{x e^{2x}}{x} \end{vmatrix} = -1 \Rightarrow u(x) = -x + C_1\) and \(v' = \begin{vmatrix} e^{2x} & 0 \\ \frac{e^{2x}}{x} & \frac{xe^{2x}}{x} \end{vmatrix} = 1 \Rightarrow v(x) = \ln x + C_2.\)

Answer: \(y = \frac{e^{2x}}{x} + C_1 e^{2x} + C_2 e^{2x} - xe^{2x} + \ln x \cdot xe^{2x} = e^{2x}(C_1 + C_2 x + x \ln x).\) □

\[\textbf{Special} \text{ cases of } r(x)\]

- \(r(x)\) is a \textbf{polynomial} in \(x\):

Use a polynomial of the \textbf{same degree} but with \textbf{undetermined coefficients} as a trial solution for \(y_p\).

\underline{EXAMPLE}: \(y'' + 2y' - 3y = x\), \(\lambda_{1,2} = 1\) and \(-3 \Rightarrow y_1 = e^x\) and \(y_2 = e^{-3x}\). \(y_p = Ax + B\), where \(A\) and \(B\) are found by substituting this \(y_p\) into the full equation and getting: \(2A - 3Ax - 3B = x \Rightarrow A = -\frac{1}{3}\) and \(B = -\frac{2}{9}\). Answer: \(y = C_1 e^x + C_2 e^{-3x} - x^2 - \frac{2}{9}\).

\underline{Exceptional case}: When \(\lambda = 0\), this will not work unless the trial solution \(y_p\) is further multiplied by \(x\) (when \(\lambda = 0\) is a multiple root, \(x\) has to be raised to the multiplicity of \(\lambda\)).

\underline{EXAMPLE}: \(y'' - 2y' = x^2 + 1\), \(\lambda_{1,2} = 0\) and \(2 \Rightarrow y_1 = 1\) and \(y_2 = e^{2x}\). \(y_p = Ax^3 + Bx^2 + Cx\) where \(A\), \(B\) and \(C\) are found by substituting: \(6Ax + 2B - 2(3Ax^2 + 2Bx + C) = x^2 + 1 \Rightarrow A = -\frac{1}{6}, B = -\frac{1}{4}\) and \(C = -\frac{3}{4}\). Answer: \(y = C_1 + C_2 e^{2x} - \frac{x^3}{6} - \frac{x^2}{4} - \frac{3x}{4}\). □

- \(r(x) \equiv ke^{ax}\) (an \textbf{exponential} term):

The trial solution is \(y_p = Ae^{ax}\) [with only \(A\) to be found; this is the \textbf{undetermined coefficient} of this case].

\underline{EXAMPLE}: \(y'' + 2y' + 3y = 3e^{-2x}\) \(\Rightarrow \lambda_{1,2} = -1 \pm \sqrt{2}i\), \(y_p = Ae^{-2x}\) substituted gives: \(A(4 - 4 + 3)e^{-2x} = 3e^{-2x} \Rightarrow A = 1\). Answer: \(y_p = e^{-x}[C_1 \sin(\sqrt{2}x) + C_2 \cos(\sqrt{2}x)] + e^{-2x}..\).
Exceptional case: When $\alpha = \lambda$ [any of the roots], the trial solution must be first multiplied by $x$ (to the power of the multiplicity of this $\lambda$).

EXAMPLE: $y'' + y' - 2y = 3e^{-2x} \Rightarrow \lambda_{1,2} = 1$ and $-2$ [same as $\alpha$]! $y_p = Ax e^{-2x}$, substituted: $A(4x - 4) + A(1 - 2x) - 2Ax = 3 \Rightarrow A = -1$ [this follows from the absolute term, the $x$-proportional terms cancel out, as they must].

Answer: $y = C_1 e^x + (C_2 - x)e^{-2x}$.

- $r(x) = k_s e^{px} \sin(qx)$ [or $k_c e^{px} \cos(qx)$, or a combination (sum) of both]:

The trial solution is $[A \sin(qx) + B \cos(qx)]e^{px}$.

EXAMPLE: $y'' + y' - 2y = 2e^{-x} \sin(4x) \Rightarrow \lambda_{1,2} = 1$ and $-2$ [as before]. $y_p = [A \sin(4x) + B \cos(4x)]e^{-x}$, substituted into the equation: $-18[A \sin(4x) + B \cos(4x)] - 4[A \cos(4x) - B \sin(4x)] = 2 \sin(4x) \Rightarrow -18A + 4B = 2$ and $-4A - 18B = 0 \Rightarrow \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} -18 & 4 \\ -4 & -18 \end{bmatrix}^{-1} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{9}{35} \\ \frac{2}{35} \end{bmatrix} \Rightarrow y = C_1 e^x + C_2 e^{-2x} + \left(\frac{2}{35} \cos(4x) - \frac{9}{35} \sin(4x)\right) e^{-2x}$.

Exceptional case: When $\lambda = p + iq$ [both the real and purely imaginary parts must agree], the trial solution acquires the standard factor of $x$.

'Special-case’ summary:

We would like to mention that all these special case can be covered by one and the same rule: When $r(x) = P_n(x)e^{\beta x}$, where $P_n(x)$ is an $n$-degree polynomial in $x$, the trial solution is $Q_n(x)e^{\beta x}$, where $Q_n(x)$ is also an $n$-degree polynomial, but with ‘undetermined’ (i.e. yet to be found) coefficients.

And the same exception: When $\beta$ coincides with a root of the characteristic polynomial (of multiplicity $\ell$) the trial solution must be further multiplied by $x^{\ell}$.

If we allowed complex solutions, these rules would have covered it all. Since we don’t, we have to spell it out differently for $\beta = p + iq$:

When $r(x) = [P_s(x) \sin(qx) + P_c(x) \cos(qx)]e^{px}$ where $P_{s,c}$ are two polynomials of degree not higher than $n$ [i.e. $n$ is the higher of the two; also: one $P$ may be identically equal to zero], the trial solution is: $[Q_s(x) \sin(qx) + Q_c(x) \cos(qx)]e^{px}$ with both $Q_{s,c}$ being polynomials of degree $n$ [no compromise here – they both have to be there, with the full degree, even if one $P$ is missing].

Exception: If $p + iq$ coincides with one of the $\lambda$s, the trial solution must be further multiplied by $x$ raised to the $\lambda$’s multiplicity [note that the conjugate root $p - iq$ will have the same multiplicity; use the multiplicity of one of these – don’t double it] $\Box$

Finally, if the right hand side is a linear combination (sum) of such terms, we use the superposition principle to construct the overall $y_p$. This means we find $y_p$ individually for each of the distinct terms of $r(x)$, then add them together to build the final solution.
EXAMPLE: \( y'' + 2y' - 3y = x + e^{-x} \Rightarrow \lambda_{1,2} = 1, -3 \). We break the right hand side into \( r_1 \equiv x \) and \( r_2 \equiv e^{-x} \), construct \( y_{1p} = Ax + B \Rightarrow [ \text{when substituted into the equation with only } x \text{ on the right hand side} ] \ 2A - 3Ax - 3B = x \Rightarrow A = -\frac{1}{3} \) and \( B = -\frac{2}{9} \), and then \( y_{2p} = Ce^{-x} \), substituted into the equation with \( r_2 \) only: \( C - 2C - 3C = 1 \Rightarrow C = -\frac{1}{4} \).

Answer: \( y = C_1e^x + C_2e^{-3x} - \frac{x}{3} - \frac{2}{9} - \frac{1}{4}e^{-x} \).

Cauchy equation looks like this:
\[(x - x_0)^2y'' + a(x - x_0)y' + by = r(x)\]
where \( a, b \) and \( x_0 \) are specific constants (\( x_0 \) is usually equal to 0, e.g. \( x^2y'' + 2xy' - 3y = x^5 \)).

There are two ways of solving it:

- **Converting**

it to the previous case of a **constant-coefficient** equation (convenient when \( r(x) \) is a polynomial in either \( x \) or \( \ln x \) – try to figure out why). This conversion is achieved by introducing a new independent variable \( t = \ln(x - x_0) \). We have already derived the following set of formulas for performing such a conversion: \( y' = \frac{dy}{dt} \cdot \frac{dt}{dx} = \frac{\dot{y}}{x-x_0} \) and \( y'' = \frac{d^2y}{dt^2} \cdot \left( \frac{dt}{dx} \right)^2 + \frac{d}{dt} \cdot \frac{dy}{dx} \cdot \frac{dt}{dx} = \frac{\ddot{y}}{(x-x_0)^2} - \frac{\dot{y}}{(x-x_0)^2} \). The original Cauchy equation thus becomes
\[ \ddot{y} + (a - 1)\dot{y} + by = r(x_0 + e^t) \]
which we solve by the old technique.

EXAMPLE: \( x^2y'' - 4xy' + 6y = \frac{42}{3} \Rightarrow \ddot{y} - 5\dot{y} + 6y = 42e^{-4t} \Rightarrow \lambda_{1,2} = 2, 3 \) and \( y_p = Ae^{-4t} \) substituted: \( 16A + 20A + 6A = 42 \Rightarrow A = 1 \).

Answer: \( y = C_1e^{3t} + C_2e^{2t} + e^{-4t} = C_1x^3 + C_2x^2 + \frac{1}{2x} \) [since \( x = e^t \)].

Problems for extra practice

(solve by **undetermined coefficients** – not V of P):

1. \( y'' - \frac{y'}{x} - \frac{3y}{x^2} = \ln x + 1 \Rightarrow [ \text{must be multiplied by } x^2 \text{ first} ] \ \ddot{y} - 2\dot{y} - 3y = te^{2t} + e^{2t} \Rightarrow \ldots \)

Answer: \( y = C_1x + \frac{C_2}{x} - \left( \frac{5}{y} + \frac{\ln x}{3} \right)x^2. \)

2. \( x^2y'' - 2xy' + 2y = 4x + \sin(\ln x) \Rightarrow \ddot{y} - 3\dot{y} + 2y = 4e^t + \sin(t) \Rightarrow \ldots \)

Answer: \( y = C_1x + C_2x^2 - 4x \ln x + \frac{1}{10} \sin(\ln x) + \frac{3}{10} \cos(\ln x). \)

**Warning:** To use the undetermined-coefficients technique (via the \( t \) transformation) the equation must have (or must be brought to) the form of: \( x^2y'' + axy' + \ldots \); to use the V-of-P technique, the equation must have the \( y'' + \frac{a}{x}y' + \ldots \) form.
(This is more convenient when \( r(x) \) is either equal to zero, or does not have the special form mentioned above). We substitute a trial solution \((x - x_0)^m\), with \( m \) yet to be determined, into the homogeneous Cauchy equation, and divide by \((x - x_0)^m\). This results in:

\[
m^2 + (a - 1)m + b = 0
\]

a characteristic polynomial for \( m \). With two distinct real roots, we get our two basic solutions right away; with a duplicate root, we need an extra factor of \( \ln(x - x_0) \) to construct the second basic solution; with two complex roots, we must go back to the 'conversion' technique.

**EXAMPLES:**

1. \( x^2y'' + xy' - y = 0 \) \( \Rightarrow m^2 - 1 = 0 \) \( \Rightarrow m_{1,2} = \pm 1 \)
   
   Answer: \( y = C_1x + C_2 \frac{1}{x} \).

2. \( x^2y'' + 3xy' + y = 0 \) \( \Rightarrow m^2 + 2m + 1 = 0 \) \( \Rightarrow m_{1,2} = -1 \) (duplicate) \( \Rightarrow y = \frac{C_1}{x} + \frac{C_2}{x} \ln x \).

3. \( 3(2x - 5)^2y'' - (2x - 5)y' + 2y = 0 \). First we have to rewrite it in the standard form of: \( (x - \frac{5}{2})^2y'' - \frac{1}{5}(x - \frac{5}{2})y' + \frac{1}{5}y = 0 \) \( \Rightarrow m^2 - \frac{7}{6}m + \frac{1}{6} = 0 \) \( \Rightarrow m_{1,2} = \frac{1}{6}, \frac{1}{2} \); \( \Rightarrow y = \tilde{C}_1(x - \frac{5}{2}) + \tilde{C}_2(x - \frac{5}{2})^{1/6} = C_1(2x - 5) + C_2(2x - 5)^{1/6} \).

4. \( x^2y'' - 4xy' + 4y = 0 \) with \( y(1) = 4 \) and \( y'(1) = 13 \). First \( m^2 - 5m + 4 = 0 \) \( \Rightarrow m_{1,2} = 1, 4 \) \( \Rightarrow y = C_1x + C_2x^4 \). Then \( C_1 + C_2 = 4 \) and \( C_1 + 4C_2 = 13 \) give \( C_2 = 3 \) and \( C_1 = 1 \).
   
   Answer: \( y = x + 3x^4 \).

5. \( x^2y'' - 4xy' + 6y = x^4 \sin x \). To solve the homogeneous version: \( m^2 - 5m + 6 = 0 \) \( \Rightarrow m_{1,2} = 2, 3 \) \( \Rightarrow y_1 = x^2 \) and \( y_2 = x^3 \). To use V-of-P formulas the equation must be first rewritten in the 'standard' form of \( y'' - \frac{4}{x}y' + \frac{6}{x^2}y = \)

\[
x^2 \sin x \Rightarrow u' = \begin{vmatrix} 0 & x^3 \\ x^2 \sin x & 3x^2 \\ x^2 \end{vmatrix} = -x \sin x \Rightarrow u(x) = x \cos x - \sin x + C_1 \text{ and}
\]

\[
v' = \begin{vmatrix} x^2 & 0 \\ 2x & x^2 \sin x \\ 2x \end{vmatrix} = \sin x \Rightarrow v(x) = -\cos(x) + C_2.
\]

Solution: \( y = (C_1 - \sin x)x^2 + C_2x^3 \) [the rest cancelled out — common occurrence when using this technique].
Chapter 4  THIRD AND HIGHER-ORDER LINEAR ODES

First we extend the general linear-equation results to higher orders. Explicitly, we mention the third order only, but the extension to higher orders is quite obvious.

A third order linear equation

\[ y''' + f(x)y'' + g(x)y' + h(x)y = r(x) \]

has the following general solution: \( y = C_1y_1 + C_2y_2 + C_3y_3 + y_p \), where \( y_1, y_2 \) and \( y_3 \) are three basic (linearly independent) solutions of the homogeneous version. There is no general analytical technique for finding them. Should these be known (obtained by whatever other means), we can construct a particular solution (to the full equation) \( y_p \) by the \textit{variation of parameters} (this time we skip the details), getting:

\[
\begin{vmatrix}
0 & y_2 & y_3 \\
0 & y_2' & y_3' \\
r & y_2'' & y_3'' \\
y_1 & y_2 & y_3 \\
y_1' & y_2' & y_3' \\
y_1'' & y_2'' & y_3''
\end{vmatrix}
\]

with a similar formula for \( u' \) and for \( w' \) (we need three of them, one for each basic solution). The pattern of these formulas should be obvious: there is the Wronskian in the denominator, and the same matrix with one of its columns (the first for \( u' \), the second for \( v' \), and the last for \( w' \)) replaced by \( \begin{bmatrix} 0 \\ 0 \\ r \end{bmatrix} \) in the numerator.

The corresponding constant-coefficient equation can be solved easily by constructing its characteristic polynomial and finding its roots, in a manner which is a trivial extension of the second-degree case. The main difficulty here is finding roots of higher-degree polynomials. We will take up this issue first.

Polynomial roots

We start with a general cubic polynomial (we will call its variable \( x \) rather than \( \lambda \))

\[ x^3 + a_2x^2 + a_1x + a_0 = 0 \]

such as, for example \( x^3 - 2x^2 - x + 2 = 0 \). Finding its three roots takes the following steps:

1. Compute \( Q = \frac{3a_1 - a_2^2}{9} \left[= -\frac{7}{9}\right] \) and \( R = \frac{9a_1a_2 - 27a_0 - 2a_3^2}{54} \left[= -\frac{10}{27}\right] \).

2. Now consider two cases:

   (a) \( Q^3 + R^2 \geq 0 \).
Compute \( s = \sqrt[3]{R + \sqrt{Q^3 + R^2}} \) and \( t = \sqrt[3]{R - \sqrt{Q^3 + R^2}} \) [each being a real but possibly negative number]. The original equation has one real root given by \( s + t - \frac{a_2}{3} \), and two complex roots given by \( -\frac{s + t}{2} - \frac{a_2}{3} \pm \frac{\sqrt{3}}{2} (s - t) i \).

(b) \( Q^3 + R^2 < 0 \) [our case \( = -\frac{1}{3} \)].
Compute \( \theta = \arccos \frac{R}{\sqrt{-Q^3}} \) [\( Q \) must be negative]. The equation has three real roots given by \( 2\sqrt{-Q} \cos \left( \frac{\theta + 2\pi k}{3} \right) - \frac{a_2}{3} \), where \( k = 0, 1 \) and \( 2 \). [In our case \( \theta = \arccos \left( -\frac{10}{27} \sqrt{\frac{729}{4323}} \right) \Rightarrow 2\sqrt{\frac{7}{9}} \cos \left( \frac{\theta}{3} \right) + \frac{2}{3} = 2, \]
\( 2\sqrt{\frac{7}{9}} \cos \left( \frac{\theta + 2\pi}{3} \right) + \frac{2}{3} = -1 \) and \( 2\sqrt{\frac{7}{9}} \cos \left( \frac{\theta + 4\pi}{3} \right) + \frac{2}{3} = 1 \) are the three roots.

Proof (for the \( s-t \) case only): Let \( x_1, x_2 \) and \( x_3 \) be the three roots of the formula. Expand \((x - x_1)(x - x_2)(x - x_3) = [x + \frac{a_2}{3} - (s + t)](x + \frac{a_2}{3} + \frac{3}{2}(s - t))^2 = x^3 + a_2x^2 + \left( \frac{a_2}{3} - 3st \right)x + \frac{a_2^2}{27} - sta_2 - s^3 - t^3 \). Since \( -st = Q \) and \( -s^3 - t^3 = -2R \), we can see quite easily that \( \frac{a_2}{3} - 3st = a_1 \) and \( \frac{a_2^2}{27} - sta_2 - s^3 - t^3 = a_0 \). \( \square \)

\[ \text{And now we tackle the forth-degree (quartic) polynomial} \]
\[ x^4 + a_3x^3 + a_2x^2 + a_1x + a_0 = 0 \]
such as, for example \( x^4 + 4x^3 + 6x^2 + 4x + 1 = 0 \).

1. First solve the following cubic
\[ y^3 - a_2y^2 + (a_1a_3 - 4a_0)y + (4a_0a_2 - a_1^2 - a_0a_3^2) = 0 \]
\[ [y^3 - 6y^2 + 12y - 8 = 0 \text{ in the case of our example – the three roots are all equal to } 2]. \]

2. Then solve the following two quadratic equations
\[ z^2 + \frac{1}{2} \left( a_3 \pm \sqrt{a_3^2 - 4a_2 + 4y_1} \right) z + \frac{1}{2} \left( y_1 \mp \sqrt{y_1^2 - 4a_0} \right) = 0 \]
when \( 2a_1 - a_3y_1 \geq 0 \), or
\[ z^2 + \frac{1}{2} \left( a_3 \pm \sqrt{a_3^2 - 4a_2 + 4y_1} \right) z + \frac{1}{2} \left( y_1 \pm \sqrt{y_1^2 - 4a_0} \right) = 0 \]
when \( 2a_1 - a_3y_1 \leq 0 \), where \( y_1 \) is the largest real root of the previous cubic \([z^2 + 2z + 1 = 0, \text{twice, in our example}]. \) The resulting four roots are those of the original quartic \([-1, -1, -1 \text{ and } -1]\).
Proof: We assume that $2a_1 - a_3 y_1 > 0$ (the other case would be a carbon copy). By multiplying the left hand sides of the corresponding two quadratic equations one gets:

$$z^4 + a_3 z^3 + a_2 z^2 + \frac{a_3 y_1 + \sqrt{a_3^2 - 4a_2 + 4y_1 \sqrt{y_1^2 - 4a_0}}}{2} z + a_0$$

It remains to be shown that the linear coefficient is equal to $a_1$. This amounts to:

$$2a_1 - a_3 y_1 = \sqrt{a_3^2 - 4a_2 + 4y_1 \sqrt{y_1^2 - 4a_0}}$$

Since each of the two expressions under square root must be non-negative (see the Extra Proof below), the two sides of the equation have the same sign. It is thus legitimate to square them, obtaining a cubic equation for $y_1$, which is identical to the one we solved in (1). □

Extra Proof: Since $y_1$ is the (largest real) solution to $(a_3^2 - 4a_2 + 4y_1) \cdot (y_1^2 - 4a_0) = (2a_1 - a_3 y_1)^2$, it is clear that both factors on the LHS must have the same sign. We thus have to prove that either of them is positive. Visualizing the graph of the cubic polynomial $(a_3^2 - 4a_2 + 4y_1) \cdot (y_1^2 - 4a_0) - (2a_1 - a_3 y_1)^2$, it is obvious that by adding $(2a_1 - a_3 y_1)^2$ to it, the largest real root can only decrease. This means that $y_1$ must be bigger than each of the real roots of $(a_3^2 - 4a_2 + 4y_1) \cdot (y_1^2 - 4a_0) = 0$, implying that $a_2 - \frac{a_3^2}{4} < y_1$ (which further implies that $4a_0 < y_1^2$). □

There is no general formula for solving fifth-degree polynomials and beyond (investigating this led Galois to discover groups); we can still find the roots numerically to any desired precision (the algorithms are fairly tricky though, and we will not go into this).

There are special cases of higher-degree polynomials which we know how to solve (or at least how to reduce them to a lower-degree polynomial). Let us mention a handful of these:

1. $x^n = a$, by finding all ($n$ distinct) complex values of $\sqrt[n]{a}$, i.e., $\sqrt[n]{a} \left( \cos \frac{2\pi k}{n} + i \sin \frac{2\pi k}{n} \right)$ when $a > 0$ and $\sqrt[n]{a} \left( \cos \frac{2\pi k + \pi}{n} + i \sin \frac{2\pi k + \pi}{n} \right)$ when $a < 0$, both with $k = 0, 1, 2, ..., n - 1$.

   Examples: (i) $\sqrt[4]{16} = 2, -2, 2i$ and $-2i$ (ii) $\sqrt[-2]{-8} = -2$ and $1 \pm \sqrt[3]{2}i$.

2. When 0 is one of the roots, it’s trivial to find it, with its multiplicity.

   Example: $x^4 + 2x^3 - 4x^2 = 0$ has obviously 0 as a double root. Dividing the equation by $x^2$ makes it into a quadratic equation which can be easily solved.

3. When coefficients of an equation add up to 0, 1 must be one of the roots. The left hand side is divisible by $(x - 1)$ [synthetic division], which reduces its order.

   Example: $x^3 - 2x^2 + 3x - 2 = 0$ thus leads to $(x^3 - 2x^2 + 3x - 2) \div (x - 1) = x^2 - x + 2$ [quadratic polynomial].
4. When coefficients of the odd powers of $x$ and coefficients of the even powers of $x$ add up to the same two answers, then $-1$ is one of the roots and $(x + 1)$ can be factored out.

Example: $x^3 + 2x^2 + 3x + 2 = 0$ leads to $(x^3 + 2x^2 + 3x + 2) ÷ (x + 1) = x^2 + x + 2$ and a quadratic equation.

5. Any lucky guess of a root always leads to the corresponding reduction in order (one would usually try $2, -2, 3, -3$, etc. as a potential root).

6. One can cut the degree of an equation in half when the equation has even powers of $x$ only by introducing $z = x^2$.

Example: $x^4 - 3x^2 - 4 = 0$ thus reduces to $z^2 - 3z - 4 = 0$ which has two roots $z_{1,2} = -1, 4$. The roots of the original equation thus are: $x_{1,2,3,4} = i, -i, 2, -2$.

7. Similarly with odd powers only.

Example: $x^5 - 3x^3 - 4x = 0$. Factoring out $x$ [there is an extra root of $0$ ], this becomes the previous case.

8. All powers divisible by $3$ ($4$, $5$, etc.). Use the same trick.

Example: $x^6 - 3x^3 - 4 = 0$. Introduce $z = x^3$, solve for $z_{1,2} = -1, 4$ [same as before]. Thus $x_{1,2,3} = \sqrt[3]{-1} = -1, \frac{1}{2} \pm \frac{\sqrt{3}}{2}i$ and $x_{4,5,6} = \sqrt[4]{4}, \sqrt[4]{(-\frac{1}{2} \pm \frac{\sqrt{3}}{2}i)} = 1.5874, -0.7937 + 1.3747i, = -0.7937 - 1.3747i$.

9. When a multiple root is suspected (the question may indicate: 'there is a triple root'), the following will help: each differentiation of a polynomial reduces the multiplicity of its every root by one. This means, for example, that a triple root becomes a single root of the polynomial's second derivative.

Example: $x^4 - 5x^3 + 6x^2 + 4x - 8 = 0$, given there is a triple root. Differentiating twice: $12x^2 - 30x + 12 = 0 \Rightarrow x_{1,2} = \frac{1}{2}, 2$. These must be substituted back into the original equation [only one of these is its triple root, the other is not a root at all]. Since $2$ meets the original equation ($\frac{1}{2}$ does not), it must be its triple root, which can be factored out, thus: $(x^4 - 5x^3 + 6x^2 + 4x - 8) ÷ (x - 2)^3 = x + 1$ [ending up with a linear polynomial]. The last root is thus, trivially, equal to $-1$.

10. Optional: One an take a slightly more sophisticated approach when it comes to multiple roots. As was already mentioned: each differentiation of the polynomial reduces the multiplicity of every root by one, but may (and usually does) introduce a lot of extra 'phoney' roots. These can be eliminated by taking the greatest common divisor (GCD) of the polynomial and its derivative, by using Euclid's algorithm, which works as follows:

To find the GCD of two polynomials $p$ and $q$, we divide one into the other to find the remainder (residue) of this operation (we are allowed to multiply the result by a constant to make it a monic polynomial): $r_1 = \text{Res}(p ÷ q)$, then $r_2 = \text{Res}(q ÷ r_1), r_3 = \text{Res}(r_1 ÷ r_2), ...$ until the remainder becomes zero. The GCD is the previous (last nonzero) $r$. 
Example: \( p(x) = x^8 - 28x^7 + 337x^6 - 2274x^5 + 9396x^4 - 24312x^3 + 38432x^2 - 33920x + 12800. \)
\( s(x) = \gcd(p, p') = x^5 - 17x^4 + 112x^3 - 356x^2 + 544x - 320 \)

[the \( r \)-sequence was: \( x^6 - \frac{35}{4} x^4 + \frac{733}{4} x^3 - 832 x^2 + 2057 x^2 - 2632 x + 1360 \), \( s, 0 \)]. \( u(x) = \gcd(s, s') = x^2 - 6x + 8 \) [the \( r \)-sequence: \( x^3 - \frac{31}{3} x^2 + 34 x - \frac{104}{3} \), \( u, 0 \)]. The last (quadratic) polynomial can be solved (and thus factorized) easily: \( u = (x - 2)(x - 4). \) Thus 2 and 4 (each) must be a double root of \( s \) and a triple root of \( p \). Taking \( s \div (x - 2)^2(x - 4)^2 \) reveals that 5 is a single root of \( s \) and therefore a double root of \( p \). Thus, we have found all eight roots of \( p: 2, 2, 2, 4, 4, 4, 5 \) and 5. ■

**Constant-coefficient equations**

Similarly to solving second-order equations of this kind, we

- find the roots of the characteristic polynomial,
- based on these, construct the basic solutions of the homogeneous equation,
- find \( y_p \) by either V-P or (more commonly) undetermined-coefficient technique (which requires only a trivial and obvious extension). ■

Since the Cauchy equation is effectively a linear equation in disguise, we know how to solve it (beyond the second order) as well.

**EXAMPLES:**

1. \( y^{iv} + 4y'' + 4y = 0 \Rightarrow \lambda^4 + 4\lambda^2 + 4 = 0 \Rightarrow z(= \lambda^2)_{1,2} = -2 \) [double] \( \Rightarrow \lambda_{1,2,3,4} = \pm \sqrt{2} i \) [both are double roots] \( \Rightarrow y = C_1 \sin(\sqrt{2}x) + C_2 \cos(\sqrt{2}x) + C_3 x \sin(\sqrt{2}x) + C_4 x \cos(\sqrt{2}x). \)

2. \( y''' + y'' + y' + y = 0 \Rightarrow \lambda_1 = -1, \) and \( \lambda^2 + 1 = 0 \Rightarrow \lambda_{2,3} = \pm i \Rightarrow y = C_1 e^{-x} + C_2 \sin x + C_3 \cos(x). \)

3. \( y'' - 3y' + 3y'' - y = 0 \Rightarrow \lambda_{1,2} = 0, \) \( \lambda_3 = 1 \) and \( \lambda^2 - 2\lambda + 1 = 0 \Rightarrow \lambda_{4,5} = 1 \) \( \Rightarrow y = C_1 + C_2 x + C_3 e^x + C_4 x e^x + C_5 x^2 e^x. \)

4. \( y^{iv} - 5y'' + 4y = 0 \Rightarrow z(= \lambda^2)_{1,2} = 1, \) \( 4 \Rightarrow \lambda_{1,2,3,4} = \pm 2, \) \( \pm 1 \Rightarrow y = C_1 e^x + C_2 e^{-2x} + C_3 e^{2x} + C_4 e^{2x}. \)

5. \( y''' - y' = 10 \cos(2x) \Rightarrow \lambda_{1,2,3} = 0, \pm 1. \) \( y_p = A \sin(2x) + B \cos(2x) \Rightarrow -10 \cos(2x) = 10 \sin(2x) \Rightarrow A = -1, B = 0 \Rightarrow y = C_1 + C_2 e^x + C_3 e^{-x} - \sin(2x). \)

6. \( y''' + 2y'' = x^2 - 1 \Rightarrow \lambda_{1,2} = 0, \) \( \lambda_3 = 2. \) \( y_p = A x^4 + B x^3 + C x^2 \Rightarrow -24 Ax^2 + 24 A x - 12 B x + 6 B - 4 C = x^2 - 1 \Rightarrow A = -\frac{1}{24}, B = -\frac{1}{12}, C = \frac{1}{8} \Rightarrow y = C_1 + C_2 x + C_3 x^2 e^{x^2} - \frac{x^4}{24} + \frac{x^3}{12} + \frac{x^2}{8}. \)

7. \( x^3 y''' + x^2 y'' - 2xy' + 2y = 0 \) (homogeneous Cauchy). \( m(m - 1)(m - 2) + m(m - 1) - 2m + 2 = 0 \) is the characteristic polynomial. One can readily notice \( (m - 1) \) being a common factor, which implies \( m_1 = 1 \) and \( m^2 - m - 2 = 0 \Rightarrow m_{2,3} = -1, \) \( 2 \Rightarrow y = C_1 x + C_2 x + C_3 x^2. \)
Chapter 5  SETS OF LINEAR, FIRST-ORDER, CONSTANT-COEFFICIENT ODES

First we need to complete our review of Matrix Algebra

There are two ways of finding these

▷ The ‘classroom’ algorithm

which requires \( \propto n! \) number of operations and becomes not just impractical, but virtually impossible to use (even for supercomputers) when \( n \) is large (>20). Yet, for small matrices \( (n \leq 4) \) it’s fine, and we actually prefer it. It works like this:

- In a 2×2 case, it’s trivial: 
  \[
  \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \frac{1}{ad - bc}
  \]
  where the denominator is the determinant.
  
  Example:
  \[
  \begin{pmatrix} 2 & 4 \\ -3 & 5 \end{pmatrix}^{-1} = \begin{pmatrix} 5 & -4 \\ -6 & 2 \end{pmatrix}
  \]
  \( 22 \) being the determinant.

- The 3×3 case is done in four steps.

  Example:
  \[
  \begin{pmatrix} 2 & 4 & -1 \\ 0 & 3 & 2 \\ 2 & 1 & 4 \end{pmatrix}^{-1}
  \]

  1. Construct a 3×3 matrix of all 2×2 subdeterminants (striking out one row and one column – organize the answers accordingly):

  \[
  \begin{pmatrix} 10 & -4 & -6 \\ 17 & 10 & -6 \\ 11 & 4 & 6 \end{pmatrix}
  \]

  2. Transpose the answer:

  \[
  \begin{pmatrix} 10 & 17 & 11 \\ -4 & 10 & 4 \\ -6 & -6 & 6 \end{pmatrix}
  \]

  3. Change the sign of every other element (using the following checkerboard scheme: 

  \[
  \begin{array}{ccc}
  + & - & + \\
  - & + & - \\
  + & - & + \\
  \end{array}
  \]

  thus:

  \[
  \begin{pmatrix} 10 & -17 & 11 \\ 4 & 10 & -4 \\ -6 & 6 & 6 \end{pmatrix}
  \]
4. Divide by the determinant, which can be obtained easily by multiplying ('scalar' product) the first row of the original matrix by the first column of the last matrix (or vice versa — one can also use the second or third row/column — column/row):

\[
\begin{bmatrix}
10 & 17 & 11 \\
14 & 17 & 12 \\
16 & 17 & 12 \\
12 & 17 & 12 \\
\end{bmatrix}
\]

If we need the determinant only, there is an easier scheme:

\[
\begin{bmatrix}
2 & 4 & -1 \\
0 & 3 & 2 \\
2 & 4 & -1 \\
0 & 3 & 2 \\
\end{bmatrix}
\]

resulting in

\[
2 \times 3 \times 4 + 0 \times 1 \times (-1) + 2 \times 4 \times 2 - (-1) \times 3 \times 2 - 2 \times 1 \times 2 - 4 \times 4 \times 0 = 42.
\]

- Essentially the same algorithm can be used for 4 × 4 matrices and beyond, but it becomes increasingly impractical and soon enough virtually impossible to carry out.

▷ The ‘practical’ algorithm

requires \( \propto n^3 \) operations and can be easily converted into a computer code:

1. The original \((n \times n)\) matrix is extended to an \(n \times 2n\) matrix by appending it with the \(n \times n\) unit matrix.

2. By using one of the following three ‘elementary’ operations we make the original matrix into the unit matrix, while the appended part results in the desired inverse:

   (a) A (full) row can be divided by any nonzero number [this is used to make the main-diagonal elements equal to 1, one by one].

   (b) A multiple of a row can be added to (or subtracted from) any other row [this is used to make the non-diagonal elements of each column equal to 0].

   (c) Two rows can be interchanged whenever necessary [when a main-diagonal element is zero, interchange the row with any subsequent row which has a nonzero element in that position - if none exists the matrix is singular].

The product of the numbers we found on the main diagonal (and had to divide by), further multiplied by \(-1\) if there has been an odd number of interchanges, is the matrix’ determinant.

- A \(4 \times 4\) EXAMPLE:
The last matrix is the inverse of the original matrix, as can be easily verified [no interchanges were needed]. The determinant is $3 \times (-1) \times (-\frac{1}{3}) \times 7 = 7$.

Solving $n$ equations for $m$ unknowns

For an $n \times n$ non-singular problems with $n$ 'small' we can use the matrix inverse: $A\mathbf{x} = \mathbf{b} \Rightarrow \mathbf{x} = A^{-1}\mathbf{b}$, but this is not very practical beyond $2 \times 2$.

The fully general technique

which is applicable to singular as well as $n$ by $m$ problems works like this:

1. Extend $A$ by an extra column $\mathbf{b}$.

2. Using 'elementary operations' attempt to make the original $A$-part of the matrix into the unit matrix (no need to keep track of interchanges). If you succeed, the $\mathbf{b}$-part of the matrix is the (unique) solution. This of course cannot work when the number of equations and the number of unknowns don’t match. Furthermore, we may run into difficulty for the following two reasons:

   (a) We may come to a column which has 0 on the main diagonal and all elements below it (in the same column). This column will be then skipped (as if it never existed, i.e. we will try to get 1 in the same position of the next column).

   (b) Discarding the columns we skipped, we may end up with fewer columns than rows [resulting in some extra rows with only zeros in their $A$-part], or the other way round [resulting in some (nonzero) extra columns, which we treat in the same manner as those columns which were skipped]. The final number of 1’s [on the main diagonal] is the rank of $A$.

We will call the result of this part the matrix echelon form of the equations.

3. To interpret the answer we do this:
(a) If there are any 'extra' (zero A-part) rows, we check the corresponding b elements. If they are all equal to zero, we delete the extra (redundant) rows and go to the next step; if we find even a single non-zero element among them, the original system of equations is inconsistent, and there is no solution.

(b) Each of the 'skipped' columns represents an unknown whose value can be chosen arbitrarily. Each row then provides an expression for one of the remaining unknowns (in terms of the 'freely chosen' ones). Note that when there are no 'skipped' columns, the solution is just a point in m (number of unknowns) dimensions, one 'skipped' column results in a straight line, two 'skipped' columns in a plane, etc.

Since the first two steps of this procedure are quite straightforward, we give EXAMPLES of the interpretation part only:

1. 
\[
\begin{bmatrix}
1 & 3 & 0 & 2 & 0 & 2 \\
0 & 0 & 1 & 3 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 4 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

means that \(x_2\) and \(x_4\) are the 'free' parameters (often, they would be renamed \(c_1\) and \(c_2\), or \(A\) and \(B\)). The solution can thus be written as
\[
\begin{align*}
x_1 &= 2 - 3x_2 - 2x_4 \\
x_3 &= 1 - 3x_4 \\
x_5 &= 4
\end{align*}
\]
or, in a vector-like manner:
\[
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
x_5 \\
x_6
\end{bmatrix} =
\begin{bmatrix}
2 \\
0 \\
1 \\
0 \\
4 \\
0
\end{bmatrix} +
\begin{bmatrix}
-3 \\
1 \\
0 \\
0 \\
0 \\
0
\end{bmatrix}c_1 +
\begin{bmatrix}
-2 \\
0 \\
-3 \\
1 \\
0 \\
0
\end{bmatrix}c_2
\]

Note that this represents a (unique) plane in a five-dimensional space; the 'point' itself and the two directions (coefficients of \(c_1\) and \(c_2\)) can be specified in infinitely many different (but equivalent) ways.

2. 
\[
\begin{bmatrix}
1 & 3 & 0 & 0 & 0 & -2 & 5 \\
0 & 0 & 1 & 0 & 0 & 3 & 2 \\
0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 1 & 4 & 3
\end{bmatrix} \Rightarrow
\begin{bmatrix}
x_1 \\
x_2 \\
x_3 \\
x_4 \\
x_5 \\
x_6
\end{bmatrix} =
\begin{bmatrix}
5 \\
0 \\
2 \\
0 \\
3 \\
0
\end{bmatrix} +
\begin{bmatrix}
-3 \\
1 \\
0 \\
0 \\
-3 \\
-1
\end{bmatrix}c_1 +
\begin{bmatrix}
2 \\
0 \\
-3 \\
0 \\
-4 \\
1
\end{bmatrix}c_2
\]

Eigenvalues & Eigenvectors

If, for a square \((n \times n)\) matrix \(A\), we can find a non-zero [column] vector \(x\) and a (scalar) number \(\lambda\) such that
\[
Ax = \lambda x
\]
then \( \lambda \) is the matrix’ eigenvalue and \( x \) is its right **eigenvector** (similarly \( y^T A = \lambda y^T \) would define its left **eigenvector** \( y^T \), this time a row vector). This means that we seek a non-zero solutions of

\[
(A - \lambda I) x = 0
\]

which further implies that \( A - \lambda I \) must be **singular**: \( \det(A - \lambda I) = 0 \).

The left hand side of the last equation is an \( n^{th} \)-degree polynomial in \( \lambda \) which has (counting multiplicity) \( n \) [possibly complex] roots. These roots are the **eigenvalues** of \( A \); one can easily see that for each distinct root one can find at least one right (and at least one left) eigenvector, by solving (*) for \( x \) (\( \lambda \) being known now).

It is easy to verify that

\[
\det(\lambda I - A) = \lambda^n - \lambda^{n-1} \cdot \text{Tr}(A) + \lambda^{n-2} \cdot \text{sum of all } 2 \times 2 \text{ major subdeterminants} - \lambda^{n-3} \cdot \text{sum of all } 3 \times 3 \text{ major subdeterminants} + \ldots \pm \det(A)
\]

where \( \text{Tr}(A) \) is the sum of all main-diagonal elements. This is called the **characteristic polynomial** of \( A \), and its roots are the only eigenvalues of \( A \).

**EXAMPLES:**

1. \[
\begin{bmatrix} 2 & 3 \\ 1 & -2 \end{bmatrix}
\]
   has \( \lambda^2 - 0 \cdot \lambda - 7 = 0 \) as its characteristic polynomial, which means that the eigenvalues are \( \lambda_{1,2} = \pm \sqrt{7} \).

2. \[
\begin{bmatrix} 3 & -1 & 2 \\ 0 & 4 & 2 \\ 2 & -1 & 3 \end{bmatrix}
\]
   \( \Rightarrow \lambda^3 - 10\lambda^2 + (12 + 14 + 5)\lambda - 22 \) [we know how to find the determinant]. The coefficients add up to 0. This implies that \( \lambda_1 = 1 \) and \( \lambda_{2,3} = \frac{9}{2} \pm \frac{\sqrt{7}}{2} i \).

3. \[
\begin{bmatrix} 2 & 4 & -2 & 3 \\ 3 & 6 & 1 & 4 \\ -2 & 4 & 0 & 2 \\ 8 & 1 & -2 & 4 \end{bmatrix}
\]
   \( \Rightarrow \lambda^4 - 12\lambda^3 + (0 - 4 + 4 - 4 + 20 - 16)\lambda^2 - (-64 - 15 - 28 - 22)\lambda + (-106) = 0 \) [note there are \( \binom{4}{2} = 6 \) and \( \binom{4}{3} = 4 \) major subdeterminants of the \( 2 \times 2 \) and \( 3 \times 3 \) size, respectively] \( \Rightarrow \lambda_1 = -3.2545, \lambda_2 = 0.88056, \lambda_3 = 3.3576 \) and \( \lambda_4 = 11.0163 \) [these were obtained from our general formula for fourth-degree polynomials – let’s hope we don’t have to use it very often].

The corresponding (right) **eigenvectors** can be now found by solving (*), a **homogenous** set of equations with a **singular** matrix of coefficients [therefore, there must be at least one nonzero solution – which, furthermore, can be multiplied by an arbitrary constant]. The number of **linearly independent** (LI) solutions cannot be bigger than the **multiplicity** of the corresponding eigenvalue; establishing their correct number is an important part of the answer.

**EXAMPLES:**
1. Using \( A = \begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix} \) [one of our previous examples] \((A - \lambda_1 I)x = 0\) amounts to
\[
\begin{pmatrix} 2-\sqrt{7} & 3 \\ 1 & -2-\sqrt{7} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix},
\]
with the second equation being a multiple of the first [check it!]. We thus have to solve only \( x_1 - (2 + \sqrt{7})x_2 = 0 \), which has the following general solution: \( \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2+\sqrt{7} \\ 1 \end{pmatrix} c \), where \( c \) is arbitrary [geometrically, the solution represents a straight line in the \( x_1-x_2 \) plane, passing through the origin]. Any such vector, when pre-multiplied by \( A \), increases in length by a factor of \( \sqrt{7} \), without changing direction (check it too). Similarly, replacing \( \lambda_1 \) by \( \lambda_2 = -\sqrt{7} \), we would be getting \( \begin{pmatrix} 2-\sqrt{7} \\ 1 \end{pmatrix} c \) as the corresponding eigenvector. There are many equivalent ways of expressing it, \( \begin{pmatrix} -3 \\ 2+\sqrt{7} \end{pmatrix} \tilde{c} \) is one of them.

2. A **double eigenvalue** may possess either one or two linearly independent eigenvectors:

   (a) The unit \( 2 \times 2 \) matrix has \( \lambda = 1 \) as its duplicate eigenvalue, \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) and \( \begin{pmatrix} 0 \\ 1 \end{pmatrix} \) are two LI eigenvectors [the general solution to \( \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \)]. This implies that any vector is an eigenvector of the unit matrix.

   (b) The matrix \( \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix} \) has the same duplicate eigenvalue of \(+1\) [in general, the main diagonal elements of an **upper-triangular** matrix are its eigenvalues], but solving \( \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \) i.e. \( 2x_2 = 0 \) has only one LI solution, namely \( \begin{pmatrix} 1 \\ 0 \end{pmatrix} c \).

Finding eigenvectors and eigenvalues of a matrix represents the main step in solving **sets of ODEs**; we will present our further examples in that context. So let us now return to these:

**Set (system) of differential equations**

of **first order**, **linear**, and with **constant coefficients** typically looks like this:

\[
y'_1 = 3y_1 + 4y_2 \\
y'_2 = 3y_1 - y_2
\]

[the example is of the **homogeneous** type, as each term is either \( y_i \) or \( y'_i \) proportional]. The same set can be conveniently expressed in **matrix notation** as

\[
y' = A y
\]

where \( A = \begin{pmatrix} 3 & 4 \\ 3 & -1 \end{pmatrix} \) and \( y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \) [both \( y_1 \) and \( y_2 \) are function of \( x \)].
The Main Technique

for constructing a solution to any such set of \( n \) DEs is very similar to what we have seen in the case of one (linear, constant-coefficient, homogeneous) DE, namely:

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}
\]

We first try to find \( n \) linearly independent basic solutions (all having the form), then build the general solution as a linear combination (with arbitrary coefficients) of these.

It happens that the basic solutions can be constructed with the help of matrix algebra. To find them, we use the following trial solution:

\[ y_T = q \cdot e^{\lambda x} \]

where \( q \) is a constant \((n\text{-dimensional})\) vector. Substituting into \( y' = A y \) and cancelling the (scalar) \( e^{\lambda x} \) gives: \( \lambda q = A q \), which means \( \lambda \) can be any one of the eigenvalues of \( A \) and \( q \) be the corresponding eigenvector. If we find \( n \) of these (which is the case with simple eigenvalues) the job is done; we have effectively constructed a general solution to our set of DEs.

EXAMPLES:

1. Solve \( y' = A y \), where \( A = \begin{bmatrix} 3 & 4 \\ 3 & -1 \end{bmatrix} \). The characteristic equation is: \( \lambda^2 - 2\lambda - 15 = 0 \Rightarrow \lambda_{1,2} = 1 \pm \sqrt{16} = -3 \) and 5. The corresponding eigenvectors (we will call them \( q^{(1)} \) and \( q^{(2)} \)) are the solutions to \( \begin{bmatrix} 6 & 4 & 0 \\ 3 & 2 & 0 \end{bmatrix} \Rightarrow 3q^{(1)}_1 + 2q^{(1)}_2 = 0 \Rightarrow q^{(1)} = \begin{bmatrix} 2 \\ -3 \end{bmatrix} c_1 \), and \( \begin{bmatrix} 2 & 4 \\ 3 & -6 \end{bmatrix} \) \([\text{from now on we will assume a zero right hand side}]\) \( \Rightarrow 3q^{(2)}_1 - 2q^{(2)}_2 = 0 \Rightarrow q^{(2)} = \begin{bmatrix} 2 \\ 1 \end{bmatrix} c_2 \).

The final, general solution is thus \( y = c_1 \begin{bmatrix} 2 \\ -3 \end{bmatrix} e^{-3x} + c_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{5x} \). Or, if you prefer, more explicitly:

\[
\begin{align*}
y_1 &= 2c_1 e^{-3x} + 2c_2 e^{5x} \\
y_2 &= -3c_1 e^{-3x} + c_2 e^{5x}
\end{align*}
\]

where \( c_1 \) and \( c_2 \) can be chosen arbitrarily.

Often, they are specified via initial conditions, e.g. \( y_1(0) = 2 \) and \( y_2(0) = -3 \) imply \( 2c_1 + 2c_2 = 2 \Rightarrow c_1 - 3c_2 = -3 \Rightarrow c_1 = 1 \) and \( c_2 = 0 \) imply \( y_1 = 2e^{-3x} \) and \( y_2 = -3e^{-3x} \).

2. Let us now tackle a three-dimensional problem, with \( A = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 1 & -1 \\ 2 & -1 & 0 \end{bmatrix} \). The characteristic equation is \( \lambda^3 - 2\lambda^2 - \lambda + 2 = 0 \Rightarrow \lambda_1 = -1 \) and the roots of \( \lambda^2 - 3\lambda + 2 = 0 \Rightarrow \lambda_{2,3} = 1 \) and 2. The respective eigenvectors are:
The general solution is thus \( y = c_1 \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix} e^{-x} + c_2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} e^x + c_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} e^{2x} \).

The case of Double (Multiple) Eigenvalue

For each such eigenvalue we must first find all possible solutions of the type \( q e^{\lambda x} \)

\[
(qx + s)e^{\lambda x}
\]

where \( q \) and \( s \) are two constant vectors to be found by substituting this (trial) solution into the basic equation \( y' = Ay \). As a result we get \( q = (A - \lambda I)q + (A - \lambda I)s \) \( \Rightarrow q \) is such a linear combination of the \( q \)-vectors found in the previous step which allows \((A - \lambda I)s = q\) to be solved in terms of \( s \). Note that both the components of \( s \) and the coefficients of the linear combination of the \( q \) vectors are the unknowns of this problem. One thus needs to append all the \( q \) vectors found in the previous step to \((A - \lambda I)\) (as extra columns) and reduces the whole (thus appended) matrix to its echelon form. It is the number of 'skipped' \( q \)-columns which tells us how many distinct solutions there are (the 'skipped' columns \( A - \lambda I \) would be adding, to \( s \), a multiple of the \( q e^{\lambda x} \) solution already constructed). One thus cannot get more solutions than in the previous step.

And, if still not done, we have to proceed to

\[
(q \frac{x^2}{2!} + sx + u)e^{\lambda x}
\]

where \( q \) and \( s \) [in corresponding pairs] is a combination of solutions from the previous step such that \((A - \lambda I)u = s\) can be solved in terms of \( u \). Find how many LI combinations of the \( s \)-vectors allow a nonzero \( u \)-solution [by the 'appending' technique], solve for the corresponding \( u \)-vectors, and if necessary move on to the next \((q \frac{x^3}{3!} + s \frac{x^2}{2!} + ux + w)e^{\lambda x}\) step, until you have as many solutions as the eigenvalue's multiplicity. Note that \((A - \lambda I)^2s = 0, (A - \lambda I)^3u = 0, \ldots\)
EXAMPLES:

1. \( A = \begin{bmatrix} 5 & 2 & 2 \\ 2 & 2 & -4 \\ 2 & -4 & 2 \end{bmatrix} \) has \( \lambda^3 - 9\lambda^2 + 108 \) as its characteristic polynomial [hint: there is a double root] \( \Rightarrow 3\lambda^2 - 18\lambda = 0 \) has two roots, 0 [does not check] and 6 [checks]. Furthermore, \( \lambda^3 - 9\lambda^2 + 108 \div (\lambda - 6)^2 = \lambda + 3 \Rightarrow \) the three eigenvalues are \(-3\) and 6 [duplicate]. Using \( \lambda = -3 \) we get:

\[
\begin{bmatrix} 1 & 0 & \frac{1}{7} \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow c_1 \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix},
\]
which, when multiplied by \( e^{-3x} \), gives the first basic solution. Using \( \lambda = 6 \) yields:

\[
\begin{bmatrix} -1 & 2 & 2 \\ 2 & -4 & -4 \\ 2 & -4 & -4 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \\ 0 \end{bmatrix} \Rightarrow c_2 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix} + c_3 \begin{bmatrix} 2 \\ 0 \\ 1 \end{bmatrix}
\]
which, when multiplied by \( e^{6x} \), supplies the remaining two basic solutions.

2. \( A = \begin{bmatrix} 1 & -3 & 1 \\ 2 & -1 & -2 \\ 2 & -3 & 0 \end{bmatrix} \) has \( \lambda^3 - 3\lambda^2 - 2 \) as its characteristic polynomial, with roots: \( \lambda_1 = -1 \) [one of our rules] \( \Rightarrow (\lambda^3 - 3\lambda^2 - 2) \div (\lambda + 1) = \lambda^2 - \lambda - 2 \Rightarrow \lambda_2 = -1 \) and \( \lambda_3 = 2 \). So again, there is one duplicate root.

For \( \lambda = 2 \) we get:

\[
\begin{bmatrix} -1 & -3 & 1 \\ 2 & -3 & -2 \\ 2 & -3 & -2 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \\ 0 \end{bmatrix} \Rightarrow c_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \Rightarrow e^{2x} \text{ [a single solution only]}
\]
we get:

\[
\begin{bmatrix} 2 & -3 & 1 \\ 2 & 0 & -2 \\ 2 & -3 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow c_2 \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} e^{-x}
\]
The challenge is to construct the other (last) solution. We have to solve

\[
\begin{bmatrix} 2 & -3 & 1 \\ 2 & 0 & -2 \\ 2 & -3 & 1 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}, \text{ getting } s = \begin{bmatrix} \frac{1}{7} \\ 1 \\ 0 \end{bmatrix}
\]
second part just duplicates the previous basic solution and can be discarded.

The third basic solution is thus: \( c_3 \begin{bmatrix} \frac{1}{7} & x \\ 1 & 0 \\ \frac{2}{7} & x \end{bmatrix} e^{-x} \equiv c_3 \begin{bmatrix} x + \frac{2}{7} \\ 1 \\ x \end{bmatrix} e^{-x} \).

3. \( A = \begin{bmatrix} 42 & -9 & 9 \\ -12 & 39 & -9 \\ -28 & 21 & 9 \end{bmatrix} \Rightarrow \lambda^3 - 90\lambda^2 + 2700\lambda - 27000 \) [hint: triple root] \( \Rightarrow 6\lambda - 180 = 0 \) has a single root of 30 [\( \Rightarrow \) triple root of the original polynomial].

Finding eigenvectors:

\[
\begin{bmatrix} 12 & -9 & 9 \\ -12 & 9 & -9 \\ -28 & 21 & -21 \end{bmatrix} \Rightarrow \begin{bmatrix} 1 & -\frac{3}{7} & \frac{3}{7} \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Rightarrow c_1 \begin{bmatrix} 3 \\ 4 \\ -3 \end{bmatrix} \text{ and } c_2 \begin{bmatrix} 0 \\ 0 \\ 4 \end{bmatrix}
\]
are the corresponding eigenvectors [only two] which, when multiplied by \( e^{30x} \), yield the first two basic solutions. To construct the third, we set up the
equations for individual components of \( \mathbf{s} \), and for the \( a_1 \) and \( a_2 \) coefficients of \( \mathbf{q} \equiv a_1 \begin{pmatrix} 3 \\ 4 \\ 0 \end{pmatrix} + a_2 \begin{pmatrix} 0 \\ 4 \\ 0 \end{pmatrix} \), thus:

\[
\begin{pmatrix} -3 & -12 & -9 & 9 & 3 & -3 \\ -12 & 9 & -9 & 4 & 0 \\ -28 & 21 & -21 & 0 & 4 \end{pmatrix}
\]

We bring this to its matrix echelon form:

\[
\begin{pmatrix} 1 & -\frac{3}{7} & \frac{2}{7} & 0 & -\frac{1}{7} \\ 0 & 0 & 0 & 1 & -\frac{3}{7} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}
\]

which first implies that \( a_1 - \frac{3}{7} a_2 = 0 \) ⇔ \( a_1 = 3c_3 \), \( a_2 = 7c_3 \). This results in

\[
\begin{pmatrix} 3 & 4 & 0 \end{pmatrix} + \begin{pmatrix} -3 \\ 0 \\ 4 \end{pmatrix} x + \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} e^{30x} = c_3 \begin{pmatrix} -12x - 1 \\ 12x \\ 28x \end{pmatrix} e^{30x}.
\]

The second basic solution is thus \( c_2 (\mathbf{q} x + \mathbf{s}) e^{-55x} = c_2 \begin{pmatrix} x - \frac{9}{220} \\ -4x + \frac{1}{55} \\ -4x \end{pmatrix} e^{-55x} \). Finally, and the corresponding third basic solution:

\[
\begin{pmatrix} 3 & \mathbf{q}^2 \mathbf{s} + \mathbf{u} \end{pmatrix} e^{-55x} = c_3 \begin{pmatrix} \frac{x^2}{2} - \frac{9}{220} x - \frac{131}{12100} \\ \frac{x^2}{2} + \frac{1}{55} + \frac{39}{12100} \\ -2x^2 \end{pmatrix} e^{-55x}.
\]

To deal with \( \text{Complex Eigenvalues/Vectors} \):

we first write the corresponding solution in a complex form, using the regular procedure. We then replace each conjugate pair of basic solutions by the real and imaginary part (of either solution).

**EXAMPLE:**
\[ y' = \begin{pmatrix} 2 & -1 \\ 3 & 4 \end{pmatrix} y \Rightarrow \lambda^2 - 6\lambda + 11 \Rightarrow \lambda_{1,2} = 3 \pm \sqrt{2}i \Rightarrow \begin{pmatrix} -1-\sqrt{2}i \\ 3 \end{pmatrix} \begin{pmatrix} 1-\sqrt{2}i \\ -3 \end{pmatrix} \]

is the eigenvector corresponding to \( \lambda_1 = 3 + \sqrt{2}i \) [its complex conjugate corresponds to \( \lambda_2 = 3 - \sqrt{2}i \)]. This means that the two basic solutions (in their complex form) are \( e^{(3+\sqrt{2}i)x} \) and its complex conjugate \( \rightarrow-i \). Equivalently, we can use the real and imaginary part of either of these [up to a sign, the same answer] to get:

\[
y = c_1 \begin{pmatrix} \cos(\sqrt{2}x) + \sqrt{2}\sin(\sqrt{2}x) \\ -3\cos(\sqrt{2}x) \end{pmatrix} e^{3x} + c_2 \begin{pmatrix} -\sqrt{2}\cos(\sqrt{2}x) + \sin(\sqrt{2}x) \\ -3\sin(\sqrt{2}x) \end{pmatrix} e^{3x}. \]

This is the fully general, real solution to the original set of DEs.

Now, we extend our results to the

Non-homogeneous case

of

\[ y' - \lambda y = r(x) \]

where \( r \) is a given vector function of \( x \) (effectively \( n \) functions, one for each equation). We already know how to solve the corresponding homogeneous version.

There are two techniques to find a particular solution \( y^{(p)} \) to the complete equation; the general solution is then constructed in the usual manner.

The first of these techniques (for constructing \( y_p \)) is:

\[ \text{Variation of Parameters} \]

As a trial solution, we use

\[ y^{(T)} = Y \cdot c \]

where \( Y \) is an \( n \) by \( n \) matrix of functions, with the \( n \) basic solutions of the homogeneous equation comprising its individual columns:

\[ Y \equiv \begin{pmatrix} y^{(1)} & y^{(2)} & \ldots & y^{(n)} \end{pmatrix} \]

and \( c \) being a single column of the \( c_i \) coefficients: \( c \equiv \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \), each now considered a function of \( x \) [\( Y \cdot c \) is just a matrix representation of \( c_1 y^{(1)} + c_2 y^{(2)} + \ldots + c_n y^{(n)} \), with the \( c_i \) coefficients now being 'variable'].


Substituting in the full (non-homogeneous) equation, and realizing that \( Y' \equiv A \cdot Y \) [\( Y' \) represents differentiating, individually, every element of \( Y \)] we obtain: 
\[
Y' + Y \cdot c' - A \cdot Y \cdot c = r \Rightarrow c' = Y^{-1} \cdot r .
\]
Integrating the right hand side (component by component) yields \( c \). The general solution is thus
\[
y^{(p)} = Y \int Y^{-1} \cdot r(x) \, dx
\]
EXAMPLE:
\[
y' = \begin{bmatrix} 3 & 2 \\ 1 & 2 \end{bmatrix} y + \begin{bmatrix} 4e^{3x} \\ 0 \end{bmatrix} \Rightarrow \lambda^2 - 5\lambda + 4 = 0 \Rightarrow \lambda_{1,2} = 1 \text{ and } 4 \text{ with the respective}
\]
eigenvectors [easy to construct]: \( \begin{bmatrix} 1 \\ -1 \end{bmatrix} \text{ and } \begin{bmatrix} 2 \\ 1 \end{bmatrix} \). Thus \( Y = \begin{bmatrix} e^x & 2e^{4x} \\ -e^x & e^{4x} \end{bmatrix} \Rightarrow \)
\[
Y^{-1} = \begin{bmatrix} \frac{1}{3}e^{-x} & -\frac{2}{3}e^{-x} \\ \frac{1}{3}e^{-4x} & \frac{2}{3}e^{-4x} \end{bmatrix} \text{ This matrix, multiplied by } r(x), \text{ yields } \begin{bmatrix} \frac{1}{3}e^{4x} \\ \frac{1}{3}e^x \end{bmatrix} . \text{ The componentwise integration of the last vector is trivial [the usual additive}
\]
constants can be omitted to avoid duplication]: \( \begin{bmatrix} \frac{1}{3}e^{4x} \\ \frac{1}{3}e^x \end{bmatrix} \), (pre)multiplied by \( Y \) finally results in: \( y^{(p)} = \begin{bmatrix} 3e^{3x} \\ e^{5x} \end{bmatrix} \). The general solution is thus \( y = c_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} \)
e\]
\( e^x + c_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix} e^{4x} + \begin{bmatrix} 3e^{3x} \\ e^{5x} \end{bmatrix} \). Let us make this into an initial-value problem: \( y_1(0) = 1 \text{ and } y_2(0) = -1 \Leftrightarrow \)
y\( (0) = Y(0)c + y^{(p)}(0) = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \Rightarrow \)
y\( (0) = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix} \begin{bmatrix} 1 \\ -3 \end{bmatrix} = \begin{bmatrix} \frac{7}{3} \\ -\frac{2}{3} \end{bmatrix} \Rightarrow y = \begin{bmatrix} \frac{7}{3} \\ -\frac{2}{3} \end{bmatrix} e^x + \begin{bmatrix} \frac{7}{3} \\ \frac{2}{3} \end{bmatrix} e^{4x} + \begin{bmatrix} 3e^{3x} \\ e^{5x} \end{bmatrix} \).  

The second technique for building \( y_p \) works only for two

\( \Rightarrow \) Special Cases of \( r(x) \)

\( \Rightarrow \) When the non-homogeneous part of the equation has the form of
\( \left( a_kx^k + a_{k-1}x^{k-1} + \ldots + a_1x + a_0 \right) e^{\beta x} \)
we use the following 'trial' solution (which is guaranteed to work) to construct \( y^{(p)} \):
\( \left( b_mx^m + b_{m-1}x^{m-1} + \ldots + b_1x + b_0 \right) e^{\beta x} \)
where \( m \) equals \( k \) plus the multiplicity of \( \beta \) as an eigenvalue of \( A \) (if \( \beta \) is not an

eigenvalue, \( m = k \), if it is a simple eigenvalue, \( m = k + 1 \), etc.).

When \( \beta \) does not coincide with any eigenvalue of \( A \), the equations to solve to obtain \( b_k, b_{k-1}, \ldots, b_1 \) are
\[
\begin{align*}
(A - \beta I) b_k &= -a_k \\
(A - \beta I) b_{k-1} &= k b_k - a_{k-1} \\
(A - \beta I) b_{k-2} &= (k-1)b_{k-1} - a_{k-2} \\
& \quad \vdots \\
(A - \beta I) b_0 &= b_1 - a_0
\end{align*}
\]
Since \((A - \beta I)\) is a regular matrix (having an inverse), solving these is quite routine (as long as we start from the top).

When \(\beta\) coincides with a simple (as opposed to multiple) eigenvalue of \(A\), we have to solve

\[
\begin{align*}
(A - \beta I) b_{k+1} &= 0 \\
(A - \beta I) b_k &= (k + 1)b_{k+1} - a_k \\
(A - \beta I) b_{k-1} &= kb_k - a_{k-1} \\
&\vdots \\
(A - \beta I) b_0 &= b_1 - a_0
\end{align*}
\]

Thus, \(b_{k+1}\) must be the corresponding eigenvector, multiplied by such a constant as to make the second equation solvable [remember that now \((A - \beta I)\) is singular]. Similarly, when solving the second equation for \(b_k\), a \(c\)-multiple of the same eigenvector must be added to the solution, with \(c\) chosen so that the third equation is solvable, etc. Each \(b_i\) is thus unique, even though finding it is rather tricky.

We will not try extending this procedure to the case of \(\beta\) being a double (or multiple) eigenvalue of \(A\).

\(\triangleright\) On the other hand, the extension to the case of \(r(x) = P(x)e^{px}\cos(qx) + Q(x)e^{px}\sin(qx)\), where \(P(x)\) and \(Q(x)\) are polynomials in \(x\) (with vector coefficients), and \(p + iq\) is not an eigenvalue of \(A\) is quite simple: The trial solution has the same form as \(r(x)\), except that the two polynomials will have undetermined coefficients, and will be of the same degree (equal to the degree of \(P(x)\) or \(Q(x)\), whichever is larger). This trial solution is then substituted into the full equation, and the coefficients of each power of \(x\) are matched, separately for the \(\cos(qx)\)-proportional and \(\sin(qx)\)-proportional terms.

In addition, one can also use the superposition principle [i.e. dividing \(r(x)\) into two or more manageable parts, getting a particular solution for each part separately, and then adding them all up].

EXAMPLES:

1. \(y' = \begin{bmatrix} -4 & -4 \\ 1 & 2 \end{bmatrix} y + \begin{bmatrix} 1 \\ 0 \end{bmatrix} e^{2x} + \begin{bmatrix} 0 \\ -2 \end{bmatrix} e^{-x} \Rightarrow \lambda^2 + 2\lambda - 4 = 0 \Rightarrow \lambda_{1,2} = -1 \pm \sqrt{5}.

   We already know how to construct the solution to the homogeneous part of the equation, we show only how to deal with \(y^{(p)} = y^{(p_1)} + y^{(p_2)}\) [for each of the two \(r(x)\) terms]:

   \[y^{(p_1)} = be^{2x},\] substituted back into the equation gives \[\begin{bmatrix} -6 & -4 \\ 1 & 0 \end{bmatrix} b = \begin{bmatrix} -1 \\ 0 \end{bmatrix} \Rightarrow b = \begin{bmatrix} 0 \\ \frac{1}{2} \end{bmatrix}\]

   Similarly \(y^{(p_2)} = be^{-x}\) [a different \(b\)], substituted, gives \[\begin{bmatrix} -3 & -4 \\ 1 & 3 \end{bmatrix} b = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \Rightarrow b = \begin{bmatrix} \frac{8}{5} \\ \frac{6}{5} \end{bmatrix}\]
The full particular solution is thus $y(p) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} e^{2x} + \begin{pmatrix} \frac{7}{5} \\ -\frac{6}{5} \end{pmatrix} e^{-x}$.

2. $y' = \begin{pmatrix} -1 & 2 & 3 \\ 5 & -1 & -2 \\ 5 & 3 & 3 \end{pmatrix} y + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} e^x + \begin{pmatrix} 4 \\ 0 \\ 0 \end{pmatrix} \Rightarrow \lambda^3 - \lambda^2 - 24\lambda - 7 = 0$. If we are interested in the particular solution only, we need to check that neither $\beta = 1$ nor $\beta = 0$ are the roots of the characteristic polynomial [true].

Thus $y(p_1) = b e^x$ where $b$ solves $\begin{pmatrix} -2 & 2 & 3 \\ 5 & -2 & -2 \\ 5 & 3 & 2 \end{pmatrix} \begin{pmatrix} b \\ b \\ b \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \Rightarrow b = \begin{pmatrix} -\frac{2}{5} \\ \frac{20}{31} \\ -\frac{20}{31} \end{pmatrix}$.

Similarly $y(p_2) = b$ where $b = \begin{pmatrix} -4 \\ 0 \\ 0 \end{pmatrix} \Rightarrow b = \begin{pmatrix} -\frac{7}{72} \\ \frac{7}{72} \\ \frac{7}{72} \end{pmatrix}$.

Answer: $y(p) = \begin{pmatrix} -\frac{2}{5} \\ \frac{20}{31} \\ -\frac{20}{31} \end{pmatrix} e^x + \begin{pmatrix} -\frac{7}{72} \\ \frac{7}{72} \\ \frac{7}{72} \end{pmatrix}$.

3. $y' = \begin{pmatrix} -1 & 2 & 3 \\ 5 & -1 & -2 \\ 5 & 3 & 3 \end{pmatrix} y + \begin{pmatrix} x - 1 \\ 2 \\ -2x \end{pmatrix}$ [characteristic polynomial same as previous example]. $y(p) = b_1 x + b_0$ with $\begin{pmatrix} -1 & 2 & 3 \\ 5 & -1 & -2 \\ 5 & 3 & 3 \end{pmatrix} \begin{pmatrix} b_1 \\ b_0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \end{pmatrix} \Rightarrow b_1 = \begin{pmatrix} -\frac{3}{5} \\ \frac{2}{5} \\ \frac{2}{5} \end{pmatrix}$ and $\begin{pmatrix} -1 & 2 & 3 \\ 5 & -1 & -2 \\ 5 & 3 & 3 \end{pmatrix} \begin{pmatrix} b_0 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix} \Rightarrow b_0 = \begin{pmatrix} -\frac{150}{7} \\ \frac{60}{7} \\ \frac{60}{7} \end{pmatrix}$. Thus $y(p) = \begin{pmatrix} -\frac{3}{5}x + \frac{100}{7} \\ \frac{2}{5}x - \frac{120}{7} \\ -\frac{2}{5}x + \frac{60}{7} \end{pmatrix}$.

4. $y' = \begin{pmatrix} -4 & -3 \\ 2 & 1 \end{pmatrix} y + \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{-x} \Rightarrow \lambda^2 + 3\lambda + 2 = 0 \Rightarrow \lambda_{1,2} = -1, -2$. Now our $\beta = -1$ 'coincides' with a simple eigenvalue.

$y(p) = (b_1 x + b_0) e^{-x}$ where $\begin{pmatrix} -3 & -3 \\ 2 & 2 \end{pmatrix} b_1 = 0 \Rightarrow$

$b_1 = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} -3 & -3 \\ 2 & 2 \end{pmatrix} b_0 = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -3 & -3 & 1 & -1 \\ 2 & 2 & -1 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$.

This fixes the value of $c_1$ at $-8 \Rightarrow b_1 = \begin{pmatrix} -8 \\ 8 \end{pmatrix}$ and $b_0 = \begin{pmatrix} 3 \\ 0 \end{pmatrix} + c_0 \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Being the last $b$, we can set $c_0 = 0$ (not to duplicate the homogeneous part of the solution).

Answer: $y(p) = \begin{pmatrix} -8x + 3 \\ 8x \end{pmatrix} e^{-x}$. 


5. \( y' = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 4 & -1 \\ -5 & -8 & -3 \end{pmatrix} y + \begin{pmatrix} x \\ 0 \\ 4 \end{pmatrix} \Rightarrow \lambda^3 - 2\lambda^2 - 15\lambda = 0 \Rightarrow \beta = 0 \) is a simple eigenvalue.

We construct \( y(p) = b_2x^2 + b_1x + b_0 \) where

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 4 & -1 \\
-5 & -8 & -3 
\end{bmatrix} b_2 = 0 \Rightarrow b_2 = \begin{bmatrix}
-5 \\
2 \\
3 
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 4 & -1 \\
-5 & -8 & -3 
\end{bmatrix} b_1 = 2b_2 - 0 \Rightarrow b_1 = \begin{bmatrix}
10 & 10 & -3 \\
-5 & -8 & 3 \\
6 & 0 
\end{bmatrix} + c_2 = \begin{bmatrix}
\frac{28}{15} \\
\frac{13}{15} \\
\frac{1}{15} 
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 4 & -1 \\
-5 & -8 & -3 
\end{bmatrix} b_0 = b_1 - 0 \Rightarrow b_0 = \begin{bmatrix}
12 & 12 & -3 \\
-5 & -8 & 3 \\
3 & -20 & 15 
\end{bmatrix} + c_1 = \begin{bmatrix}
\frac{1219}{15} \\
\frac{672}{15} \\
\frac{672}{15} 
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 4 & -1 \\
-5 & -8 & -3 
\end{bmatrix} b_0 = b_1 - 0 \Rightarrow b_0 = \begin{bmatrix}
12 & 12 & -3 \\
-5 & -8 & 3 \\
3 & -20 & 15 
\end{bmatrix} + c_1 = \begin{bmatrix}
\frac{1219}{15} \\
\frac{672}{15} \\
\frac{672}{15} 
\end{bmatrix}
\]

\[
\begin{bmatrix}
1 & 1 & 1 \\
1 & 4 & -1 \\
-5 & -8 & -3 
\end{bmatrix} b_0 = b_1 - 0 \Rightarrow b_0 = \begin{bmatrix}
12 & 12 & -3 \\
-5 & -8 & 3 \\
3 & -20 & 15 
\end{bmatrix} + c_1 = \begin{bmatrix}
\frac{1219}{15} \\
\frac{672}{15} \\
\frac{672}{15} 
\end{bmatrix}
\]

Answer: \( y(p) = -5x^2 + \frac{28}{15}x + \frac{1219}{15} \\
-\frac{13}{15}x^2 - \frac{119}{225}x - \frac{672}{15} \\
-\frac{1}{15}x^2 - \frac{4}{225}x 
\]

Two Final Remarks

Note that an equation of the type

\[ B y' = A y + r \]

can be converted to the regular type by (pre)multiplying it by \( B^{-1} \).

EXAMPLE:

\[
2y_1' - 3y_2' = y_1 - 2y_2 + x \quad \iff \quad y_1' = \frac{7}{5} y_1 + \frac{7}{5} y_2 + \frac{7}{5} x - \frac{12}{5} \\
y_2' = \frac{5}{3} y_1 + \frac{5}{3} y_2 - \frac{5}{3} x - \frac{8}{5}
\]

which we know how to solve.

And, finally: Many important systems of differential equations you encounter in Physics are nonlinear, of second order. A good example is: \( \ddot{r} + \mu \frac{\dot{r}^3}{r^2} = 0 \) (\( \mu \) is a constant) describing motion of a planet around the sun. Note that \( r \) actually stands for three dependent variables, \( x(t) \), \( y(t) \) and \( z(t) \). There is no general method for solving such equations, specialized techniques have to be developed for each particular case. We cannot pursue this topic here.
Chapter 6  POWER-SERIES SOLUTION

of the following equation

\[ y'' + f(x)y' + g(x)y = 0 \]  \hspace{1cm} (MAIN)

So far we have not discovered any general procedure for solving such equations. The technique we develop in this chapter can do the trick, but it provides the Maclaurin expansion of the solution only (which, as we will see, is quite often sufficient to identify the function).

We already know that once the solution to a homogeneous equation is found, the V of P technique can easily deal with the corresponding nonhomogeneous case (e.g. any \( r(x) \) on the right hand side of MAIN). This is why, in this chapter, we restrict our attention to homogeneous equations only.

The main idea is to express \( y \) as a power series in \( x \), with yet to be determined coefficients:

\[ y(x) = \sum_{i=0}^{\infty} c_i x^i \]

then substitute this expression into the differential equation to be solved [\( y' = \sum_{i=1}^{\infty} i c_i x^{i-1} \) needs to be multiplied by \( f(x) \) and \( y'' = \sum_{i=2}^{\infty} i(i-1)c_i x^{i-2} \) by \( g(x) \), where both \( f(x) \) and \( g(x) \) must be expanded in the same manner – they are usually in that form already] and make the overall coefficient of each power of \( x \) is equal to zero.

This results in (infinitely many, but regular) equations for the unknown coefficients \( c_0, c_1, c_2, \ldots \). These can be solved in a recurrent [some call it recursive] manner (i.e. by deriving a simple formula which computes \( c_k \) based on \( c_0, c_1, \ldots, c_{k-1} \); \( c_0 \) and \( c_1 \) can normally be chosen arbitrarily).

EXAMPLES:

1. \( y'' + y = 0 \Rightarrow \sum_{i=2}^{\infty} i(i-1)c_i x^{i-2} + \sum_{i=0}^{\infty} c_i x^i \equiv 0 \). The main thing is to express the left hand side as a single infinite summation, by replacing the index \( i \) of the first term by \( i^* + 2 \), thus: \( \sum_{i^*-0}^{\infty} (i^* + 2)(i^* + 1)c_{i^*+2} x^{i^*} \) [note that the lower limit had to be adjusted accordingly]. But \( i^* \) is just a dummy index which can be called \( j, k \) or anything else including \( i \). This way we get (combining both terms): \( \sum_{i=0}^{\infty} [(i+2)(i+1)c_{i+2} + c_i] x^i \equiv 0 \) which implies that the expression in square brackets must be identically equal to zero. This yields the following recurrent formula

\[ c_{i+2} = \frac{-c_i}{(i+2)(i+1)} \]

where \( i = 0, 1, 2, \ldots \), from which we can easily construct the compete sequence of the \( c \)-coefficients, as follows: Starting with \( c_0 \) arbitrary, we get
\[ c_2 = \frac{-c_0}{2}, \quad c_4 = \frac{-c_2}{4!}, \quad c_6 = \frac{-c_4}{6!}, \quad \ldots, \quad c_{2k} = \frac{(-1)^k}{(2k)!} \text{ in general. Similarly, choosing an arbitrary value for } c_1 \text{ we get } c_3 = \frac{-c_1}{3!}, \quad c_5 = \frac{c_3}{5!}, \quad \ldots \times \text{ The complete solution is thus} \]

\[ y = c_0 (1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots) + c_1 (x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots) \]

where the infinite expansions can be easily identified as those of \( \cos x \) and \( \sin x \), respectively. We have thus obtained the expected \( y = c_0 \cos x + c_1 \sin x \) [check].

We will not always be lucky enough to identify each solution as a combination of simple functions, but do learn to recognize at least the following expansions:

\[ (1 - ax)^{-1} = 1 + ax + a^2x^2 + a^3x^3 + \ldots \]
\[ e^{ax} = 1 + ax + \frac{a^2x^2}{2!} + \frac{a^3x^3}{3!} + \ldots \]
\[ \ln(1 - ax) = -ax - \frac{a^2x^2}{2} - \frac{a^3x^3}{3} - \ldots \quad \text{(no factorials)} \]

with \( a \) being any number \( \lceil \text{often } a = 1 \rceil \).

And realize that \( 1 - \frac{x^2}{3!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots \) [a power of \( x \) missing] must be \( \sin x \), \( 1 - \frac{3x^2}{2!} + \frac{9x^4}{4!} - \frac{27x^6}{6!} + \ldots \) is the expansion of \( \cos(\sqrt{3}x) \), \( 1 + x^2 + \frac{x^4}{2!} + \frac{x^6}{4!} + \frac{x^8}{6!} + \ldots \) must be \( \exp(x^2) \), and \( 1 - \frac{x}{\sqrt{3}} + \frac{x^2}{3!} - \frac{x^3}{7!} + \ldots \) is \( \sin \frac{\sqrt{x}}{\sqrt{3}} \).

2. \( (1 - x^2)y'' - 2xy' + 2y = 0 \Rightarrow \sum \limits_{i=2}^{\infty} i(i - 1)c_i x^{i - 2} - \sum \limits_{i=2}^{\infty} i(i - 1)c_i x^i - 2 \sum \limits_{i=1}^{\infty} ic_i x^i \]
\[ 2 \sum \limits_{i=0}^{\infty} c_i x^i \equiv 0. \text{ By reindexing (to get the same } x^i \text{ in each term) we get } \sum \limits_{i=0}^{\infty} (i^* + 2)(i^* + 1)c_{i^* + 2} x^{i^*} - \sum \limits_{i=2}^{\infty} i(i - 1)c_i x^i - 2 \sum \limits_{i=1}^{\infty} ic_i x^i + 2 \sum \limits_{i=0}^{\infty} c_i x^i \equiv 0. \]

Realizing that, as a dummy index, \( i^* \) can be called \( i \) (this is the last time we introduced \( i^* \), form now on we will call it \( i \) directly), our equation becomes:

\[ \sum \limits_{i=0}^{\infty} [(i + 2)(i + 1)c_{i + 2} - i(i - 1)c_i - 2ic_i + 2c_i] x^i \equiv 0 \]

[we have adjusted the lower limit of the second and third term down to 0 without affecting the answer – careful with this though, things are not always that simple]. The square brackets must be identically equal to zero which implies:

\[ c_{i + 2} = \frac{i^2 + i - 2}{(i + 2)(i + 1)} c_i = \frac{i - 1}{i + 1} c_i \]

where \( i = 0, 1, 2, \ldots \) Starting with an arbitrary \( c_0 \) we get \( c_2 = -c_0, \quad c_4 = \frac{1}{3} c_2 = -\frac{1}{3} c_0, \quad c_6 = \frac{3}{5} c_4 = -\frac{3}{5} c_0, \quad c_8 = -\frac{1}{7} c_0, \quad \ldots \) Starting with \( c_1 \) we get \( c_3 = 0, \quad c_5 = 0, \quad c_7 = 0, \quad \ldots \)

The solution is thus

\[ y = c_0 (1 - x^2 - \frac{x^4}{3} - \frac{x^6}{5} - \frac{x^8}{7} - \ldots) + c_1 x \]
One of the basic solutions is thus simply equal to $x$, once we know that we can use the V of P technique to get an analytic expression for the other solution, thus: $y'(x) = c(x) \cdot x$ substituted into the original equation gives:

$$c'' x(1-x^2) + 2c' (1-2x^2) = 0$$

With $c' = z$ this gives $rac{4c}{x} = -2 \left( \frac{1}{x(1-x^2)} \right) = -2 \left[ \frac{A}{x} + \frac{B}{1-x} + \frac{C}{1+x} \right]$ (partial fractions). To solve for $A$, $B$ and $C$ we substitute 0, 1 and -1 into $A(1-x^2) + Bx(1+x) + Cx(1-x) = 0 \Rightarrow A = 1$, $B = -\frac{1}{2}$ and $C = \frac{1}{2}$.

Thus $\ln z = -2 \left[ \ln x + \frac{1}{2} \ln(1-x) + \frac{1}{2} \ln(1+x) \right] + \hat{c} \Rightarrow z = \frac{\hat{c}}{x^2(1-x^2)}$.

Thus $c(x) = \hat{c} \left( -\frac{1}{2} + \frac{1}{2} \ln \frac{1+x}{1-x} \right) + c_1 \Rightarrow y' = c_0 \left( 1 - \frac{x}{2} \ln \frac{1+x}{1-x} \right) + c_1 x$. One can easily verify that this agrees with the previous expansion.

3. $y'' - 3y' + 2y = 0 \Rightarrow \sum_{i=2}^{\infty} i (i-1) c_i x^{i-2} - 3 \sum_{i=1}^{\infty} i c_i x^{i-1} + 2 \sum_{i=0}^{\infty} c_i x^i \equiv 0 \Rightarrow \sum_{i=0}^{\infty} \left[ (i+2)(i+1) c_{i+2} - 3(i+1) c_{i+1} + 2c_i \right] x^i \equiv 0 \Rightarrow c_{i+2} = \frac{3(i+1) c_{i+1} - 2c_i}{(i+2)(i+1)}$. By choosing $c_0 = 1$ and $c_1 = 0$ we can generate the first basic solution $[c_2 = \frac{3x^2-2x}{2} = -1, c_3 = \frac{3x^2(-1)-2x}{3x^2} = -1, ...]$:

$$c_0(1-x^2 - x^3 - \frac{7}{12} x^4 - \frac{1}{4} x^5 - ...)$$

similarly with $c_0 = 0$ and $c_1 = 1$ the second basic solution is:

$$c_1(x + \frac{3}{2} x^2 + \frac{7}{6} x^3 + \frac{5}{8} x^4 + \frac{31}{120} x^5 + ...)$$

There is no obvious pattern to either sequence of coefficients. Yet we know that, in this case, the two basic solutions should be simply $e^x$ and $e^{2x}$. The trouble is that our power-series technique presents these in a hopelessly entangled form of $2e^x - e^{2x}$ [our first basic solution] and $e^{2x} - e^x$ [the second], and we have no way of properly separating them.

Sometimes the initial conditions may help, e.g. $y(0) = 1$ and $y'(0) = 1$ [these are effectively the values of $c_0$ and $c_1$, respectively], leading to $c_2 = \frac{3-2}{2} = \frac{1}{2}$, $c_3 = \frac{3-2}{3x^2} = \frac{1}{6}$, $c_4 = \frac{3-1}{4x^3} = \frac{1}{24}$, ... form which the pattern of the $e^{2x}$-expansion clearly emerges. We can then conjecture that $c_i = \frac{1}{i!}$ and prove it by substituting into $(i+2)(i+1) c_{i+2} - 3(i+1) c_{i+1} + 2c_i = (i+2)(i+1) \frac{1}{(i+2)!} - 3(i+1) \frac{1}{(i+1)!} + 2 \frac{1}{i!} = \frac{1-3+2}{i!} = 0$. Similarly, the initial values of $y(0) = c_0 = 1$ and $y'(0) = c_1 = 2$ will lead to $1 + 2x + \frac{(2x)^2}{2} + \frac{(2x)^3}{3!} + ...$ [the expansion of $e^{2x}$]. Prove that $c_i = \frac{2^i}{i!}$ is also a solution of our recurrence equation!

In a case like this, I often choose such 'helpful' initial conditions; if not, you would be asked to present the first five nonzero terms of each basic (entangled) solution only (without identifying the function). ■
Sturm-Liouville eigenvalue problem

The basic equation of this chapter, namely \( y'' + f(x)y' + g(x)y = 0 \) can be always rewritten in the form of

\[
[p(x)y']' + q(x)y = 0
\]

where \( p(x) = e \int f(x) \, dx \) and \( q(x) = g(x) \cdot p(x) \) [verify]. In Physics, we often need to solve such an equation under the following simple conditions of \( y(x_1) = y(x_2) = 0 \).

A trivial solution to this boundary-value problem is always \( y \equiv 0 \), the real task is to find non-zero solutions, if possible.

Such nonzero solutions will be there, but only under rather special circumstances. To give ourselves more flexibility (and more chances of finding a nonzero solution), we multiply the \( q \)-function by an arbitrary number, say \(-\lambda\), to get:

\[
(py'')' + qy = 0
\]

One can easily verify that \( \frac{(py'')'}{q} \equiv \mathcal{A}y \) defines a linear operator [an operator is a 'prescription' for modifying a function into another function; linear means \( \mathcal{A} (c_1 y_1 + c_2 y_2) = c_1 \mathcal{A} y_1 + c_2 \mathcal{A} y_2 \), where \( c_1 \) and \( c_2 \) are arbitrary constants].

The situation is now quite similar to our old eigenvalue problem of \( \mathcal{A}y = \lambda y \), where \( \mathcal{A} \) is also a linear operator, but acting on vectors rather than functions. The analogy is quite appropriate: similarly to the matrix/vector case we can find a nonzero solution to (S-L) only for some values of \( \lambda \), which will also be called the problem’s eigenvalues [the corresponding \( y \)'s will be called eigenfunctions]. The only major difference is that now we normally find infinitely many of these.

One can also prove that, for two distinct eigenvalues, the corresponding eigenfunctions are orthogonal, in the following sense:

\[
\int_{x_1}^{x_2} q(x) \cdot y_1(x) \cdot y_2(x) \, dx = 0
\]

Proof: By our assumptions:

\[
\begin{align*}
(p y_1')' &= \lambda_1 y_1 q \\
(p y_2')' &= \lambda_2 y_2 q
\end{align*}
\]

Multiply the first equation by \( y_2 \) and the second one by \( y_1 \) and subtract, to get: \( y_2 (py_1')' - y_1 (py_2')' = (\lambda_1 - \lambda_2) q y_1 y_2 \). Integrate this (the left hand side by parts) from \( x_1 \) to \( x_2 \):

\[
y_2 py_1'|_{x_1}^{x_2} - y_1 py_2'|_{x_1}^{x_2} = (\lambda_1 - \lambda_2) \int_{x_1}^{x_2} q y_1 y_2 \, dx
\]

The left hand side is zero due to our boundary conditions, which implies the rest. \( \square \)

Notes:

• When \( p(x_1) = 0 \), we can drop the initial condition \( y(x_1) = 0 \) [same with \( p(x_2) = 0 \)].
• We must always insist that each solution be integrable in the $\int_{x_1}^{x_2} q y^2 \, dx$ sense [to have a true eigenvalue problem]. From now on, we allow only such integrable functions as solutions to a S-L problem, without saying.

• Essentially the same proof would hold for a slightly more complicated equation

$$(py')' + r y = \lambda q y$$

where $r$ is yet another specific function of $x$ (we are running out of letters – nothing to do with the old nonhomogeneous term, also denoted $r$). ■

EXAMPLES:

1. $[(1 - x^2)y']' + \lambda y = 0$ is a so-called LEGENDRE equation. Its (integrable, between $x_1 = -1$ and $x_2 = 1$) solutions must meet $\int_{-1}^{1} y_1(x) \cdot y_2(x) \, dx = 0$ [since $1 - x^2 = 0$ at each $x_1$ and $x_2$, we don’t need to impose any boundary conditions on $y$].

2. $[\sqrt{1 - x^2}y']' + \frac{\lambda y}{\sqrt{1 - x^2}} = 0$ is the CHEBYSHEV equation. The solutions meet $\int_{-1}^{1} \frac{y_1(x) \cdot y_2(x)}{\sqrt{1 - x^2}} \, dx = 0$ [no boundary conditions necessary].

3. $[xe^{-x}y']' + \lambda e^{-x}y = 0$ is the LAGUERRE equation ($x_1 = 0, x_2 = \infty$). The solutions are orthogonal in the $\int_{0}^{\infty} e^{-x} y_1 \cdot y_2 \, dx = 0$ sense [no boundary conditions necessary]. ⊗

Using our power-series technique, we are able to solve the above equations (and, consequently, the corresponding eigenvalue problem). Let us start with the

---

**Legendre Equation**

$$(1 - x^2)y'' - 2xy' + \lambda y = 0$$

(Note that we already solved this equation with $\lambda = 2$, see Example 2 of the ‘main idea’ section).

The expression to be identically equal to zero is $$(i + 2)(i + 1)c_{i+2} - i(i - 1)c_i - 2ic_i + \lambda c_i \Rightarrow c_{i+2} = -\frac{\lambda - (i + 1)i}{(i + 2)(i + 1)}c_i.$$ If we allow the $c_i$-sequence to be infinite, the corresponding function is not integrable in the $\int_{-1}^{1} y^2 \, dx$ sense [we skip showing that, they do it in Physics], that is why we have to insist on finite, i.e. polynomial solution. This can be arranged only if the numerator of the $c_{i+2} = ...$ formula is zero for some integer value of $i$, i.e. iff

$$\lambda = (n + 1)n$$
[these are the eigenvalues of the corresponding S-L problem].

We then get the following polynomial solutions – \( P_n(x) \) being the standard notation:

\[
P_0(x) \equiv 1 \\
P_1(x) = x \\
P_2(x) = 1 - 3x^2 \\
P_3(x) = x - \frac{5}{3}x^3 \\
P_4(x) = 1 - 10x^2 + \frac{35}{3}x^4 \\
\]

i.e., in general, \( P_n(x) = 1 - \frac{n\cdot(n+1)}{2!}x^2 + \frac{(n-2)n\cdot(n+1)(n+3)}{4!}x^4 - \frac{(n-4)n\cdot(n-2)n\cdot(n+1)(n+3)(n+5)}{6!}x^6 + \ldots \) when \( n \) is even, and \( P_n(x) = x - \frac{(n-1)n\cdot(n+2)}{3!}x^3 + \frac{(n-3)n\cdot(n-1)n\cdot(n+2)(n+4)}{3!}x^5 - \frac{(n-5)(n-3)n\cdot(n-1)n\cdot(n+2)(n+4)(n+6)}{7!}x^7 + \ldots \) when \( n \) is odd.

Realize that, with each new value of \( \lambda \), the corresponding \( P_n \)-polynomial solves a slightly different equation (that is why we have so many solutions). But we also know that, for each \( n \), the (single) equation should have a second solution. It does, but the form of these is somehow more complicated, namely: \( Q_{n-1}^{(1)} + Q_n^{(2)} \). \( \ln \frac{1+x}{1-x} \),

where \( Q_{n-1}^{(1)} \) and \( Q_n^{(2)} \) are polynomials of degree \( n-1 \) and \( n \), respectively (we have constructed the full solution for the \( n = 1 \) case). These so-called Legendre functions of second kind are of lesser importance in Physics [since they don’t meet the integrability condition, they don’t solve the eigenvalue problem]. We will not go into further details.

Optional: ►Associate Legendre Equation►

\[
(1 - x^2)y'' - 2xy' + \left( n + 1 \right)n - \frac{m^2}{1-x^2} \right) y = 0
\]

can be seen as an eigenvalue problem with \( q(x) = \frac{1}{1-x^2} \) and \( \lambda = m^2 \) [the solutions will thus be orthogonal in the \( \int \frac{1-x^2}{1-x^2} \) sense, assuming that they share the same \( n \) but the \( m \)'s are different].

By using the substitution \( y(x) = \left(1 - x^2\right)^{m/2} \cdot u(x) \), the equation is converted to \( (1 - x^2)u'' - 2(m+1)xu' + [(n+1)n - (m+1)m]u = 0 \), which has the following polynomial solution: \( P_n^{(m)}(x) \) [the \( m \)th derivative of the Legendre polynomial \( P_n \)].

Proof: Differentiate the Legendre equation \( m \) times: \( (1 - x^2)P_n^{(m)} - 2xP_n^{(m)} + (n + 1)nP_n = 0 \) getting: \( (1 - x^2)P_n^{(m+2)} - 2m\cdot xP_n^{(m+1)} - m\cdot (m-1)P_n^{(m)} - 2xP_n^{(m+1)} - 2mP_n^{(m)} + (n+1)nP_n^{(m)} = 0 \) \( \square \)

►Chebyshev equation◄

\[
(1 - x^2)y'' - xy' + \lambda y = 0
\]
[verify that this is the same equation as in the S-L discussion].

It easy to see that the Legendre recurrence formula needs to be modified to read \((i + 2)(i + 1)c_{i+2} - i(i - 1)c_i - ic_i + \lambda c_i = 0\) [only the \(ic_i\) coefficient has changed from \(-2\) to \(-1\)] \(\Rightarrow c_{i+2} = -\frac{\lambda - i^2}{(i + 2)(i + 1)} c_i\). To make the expansion finite [and the resulting function square-integrable] we have to choose

\[
\lambda = n^2 \quad \text{(Eigenvalues)}
\]

This leads to the following Chebyshev polynomials:

\[
\begin{align*}
T_0 &= 1 \\
T_1 &= x \\
T_2 &= 1 - 2x^2 \\
T_3 &= x - \frac{4}{3}x^3 \\
&\quad \ldots
\end{align*}
\]

i.e. \(1 - \frac{n^2}{3!}x^2 + \frac{n^2(n^2-2^2)}{4!}x^4 - \frac{n^2(n^2-2^2)(n^2-4^2)}{6!}x^6 + \ldots\) in the even case, and \(x - \frac{n^2-1}{3!}x^3 + \frac{(n^2-1)(n^2-3^2)}{6!}x^5 - \frac{(n^2-1)(n^2-3^2)(n^2-5^2)}{8!}x^7 + \ldots\) in the odd case.

The corresponding set of second basic solutions would consist of functions of the \(\sqrt{1-x^2}Q_n(x)\) type, where \(Q_n\) is also a polynomial of degree \(n\).

**Method of Frobenius**

The power-series technique described so far is applicable only when both \(f(x)\) and \(g(x)\) of the MAIN equation can be expanded at \(x = 0\). This condition is violated when either \(f\) or \(g\) (or both) involve a division by \(x\) or its power [e.g. \(y'' + \frac{1}{x}y' + (1 - \frac{1}{4x^2}) y = 0\)].

To make the power-series technique work in some of these cases, we must extend it in a manner described shortly (the extension is called the method of Frobenius). The new restriction is that the singularity of \(f\) is (at most) of the first degree in \(x\), and that of \(g\) is no worse than of the second degree. We can thus rewrite the main equation as

\[
y'' + \frac{a(x)}{x}y' + \frac{b(x)}{x^2} y = 0 \quad \text{(Frobenius)}
\]

where \(a(x)\) and \(b(x)\) are regular [i.e. 'expandable': \(a(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \ldots\), and \(b(x) = b_0 a_0 + b_1 x + b_2 x^2 + b_3 x^3 + \ldots\)].

The **trial solution** now has the form of

\[
y^{(T)} = \sum_{i=0}^{\infty} c_i x^{r+i} = c_0 x^r + c_1 x^{r+1} + c_2 x^{r+2} + \ldots
\]

where \(r\) is a number (no necessarily an integer) yet to be found. When substituted into the above differential equation (which is normally simplified by multiply it by \(x^2\)), the overall coefficient of the lowest \((r^{th})\) power of \(x\) is \([r(r-1) + a_0 r + b_0] c_0\).

This must (as all the other coefficients) be equal to zero, yielding the so called **indicial equation** for \(r\)

\[
r^2 + (a_0 - 1)r + b_0 = 0
\]
Even after ignoring the possibility of complex roots [assume this never happens to us], we have to categorize the solution of the indicial (simple quadratic) equation into three separate cases:

1. Two distinct real roots which don’t differ by an integer

2. A double root

3. Two roots which differ by an integer, i.e. \( r_2 - r_1 \) is a nonzero integer (zero is covered by Case 2).

We have to develop our technique separately for each of the three cases:

Distinct Real Roots

The trial solution is substituted into the differential equation with \( r \) having the value of one of the roots of the indicial equation. Making the coefficients of each power of \( x \) cancel out, one gets the usual recurrence formula for the sequence of the \( c \)-coefficients [this time we get two such sequences, one with the first root \( r_1 \) and the other, say \( r_2 \), with \( r_2 \); this means that we don’t have to worry about intermingling the two basic solutions – the technique now automatically separates them for us]. Each of the two recurrence formula allows a free choice of the first \( c \) (called \( c_0 \) and \( c_0^* \), respectively); the rest of each sequence must uniquely follow.

EXAMPLE:

\[ x^2 y'' + (x^2 + \frac{5}{36})y = 0 \] [later on we will see that this is a special case of the so called Bessel equation]. Since \( a(x) \equiv 0 \) and \( b(x) = x^2 + \frac{5}{36} \) the indicial equation reads \( r^2 - r + \frac{5}{36} = 0 \Rightarrow r_{1,2} = \frac{1}{6} \) and \( \frac{5}{6} \) [Case 1]. Substituting our trial solution into the differential equation yields \( \sum_{i=0}^{\infty} c_i (r + i)(r + i - 1)x^{r+i} + \frac{5}{36} \sum_{i=0}^{\infty} c_i x^{r+i+2} = 0 \). Introducing a new dummy index \( i^* = i + 2 \) we get \( \sum_{i=0}^{\infty} c_i [(r + i)(r + i - 1) + \frac{5}{36}]x^{r+i} + \sum_{i^*=2}^{\infty} c_{i^*} x^{r+i^*} = 0 \) [as always, \( i^* \) can now be replaced by \( i \)]. Before we can combine the two sums together, we have to deal with the exceptional \( i = 0 \) and 1 terms. The first \( (i = 0) \) term gave us our indicial equation and was made to disappear by taking \( r \) to be one of the equation’s two roots. The second one has the coefficient of \( c_1 [(r + 1)r + \frac{5}{36}] \) which can be eliminated only by \( c_1 \equiv 0 \). The rest of the left hand side is \( \sum_{i=0}^{\infty} \{ c_i [(r + i)(r + i - 1) + \frac{5}{36}] + c_{i-2} \} x^{r+i} \Rightarrow c_i = \frac{-c_{i-2}}{(r + i)(r + i - 1) + \frac{5}{36}} \). So far we have avoided substituting a specific root for \( r \) [to be able to deal with both cases at the same time], now, to build our two basic solutions, we have to set

1. \( r = \frac{1}{6} \), getting \( c_i = \frac{-c_{i-2}}{i(i - \frac{2}{3})} \Rightarrow c_2 = \frac{-c_0}{2 \times \frac{4}{3}} \), \( c_4 = \frac{-c_0}{6 \times 4 \times \frac{1}{3}} = \frac{c_0}{4 \times 2 \times \frac{4}{3} \times \frac{1}{3}} \), \( c_6 = \frac{-c_0}{6 \times 4 \times 2 \times \frac{4}{3} \times \frac{1}{3} \times \frac{1}{3}} \), \( \ldots \)[the odd-indexed coefficients must be all equal to zero]. Even though the expansion has an obvious pattern, the function cannot be identified as a 'known'
function. Based on this expansion, one can introduce a new function [eventually a whole set of them], called Bessel, as we do in full detail later on. The first basic solution is thus \( y_1 = c_0 x^{\frac{5}{6}} (1 - \frac{3}{5} x^2 + \frac{9}{320} x^4 - \frac{9}{10240} x^6 + ...). \) [For those of you who know the \( \Gamma \)-function, the solution can be expressed in a more compact form of \( \tilde{c} \sum_{k=0}^{\infty} \frac{(-1)^k \Gamma(\frac{5}{6})2^{k+1/6}}{k!\Gamma(k+\frac{5}{6})} \).]

2. \( r = \frac{5}{6} \), getting \( c_i^* = \frac{-c_i^*-2}{i(i+\frac{5}{3})} \Rightarrow c_2^* = \frac{-c_0^*}{2x^{\frac{5}{3}}}, c_4^* = \frac{-c_0^*}{4\times 2 \times x^{\frac{10}{3}}}, c_6^* = \frac{-c_0^*}{6\times 4 \times 2 \times x^{\frac{13}{3}}}, ... \Rightarrow \)

\[
y_2 = c_0 x^{\frac{5}{6}} (1 - \frac{3}{10} x^2 + \frac{9}{896} x^4 - \frac{9}{35840} x^6 + ... ) = \tilde{c} \sum_{k=0}^{\infty} \frac{(-1)^k \Gamma(\frac{5}{6})2^{k+5/6}}{k!\Gamma(k+\frac{5}{6})}.
\]

\( \blacktriangleright \) Double root \( \blacktriangleright \)

The first basic solution is constructed in the usual manner of \( y_1 = c_0 x^r + c_1 x^{r+1} + c_2 x^{r+2} + .... \) The second basic solution has the form [guaranteed to work] of:

\[
y_2 = y_1 \ln x + c_0^* x^r + c_1^* x^{r+1} + c_2^* x^{r+2} + ....
\]

where \( y_1 \) is the first basic solution (with \( c_0 \) set equal to 1, i.e. removing the multiplicative constant). The corresponding recurrence formula (for the \( c_i^* \)'s) will offer us a free choice of \( c_i^* \), which we normally set equal to 0 (a nonzero choice would only add \( c_i^* y_1 \) to our second basic solution). After that, the rest of the \( c_i^* \)'s uniquely follows (they may turn out to be all 0 in some cases).

**EXAMPLES:**

- \( (1 + x)x^2 y'' - (1 + 2x)xy' + (1 + 2x)y = 0 \) [\( a(x) = -\frac{1+2x}{1+x} \) and \( b(x) = \frac{1+2x}{1+x} \)].

  The indicial equation is \( r^2 - 2r + 1 = 0 \Rightarrow r_{1,2} = 1 \pm 0 \) [double]. Substituting \( \sum_{i=0}^{\infty} c_i x^{i+1} \) for \( y \) yields \( \sum_{i=0}^{\infty} c_i (i+1)ix^{i+1} + \sum_{i=0}^{\infty} c_i (i+1)ix^{i+2} - \sum_{i=0}^{\infty} c_i (i+1)x^{i+1} - 2 \sum_{i=0}^{\infty} c_i (i+1)x^{i+2} + \sum_{i=0}^{\infty} c_i x^{i+1} + 2 \sum_{i=0}^{\infty} c_i x^{i+2} = 0. \) Combining terms with like powers of \( x: \sum_{i=0}^{\infty} c_i i^2 x^{i+1} + \sum_{i=0}^{\infty} c_i i (i-1) x^{i+2} = 0. \) Adjusting the index of the second sum: \( \sum_{i=0}^{\infty} c_i i^2 x^{i+1} + \sum_{i=1}^{\infty} c_{i-1} (i-1) (i-2) x^{i+1} = 0. \) The 'exceptional' \( i = 0 \) term must equal to zero automatically, our indicial equation takes care of that [check], the rest implies \( c_i = -\frac{(i-1)(i-2)}{i^2} c_{i-1} \) for \( i = 1, 2, 3, .... \). yielding \( c_1 = 0, c_2 = 0, .... \) The first basic solution is thus \( c_0 x \) [i.e. \( y_1 = x, \) verify!]. Once we have identified the first basic solution as a simple function [when lucky] we have two options:

1. (a) Use V of P: \( y(x) = c(x) \cdot x \Rightarrow (1+x)xc'' + c' = 0 \Rightarrow \frac{dx}{x} = \left( \frac{1}{1+x} - \frac{1}{x} \right) dx \Rightarrow \ln z = \ln(1+x) - \ln x + c' \Rightarrow c' = c_0^* \frac{1}{x} \Rightarrow c(x) = c_0^* (\ln x + x) + c_0. \) This makes it clear that the second basic solution is \( x \ln x + x \).

   (b) Insist on using Frobenius: Substitute \( y^{(T)} = x \ln x + \sum_{i=0}^{\infty} c_i^* x^{i+1} \) into the original equation. The sum will give you the same contribution as before,
the \( x \ln x \) term (having no unknowns) yields an extra, non-homogeneous term of the corresponding recurrence equation. There is a bit of an automatic simplification when substituting \( y_1 \ln x \) (our \( x \ln x \)) into the equation, as the \( \ln x \)-proportional terms must cancel. What we need is thus \( y \to 0, \ y' \to \frac{y}{x} \) and \( y'' \to 2 \frac{y}{x} - \frac{y}{x^2} \). This substitution results in the same old [except for \( c \to c^* \)]

\[
\sum_{i=0}^{\infty} c_i^* x^{i+1} + \sum_{i=1}^{\infty} c_{i-1}^* (i-1)(i-2)x^{i+1}
\]

on the left hand side of the equation, and

\[
-(1+x)x^2 \cdot \frac{1}{x} + (1+2x)x = x^2
\]

[don’t forget to reverse the sign] on the right hand side. This yields the same set of recurrence formulas as before, except at \( i = 1 \) [due to the nonzero right-hand-side term]. Again we get a ‘free choice’ of \( c_0^* \) [indicial equation takes care of that], which we utilize by setting \( c_0^* \) equal to zero (or anything which simplifies the answer), since a nonzero \( c_0^* \) would only add a redundant \( c_0^* y_1 \) to our second basic solution. The \( x^2 \)-part of the equation (\( i = 1 \)) then reads:

\[
c_1^* x^2 + 0 = x^2 \Rightarrow c_1^* = 1.
\]

The rest of the sequence follows from \( c_i^* = -\frac{(i-1)(i-2)}{2} c_{i-1}^* \), \( i = 2, 3, 4... \Rightarrow c_2^* = c_3^* = .... = 0 \) as before. The second basic solution is thus

\[
y_1 \ln x + c_1^* x^2 = x \ln x + x^2 \text{ [check]}
\]

- \( x(x-1)y'' + (3x-1)y' + y = 0 \) \( [a(x) = \frac{3x-1}{x-1}, b(x) = \frac{x}{x-1}] \Rightarrow r^2 = 0 \) [double root of 0]. Substituting \( y^{(T)} = \sum_{i=0}^{\infty} c_i x^{i+0} \) yields

\[
\sum_{i=0}^{\infty} i(i-1)c_i x^i - \sum_{i=0}^{\infty} i(i-1)c_i x^{i-1} + 3 \sum_{i=0}^{\infty} i c_i x^i - \sum_{i=0}^{\infty} i c_i x^{i-1} + \sum_{i=0}^{\infty} c_i x^{i+1} = 0 \Rightarrow \sum_{i=0}^{\infty} [i^2 + 2i + 1] c_i x^i - \sum_{i=0}^{\infty} i^2 c_i x^{i-1} = 0 \Rightarrow \sum_{i=-1}^{\infty} (i + 1)^2 c_i x^i - \sum_{i=-1}^{\infty} (i + 1)^2 c_{i+1} x^i = 0 .
\]

The lowest, \( i = -1 \) coefficient is zero automatically, thus \( c_0 \) is arbitrary. The remaining coefficients are \( (i+1)^2 [c_i - c_{i+1}] \), set to zero \( \Rightarrow c_{i+1} = c_i \) for \( i = 0, 1, 2, ... \Rightarrow c_0 = c_1 = c_2 = c_3 = ... \Rightarrow 1 + x + x^2 + x^3 + .... = \frac{1}{1-x} \) is the first basic solution. Again, we can get the second basic solution by either the V of P or Frobenius technique.

We demonstrate only the latter:

\[
y^{(T)} = \ln x \frac{x}{1-x} + \sum_{i=0}^{\infty} c_i^* x^{i+0}
\]

getting the same left hand side and the following right hand side:

\[
x(x-1) \left[ \frac{2}{x(1-x)^2} - \frac{1}{x^2(1-x)} \right] + (3x-1) \cdot \frac{1}{x(1-x)^2} = 0 \text{ [not typical, but it may happen]. This means that not only } c_0^*, \text{ but all the other } c^*\text{-coefficients can be set equal to zero. The second basic solution is thus } \ln x \frac{x}{1-x} \text{ [which can be verified easily by direct substitution].}
\]

\[\blacktriangleright r_1 - r_2 \text{ Equals a Positive Integer} \blacktriangleleft\]

(we choose \( r_1 > r_2 \)).

The first basic solution can be constructed, based on \( y^{(T)} = \sum_{i=0}^{\infty} c_i x^{i+r_1} \), in the usual manner (don’t forget that \( r_1 \) should be the bigger root). The second basic solution will then have the form

\[Ky_1 \ln x + \sum_{i=0}^{\infty} c_i^* x^{i+r_2}\]
where $K$ becomes one of the unknowns (on par with the $c_i^*$’s), but it may turn out to have a zero value. Note that we will first have a free choice of $c_0^*$ (must be non-zero) and then, when we reach it, we will also be offered a free choice of $c_{r_1-r_2}$ (to simplify the solution, we usually set it equal to zero – a nonzero choice would only add an extra multiple of $y_1$).

**EXAMPLES:**

- $(x^2 - 1)x^2y'' - (x^2 + 1)xy' + (x^2 + 1)y = 0$ [a(x) = $\frac{x^2 + 1}{x^2 - 1}$ and $b(x) = \frac{x^2 + 1}{x^2 - 1}$] $\Rightarrow r^2 - 1 = 0 \Rightarrow r_{1,2} = 1$ and $-1$. Using $y^{(T)} = \sum_{i=0}^{\infty} c_i x^{i+1}$ we get: $\sum_{i=0}^{\infty} (i+1)i c_i x^{i+3} - \sum_{i=0}^{\infty} (i+1) c_i x^{i+1} = 0 \Rightarrow \sum_{i=0}^{\infty} r^2 c_i x^{i+3} - \sum_{i=0}^{\infty} i (i+2) c_i x^{i+1} = 0 \Rightarrow \sum_{i=0}^{\infty} r^2 c_i x^{i+3} - \sum_{i=0}^{\infty} (i+2) i c_i x^{i+1}$ requires $\sum_{i=0}^{\infty} (i+1) i c_i x^{i+1} = 0$. The lowest $i = -2$ term is zero automatically [⇒ $c_0$ can have any value], the next $i = -1$ term [still ‘exceptional’] disappears only when $c_1 = 0$. The rest of the $c$-sequence follows from $c_{i+2} = \frac{r^2 c_i}{(i+2)(i+1)}$ with $i = 0, 1, 2, ... \Rightarrow c_2 = c_3 = c_4 = ... = 0$. The first basic solution is thus $c_0 x$ [$y_1 = x$, discarding the constant]. To construct the second basic solution, we substitute $K x \ln x + \sum_{i=0}^{\infty} c_i^* x^{i-1}$ for $y$, getting: $\sum_{i=0}^{\infty} (i-1)(i-2) c_i x^{i+1} - \sum_{i=0}^{\infty} (i-1)(i+1) c_i x^{i+1} = 0 \Rightarrow \sum_{i=0}^{\infty} (i-1)(i+1) c_i x^{i+1} = 0 \Rightarrow \sum_{i=0}^{\infty} (i-1)(i+1) c_i x^{i+1}$. On the left hand side, and $-(x^2 - 1)x^2 \cdot \frac{K}{x^2} + (x^2 + 1)x \cdot K = 2K x$ on the right hand side (the contribution of $K x \ln x$). The $i = -2$ term allows $c_0^*$ to be arbitrary, $i = -1$ requires $c_1^* = 0$, and $i = 0$ [due to the right hand side, the $x^1$-terms must be also considered ‘exceptional’] requires $4c_0^* = 2K \Rightarrow K = 2c_0^*$, and leaves $c_2^*$ free for us to choose (we take $c_2^* = 0$). After that, $c_{i+2}^* = \frac{(i-2)^2}{(i+2)(i+1)} c_i^*$ where $i = 1, 2, 3, ... \Rightarrow c_4^* = c_5^* = ... = 0$. The second basic solution is thus $c_0^* (2 x \ln x + \frac{1}{x^2})$ [verify]!

- $x^2 y'' + xy' + (x^2 - \frac{1}{4}) y = 0 \Rightarrow r^2 - \frac{1}{4} = 0 \Rightarrow r_{1,2} = \frac{1}{2}$ and $-\frac{1}{2}$. Substituting $y^{(T)} = \sum_{i=0}^{\infty} c_i x^{i+1/2}$, we get $\sum_{i=0}^{\infty} (i+\frac{1}{2})(i-\frac{1}{2}) c_i x^{i+1/2} + \sum_{i=0}^{\infty} (i+\frac{1}{2}) c_i x^{i+1/2} + \sum_{i=0}^{\infty} c_i x^{i+5/2} - \frac{1}{4} \sum_{i=0}^{\infty} c_i x^{i+1/2} = 0 \Rightarrow \sum_{i=0}^{\infty} (i+\frac{1}{2}) i c_i x^{i+1/2} + \sum_{i=0}^{\infty} c_i x^{i+5/2} = 0 \Rightarrow \sum_{i=0}^{\infty} (i+3)(i+2) c_i x^{i+5/2} + \sum_{i=0}^{\infty} c_i x^{i+5/2} = 0$, which yields: a free choice of $c_0$, $c_1 = 0$ and $c_{i+2} = -\frac{c_i}{(i+2)(i+3)}$ where $i = 0, 1, ... \Rightarrow c_2 = -\frac{c_0}{3!}$, $c_4 = \frac{c_0}{5!}$, $c_6 = -\frac{c_0}{7!}$, ..., and $c_3 = c_5 = ... = 0$. The first basic solution thus equals $c_0 x^{1/2} (1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \ldots) = c_0 \frac{\sin x}{\sqrt{x}}$ [⇒ $y_1 = \frac{\sin x}{\sqrt{x}}$]. Substituting $K y_1 x + \sum_{i=0}^{\infty} c_i^* x^{i-1/2}$ for $y$ similarly reduces the equation to $\sum_{i=0}^{\infty} (i-1) i c_i^* x^{i-1/2} + \sum_{i=0}^{\infty} c_i^* x^{i+3/2}$ on the left hand side.
and \(-x^2 \cdot \left(-\frac{y_1}{x^2} + \frac{2}{x} y_1' \right) - x \cdot \frac{w_1}{x} = -2xy_1' = K(-x^{1/2} + \frac{5}{3}x^{3/2} - \frac{9}{4}x^{5/2} + ...)\) on the right hand side or, equivalently, \(\sum_{i=-2}^{\infty} (i+1)(i+2)c_i^* x^{i+3/2} + \sum_{i=0}^{\infty} c_i^* x^{i+3/2} = K(-x^{1/2} + \frac{5}{3}x^{3/2} - \frac{9}{4}x^{5/2} + ...).\) This implies that \(c_0^*\) can have any value \((i = -2), c_i^*\) can also have any value (we make it 0), \(K\) must equal zero \((i = -1),\) and \(c_{i+2} = -\frac{c_i^*}{(i+1)(i+2)}\) for \(i = 0, 1, 2, ... \Rightarrow c_2^* = -\frac{c_0^*}{6}, c_4^* = \frac{c_2^*}{4}, c_6^* = -\frac{c_4^*}{4}, ... \Rightarrow y_2 = x^{-1/2}(1 - \frac{x^2}{2} + \frac{x^4}{4} - ... ) = \cos \frac{x}{\sqrt{x}}.\]

In each of the previous examples the second basic solution could have been constructed by \(V\) of \(P\) – try it.

Also note that so far we have avoided solving a truly non-homogeneous recurrence formula – \(K\) never appeared in more than one of its (infinitely many) equations.

A few Special functions of Mathematical Physics
This section demonstrates various applications of the Frobenius technique.

\begin{itemize}
    \item Laguerre Equation
    \end{itemize}

\[y'' + \frac{1 - x}{x}y' + \frac{n}{x}y = 0\]

or, equivalently:

\[(xe^{-x}y')' + ny e^{-x} = 0\]

which identifies it as an eigenvalue problem, with the solutions being orthogonal in the integral of \(e^{-x}L_n(x) \cdot L_{n2}(x)\) dx sense.

Since \(a(x) = 1 - x\) and \(b(x) = x\) we get \(r^2 = 0\) [duplicate roots]. Substituting \(\sum_{i=0}^{\infty} c_i x^i\) for \(y\) in the original equation (multiplied by \(x\)) results in \(\sum_{i=0}^{\infty} c_i x^{i+2} + \sum_{i=0}^{\infty} (n-i)c_i x^{i-1} = 0 \Leftrightarrow \sum_{i=1}^{\infty} (i+1)^2 c_{i+1} x^{i-1} + \sum_{i=0}^{\infty} (n-i)c_i x^{i-1} = 0 \Rightarrow c_{i+1} = -\frac{n-i}{(i+1)^2} c_i\) for \(i = 0, 1, 2, ...\). Only polynomial solutions are square integrable in the above sense (relevant to Physics), so \(n\) must be an integer, to make \(c_{n+1}\) and all subsequent \(c_i\)-values equal to 0 and thus solve the eigenvalue problem.

The first basic solution is thus \(L_n(x)\) [the standard notation for Laguerre polynomials] =

\[1 - \frac{n}{1!}x + \frac{n(n-1)}{(2!)^2} x^2 - \frac{n(n-1)(n-2)}{(3!)^2} x^3 + ... \pm \frac{1}{n!}x^n\]

The second basic solution does not solve the eigenvalue problem (it is not square integrable), so we will not bother to construct it [not that it should be difficult – try it if you like].

Optional: Based on the Laguerre polynomials, one can develop the following solution to one of the most important problems in Physics (Quantum-Mechanical treatment of Hydrogen atom):
We know that \( xL''_{n+m} + (1 - x)L'_{n+m} + (n + m)L_{n+m} = 0 \) [\( n \) and \( m \) are two integers; this is just a restatement of the Laguerre equation]. Differentiating \( 2m + 1 \) times results in \( xL^{(2m+3)}_{n+m} + (2m + 1)L^{(2m+2)}_{n+m} + (1 - x)L^{(2m+2)}_{n+m} - (2m + 1)L^{(2m+1)}_{n+m} + (n + m)L^{(2m+1)}_{n+m} = 0 \), clearly indicating that \( L^{(2m+1)}_{n+m} \) is a solution to \( xy'' + (2m + 2 - x)y' + (n + m - 1)y = 0 \). Introducing a new dependent variable \( u(x) = x^{m+1}e^{-x/2}y \), i.e. substituting \( y = x^{-m-1}e^{x/2}u \) into the previous equation, leads to \( u'' + \left[ -\frac{1}{4} + \frac{n}{z} - \frac{m(m+1)}{z^2} \right] u = 0 \). Introducing a new independent variable \( z = \frac{n}{2}x \) \( \Rightarrow u' = \frac{n}{2} \dot{u} \) and \( u'' = \left( \frac{n}{2} \right)^2 \ddot{u} \) where each dot implies a \( z \)-derivative results in \( \ddot{u} + \left[ -\frac{1}{n^2} + \frac{2}{z} - \frac{m(m+1)}{z^2} \right] u = 0 \).

We have thus effectively solved the following (S-L) eigenvalue problem:

\[
\ddot{u} + \left[ \lambda + \frac{2}{z} - \frac{m(m+1)}{z^2} \right] u = 0
\]

[\( m \) considered fixed], proving that the eigenvalues are \( \lambda = -\frac{1}{n^2} \) and constructing the respective eigenfunctions [the so called ORBITALS]: \( u(z) = \left( \frac{2z}{n} \right)^{m+1}e^{-z/2}L^{(2m+1)}_{n+m}(\frac{2z}{n}) \). Any two such functions with the same \( m \) but distinct \( n_1 \) and \( n_2 \) will be orthogonal, thus: \( \int_0^\infty u_1(z)u_2(z) \, dz = 0 \) [recall the general L-S theory relating to \((pu')' + (\lambda q + r)u = 0\)]. Understanding this short example takes care of a nontrivial chunk of modern Physics. ⊗

\[ \text{Bessel equation} \]

\[ x^2y'' + xy' + (x^2 - n^2)y = 0 \]

where \( n \) has any (non-negative) value.

The indicial equation is \( r^2 - n^2 = 0 \) yielding \( r_{1,2} = n, -n \).

To build the first basic solution we use \( y^{(T)}(z) = \sum_{i=0}^{\infty} c_i z^{i+n} \Rightarrow \sum_{i=0}^{\infty} i(i+2n)c_i z^{i+n} + \sum_{i=0}^{\infty} c_i z^{i+n+2} = 0 \Leftrightarrow \sum_{i=0}^{\infty} i(i+2n)c_i z^{i+n} + \sum_{i=2}^{\infty} c_i -2i z^{i+n} = 0 \Rightarrow c_0 \) arbitrary, \( c_1 = c_3 = c_5 = \ldots = 0 \) and \( c_i = -\frac{c_{i-2}}{i(2n+i)} \) for \( i = 2, 4, 6, \ldots \Rightarrow c_2 = -\frac{c_0}{2(2n+2)}, c_4 = \frac{c_0}{4(2n+4)} \), \( c_6 = -\frac{c_0}{6(2n+6)}, \ldots, c_{2k} = \frac{c_0}{(2k+2)(2k+4)}, \ldots, c_{2k} = \frac{c_0}{(2k+2)(2k+4)(2k+6)}, \ldots, \) \( c_{2k+2} = \frac{(-1)^k c_0}{(2k+2)(2k+4)...(2k+n)k!} \) in general, where \( k = 0, 1, 2, \ldots \). When \( n \) is an integer, the last expression can be written as \( c_{2k} = \frac{(-1)^k n! c_0}{2^{2k}(n+k)!!k!} \equiv \frac{(-1)^k c_0}{2^{2k}(n+k)!} \). The first basic solution is thus

\[
\sum_{k=0}^{\infty} \frac{(-1)^k (\frac{z}{2})^{2k+n}}{k!(n+k)!}
\]

It is called the Bessel function of the first kind of ‘order’ \( n \) [note that the ‘order’ has nothing to do with the order of the corresponding equation, which is always 2], the standard notation being \( J_n(x) \); its values (if not on your calculator) can be found in tables.

When \( n \) is a non-integer, one has to extend the definition of the factorial function to non-integer arguments. This extension is called a \( \Gamma \)-FUNCTION, and is
'shifted' with respect to the factorial function, thus: \( n! \equiv \Gamma(n+1) \). For positive \( \alpha \)
\((=n+1)\) values, it is achieved by the following integral

\[
\Gamma(\alpha) \equiv \int_{0}^{\infty} x^{\alpha-1}e^{-x} dx
\]

[note that for integer \( \alpha \) this yields \((\alpha-1)!\)], for negative \( \alpha \) values the extension is

done with the help of

\[
\Gamma(\alpha - 1) = \frac{\Gamma(\alpha)}{\alpha - 1}
\]

[its values can often be found on your calculator].

Using this extension, the previous \( J_{n}(x) \) solution (of the Bessel equation) be-
comes correct for any \( n \) [upon the \((n+k)! \to \Gamma(n+k+1)\) replacement].

When \( n \) is not an integer, the same formula with \( n \to -n \) provides the second

**basic solution** [easy to verify].

Of the non-integer cases, the most important are those with a **half-integer**
value of \( n \). One can easily verify [you will need \( \Gamma(\frac{1}{2}) = \sqrt{\pi} \)] that the corresponding

Bessel functions are elementary, e.g.

\[
\begin{align*}
J_{\frac{1}{2}}(x) &= \sqrt{\frac{2}{\pi x}} \sin x \\
J_{-\frac{1}{2}}(x) &= \sqrt{\frac{2}{\pi x}} \cos x \\
J_{\frac{3}{2}}(x) &= \sqrt{\frac{2}{\pi x}} (\frac{\sin x}{x} - \cos x)
\end{align*}
\]

Unfortunately, the most common is the case of \( n \) being an **integer**.

Constructing the **second basic solution** is then a lot more difficult. It
has, as we know, the form of \( Ky_{1}\ln x + \sum_{i=0}^{\infty} c_{i}^{*} x^{i-n} \). Substituting this into the

Bessel equation yields \( \sum_{i=0}^{\infty} i(i-2n)c_{i}^{*} x^{i-n} + \sum_{i=2}^{\infty} c_{i-2}^{*} x^{i-n} \) on the left hand side and

\[
-K \left[ x^{2} \cdot (2y_{1}' - \frac{y_{1}}{x}) + x \cdot \frac{y_{1}}{x} \right] = -2K \sum_{k=0}^{\infty} \frac{(-1)^{k}(2k+n)(\frac{\pi}{2})^{2k+n}}{k!(n+k)!} \equiv
\]

\[-2K \sum_{k=n}^{\infty} \frac{(-1)^{k-n}(2k-n)(\frac{\pi}{2})^{2k-n}}{(k-n)!!k!} \text{ on the right hand side of the recurrence formula.}
\]

One can solve it by taking \( c_{0}^{*} \) to be arbitrary, \( c_{1}^{*} = c_{3}^{*} = c_{5}^{*} = \ldots = 0 \), and

\[
c_{2}^{*} = \frac{c_{0}^{*}}{2(2n-2)}, \quad c_{4}^{*} = \frac{c_{0}^{*}}{4 \times 2 \times (2n-2) \times (2n-4)}, \quad c_{6}^{*} = \frac{c_{0}^{*}}{6 \times 4 \times 2 \times (2n-2) \times (2n-4) \times (2n-6)}, \ldots,
\]

\[
c_{2k}^{*} = \frac{c_{0}^{*}}{2^{2k}(n-1)(n-2) \ldots (n-k)k!} \equiv \frac{c_{0}^{*}(n-k-1)!}{2^{2k}(n-1)!k!}
\]

up to and including \( k = n-1 \) \([i = 2n - 2]\). When we reach \( i = 2n \) the right hand side starts contributing! The overall coefficient of \( x^{n} \) is \( c_{2n-2}^{*} = -2K \frac{1}{2^{n}(n-1)!} \Rightarrow
\]

\[
K = \frac{-c_{0}^{*}}{2^{n-1}(n-1)!}
\]
allowing a free choice of \( c_{2n}^* \).

To solve the remaining part of the recurrence formula (truly non-homogeneous) is more difficult, so we only quote (and verify) the answer:

\[
c_{2k}^* = c_0^* \frac{(-1)^{k-n} (h_{k-n} + h_k)}{2^{2k}(k-n)!k!(n-1)!}
\]

for \( k \geq n \), where \( h_k = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{k} \).

**Proof:** Substituting this \( c_{2k}^* \equiv c_i^* \) into the recurrence formula and cancelling the common part of \( \frac{(-1)^{k-n} c_0^*}{2^k(k-n)!k!(n-1)!} \) yields: \( 2k(2k-2n)(h_{k-n} + h_k) - 4k(k-n)(h_{k-n-1} + h_{k-1}) = 4(2k-n) \). This is a true identity as \( h_{k-n} - h_{k-n-1} + h_k - h_{k-1} = \frac{1}{k} + \frac{1}{k+1} = \frac{2k-n}{k(k-n)} \) [multiply by \( 4(k-n) \)]. □

The second basic solution is usually written (a slightly different normalizing constant is used, and a bit of \( J_n(x) \) is added) as:

\[
Y_n(x) = \frac{2}{\pi} J_n(x) \left[ \ln \frac{x}{2} + \gamma \right] + \frac{1}{\pi} \sum_{k=n}^{\infty} \frac{(-1)^{k+1-n} (h_{k-n} + h_k)}{(k-n)!k!} \left( \frac{x}{2} \right)^{2k-n}
\]

\[
- \frac{1}{\pi} \sum_{k=0}^{n-1} \frac{(n-k-1)!}{k!} \left( \frac{x}{2} \right)^{2k-n}
\]

where \( \gamma \) is the Euler constant \( \approx 0.557 \) [the reason for the extra term is that the last formula is derived based on yet another, possibly more elegant approach than ours, namely: \( \lim_{\nu \to \infty} \frac{J_{\nu}(\cos(\nu x)) - J_{-\nu}(\cos(\nu x))}{\sin(\nu x)} \). \( Y_n(x) \) is called the Bessel function of SECOND KIND of order \( n \).

**More on Bessel functions:**

To deal with initial-value and boundary-value problems, we have to be able to evaluate \( J_n, J'_n, Y_n \) and \( Y'_n \) [concentrating on integer \( n \)]. The tables on page A97 of your textbook provide only \( J_0 \) and \( J_1 \), the rest can be obtained by repeated application of

\[
J_{n+1}(x) = \frac{2n}{x} J_n(x) - J_{n-1}(x)
\]

and/or

\[
J'_n(x) = \frac{J_{n-1}(x) - J_{n+1}(x)}{2}
\]

[with the understanding that \( J_{-1}(x) = -J_1(x) \) and \( \lim_{x \to 0} \frac{J_n(x)}{x} = 0 \) for \( n = 1, 2, 3, \ldots \)], and the same set of formulas with \( Y \) in place of \( J \).

**Proof:** \( [(\frac{x}{2})^n J_n]' = \frac{n}{x} (\frac{x}{2})^{n-1} J_n + (\frac{x}{2})^n J'_n = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+2n-1}}{k!(n+k-1)!} = (\frac{x}{2})^n J_{n-1} \) and

\[
[(\frac{x}{2})^{-n} J_n] = -\frac{n}{x} (\frac{x}{2})^{-n-1} J_n + (\frac{x}{2})^{-n} J'_n = \sum_{k=1}^{\infty} \frac{(-1)^k x^{2k-1}}{(k-1)!(n+k)!} = \sum_{k=0}^{\infty} \frac{(-1)^{k+1} x^{2k+1}}{k!(n+k+1)!} = (\frac{x}{2})^{-n} J_{n+1} \]. Divided by \( (\frac{x}{2})^n \) and \( (\frac{x}{2})^{-n} \) respectively, these give

\[
\frac{n}{x} J_n + J'_n = J_{n-1}
\]
and

\[-\frac{n}{x} J_n + J'_n = -J_{n+1}\]

Adding and subtracting the two yields the rest [for the Y functions the proof would slightly more complicated, but the results are the same]. □

EXAMPLES:

1. \(J_3(1.3) = \frac{4}{13}J_2(1.3) - J_1(1.3) = \frac{4}{13}\left[\frac{2}{13}J_1(1.3) - J_0(1.3)\right] - J_1(1.3) = \frac{4}{13}\frac{2}{13} \times 0.52202 - 0.62009] - 0.52202 = 0.0411

2. \(J'_2(1.3) = \frac{1}{2}[J_1(1.3) - J_3(1.3)] = \frac{1}{2}[0.52202 - 0.0411] = 0.2405\]

**Modified** Bessel equation:

\[x^2 y'' + xy' - (x^2 + n^2)y = 0\]

[differs from Bessel equation by a single sign]. The two basic solutions can be developed in almost an identical manner to the 'unmodified' Bessel case [the results differ only be an occasional sign]. We will not duplicate our effort, and only mention the new notation: the two basic solutions are now \(I_n(x)\) and \(K_n(x)\) [MODIFIED BESSEL FUNCTIONS of first and second kind]. Only \(I_0\) and \(I_1\) need to be tabulated as \(I_{n+1}(x) = I_{n-1}(x) - \frac{2n}{x}I_n\) and \(I_n' = \frac{I_{n-1} + I_{n+1}}{2}\) (same with \(I_n \rightarrow K_n\)).

**Transformed** Bessel equation:

\[x^2 y'' + (1 - 2a)xy' + (b^2c^2x^{2c} - n^2c^2 + a^2)y = 0\]

where \(a, b, c\) and \(n\) are arbitrary constants [the equation could have been written as \(x^2y'' + Axy' + (B^2x^{c^2} - D)y = 0\), but the above parametrization is more convenient].

To find the solution we substitute \(y(x) = x^a \cdot u(x)\) [introducing new dependent variable \(u\)] getting: \(a(a-1)u + 2axu' + x^2u'' + (1-2a)(au + xu') + (b^2c^2x^{2c} - n^2c^2 + a^2)u = x^2u'' + xu' + (b^2c^2x^{2c} - n^2c^2)u = 0\) Then we introduce \(z = bx^c\) as a new independent variable [recall that \(u' \rightarrow \frac{du}{dz}, bx^{c-1}\) and \(u'' \rightarrow \frac{d^2u}{dz^2} \cdot (bcx^{c-1})^2 + \frac{du}{dz} \cdot bc(c-1)x^{c-2} \Rightarrow x^2 \cdot \left(\frac{d^2u}{dz^2} \cdot (bcx^{c-1})^2 + \frac{du}{dz} \cdot bc(c-1)x^{c-2}\right) + x \cdot \left(\frac{du}{dz} \cdot bcx^{c-1}\right) + (b^2c^2x^{2c} - n^2c^2)u = [after cancelling c^2]

\[z^2 \cdot \frac{d^2u}{dz^2} + z \cdot \frac{du}{dz} + (z^2 - n^2)u = 0\]

which is the Bessel equation, having \(u(z) = C_1J_n(z) + C_2Y_n(z)\) [or \(C_2J_{-n}(x)\) when \(n\) is not an integer] as its general solution.

The solution to the original equation is thus

\[C_1x^aJ_n(bx^c) + C_2x^aY_n(bx^c)\]

EXAMPLES:
1. \( xy'' - y' + xy = 0 \) [same as \( x^2y'' - xy' + x^2y = 0 \)] \( \Rightarrow a = 1 \) [from \( 1 - 2a = -1 \)], \( c = 1 \) [from \( b^2c^2x^2y = x^2y \)], \( b = 1 \) [from \( b^2c^2 = 1 \)] and \( n = 1 \) [from \( a^2 - n^2c^2 = 0 \)] \( \Rightarrow 

\[ y(x) = C_1xJ_1(x) + C_2xY_1(x) \]

2. \( x^2y'' - 3xy' + 4(x^4 - 3)y = 0 \) \( \Rightarrow a = 2 \) [from \( 1 - 2a = -3 \)], \( c = 2 \) [from \( b^2c^2x^2y = 4x^4y \)], \( b = 1 \) [from \( b^2c^2 = 4 \)] and \( n = 2 \) [from \( a^2 - n^2c^2 = -12 \)] \( \Rightarrow 

\[ y = C_1x^2J_2(x^3) + C_2x^2Y_2(x^3) \]

3. \( x^2y'' + (\frac{3}{4}x^3 - \frac{35}{4})y = 0 \) \( \Rightarrow a = \frac{1}{2} \) [from \( 1 - 2a = 0 \)], \( c = \frac{3}{2} \) [from \( x^3 \)], \( b = 3 \) [from \( b^2c^2 = \frac{81}{4} \)] and \( n = 2 \) [from \( a^2 - n^2c^2 = -\frac{35}{4} \)] \( \Rightarrow 

\[ y = C_1\sqrt{x}J_2(3x^{3/2}) + C_2\sqrt{x}Y_2(3x^{3/2}) \]

4. \( x^2y'' - 5xy' + (x + \frac{35}{4})y = 0 \) \( \Rightarrow a = 3 \) [1 - 2a = -5], \( c = \frac{1}{2} \) [xy], \( b = 2 \) [\( b^2c^2 = 1 \)] and \( n = 1 \) [\( a^2 - n^2c^2 = \frac{35}{4} \)] \( \Rightarrow 

\[ y = C_1x^3J_1(2\sqrt{x}) + C_2x^3Y_1(2\sqrt{x}) \]

**Hypergeometric equation**

\[ x(1 - x)y'' + [c - (a + b + 1)x]y' - aby = 0 \]

\( \Rightarrow r^2 + r(c - 1) = 0 \Rightarrow r_{1,2} = 0 \) and \( 1 - c. \)

Substituting \( y^{(T)} = \sum_{i=0}^{\infty} c_i x^i \) yields: \( \sum_{i=0}^{\infty} (i+1)(i+c)c_{i+1} x^i - \sum_{i=0}^{\infty} (i+a)(i+b)c_i x^i \Rightarrow 

c_1 = \frac{ab}{1 - c} c_0, \quad c_2 = \frac{a(a+1)(b+1)c_0}{1 - 2c(c+1)} c_0, \quad c_3 = \frac{a(a+1)(a+2)(b+1)(b+2)c_0}{1 - 3c(c+1)(c+2)} c_0, \ldots \) which shows that the first basic solution is

\[ 1 + \frac{ab}{1 - c} x + \frac{a(a+1)(b+1)}{1 - 2c(c+1)} x^2 + \frac{a(a+1)(a+2)(b+1)(b+2)}{1 - 3c(c+1)(c+2)} x^3 + \ldots \]

The usual notation for this series is \( F(a, b; c; x) \), and it is called the hypergeometric function. Note that \( a \) and \( b \) are interchangeable. Also note that when either of them is a negative integer (or zero), \( F(a, b; c; x) \) is just a simple polynomial (of the corresponding degree) – please learn to identify it as such!

Similarly, when \( c \) is noninteger [to avoid Case 3], we can show [skipping the details now] that the second basic solution is

\[ x^{1-c}F(a + 1 - c, b + 1 - c; 2 - c; x) \]

[this may be correct even in some Case 3 situations, but don’t forget to verify it].

**EXAMPLE:**

1. \( x(1 - x)y'' + (3 - 5x)y' - 4y = 0 \) \( \Rightarrow ab = 4 \), \( a + b + 1 = 5 \) \( \Rightarrow b^2 - 4b + 4 = 0 \) \( \Rightarrow a = 2 \), \( b = 2 \), and \( c = 3 \) \( \Rightarrow C_1F(2, 2; 3; x) + C_2x^{-2}F(0, 0; -1; x) \) [the second part is subject to verification]. Since \( F(0, 0; -1; x) \equiv 1 \), the second basic solution is \( x^{-2} \), which does meet the equation [substitute].
**Transformed** Hypergeometric equation:

\[(x - x_1)(x_2 - x)y'' + [D - (a + b + 1)x]y' - aby = 0\]

where \(x_1\) and \(x_2\) (in addition to \(a\), \(b\), and \(D\)) are specific numbers.

One can easily verify that changing the independent variable to \(z = \frac{x - x_1}{x_2 - x_1}\) transforms the equation to

\[z(1 - z)\frac{d^2y}{dz^2} + \left[ \frac{D - (a + b + 1)x_1}{x_2 - x_1} - (a + b + 1)z \right] \frac{dy}{dz} - aby = 0\]

which we know how to solve [hypergeometric].

**EXAMPLES:**

1. \(4(x^2 - 3x + 2)y'' - 2y' + y = 0 \Rightarrow (x - 1)(2 - x)y'' + \frac{1}{2}y' - \frac{1}{2}y = 0 \Rightarrow x_1 = 1, x_2 = 2, \ ab = \frac{1}{4}\) and \(a + b + 1 = 0 \Rightarrow b^2 + b + \frac{1}{4} = 0 \Rightarrow a = -\frac{1}{2}\) and \(b = -\frac{1}{2}\), and finally \(c = \frac{\frac{1}{2} - (a + b + 1)x_1}{x_2 - x_1} = \frac{1}{2}\). The solution is thus

\[y = C_1 F(-\frac{1}{2}, -\frac{1}{2}; \frac{1}{2}; x - 1) + C_2 (x - 1)^{1/2} F(0, 0; \frac{3}{2}; x - 1)\]

[since \(z = x - 1\)]. Note that \(F(0, 0; \frac{3}{2}; x - 1) \equiv 1\) [some hypergeometric functions are elementary or even trivial, e.g. \(F(1, 1; 2; x) \equiv -\frac{\ln(1 - x)}{x}\), etc.].

2. \(3x(1 + x)y'' + xy' - y = 0 \Rightarrow (x + 1) (0 - x)y'' - \frac{1}{3}xy + \frac{1}{3}y = 0 \Rightarrow x_1 = -1\) [note the sign!] \(x_2 = 0, \ ab = -\frac{1}{3}\) and \(a + b + 1 = \frac{1}{3}\) \(\Rightarrow a = \frac{1}{3}\) and \(b = -1\) \(\Rightarrow c = \frac{0 - \frac{1}{3}(-1)}{1} = \frac{1}{3}\) \(\Rightarrow\)

\[y = C_1 F(\frac{1}{3}, -1; \frac{1}{3}; x + 1) + C_2 (x + 1)^{2/3} F(1, -\frac{1}{3}; \frac{5}{3}; x + 1)\]

[the first \(F(...)\) equals to \(-x\); coincidentally, even the second \(F(...)\) can be converted to a rather lengthy expression involving ordinary functions]. ■
Part II

VECTOR ANALYSIS
Chapter 7  FUNCTIONS IN THREE DIMENSIONS – DIFFERENTIATION

3-D Geometry (overview)
It was already agreed (see Prerequisites) that everyone understands the concept of Cartesian (right-handed) coordinates, and is able to visualize points and (free) vectors within this framework. Don’t forget that both vectors and point are represented by a triplet on numbers, e.g. (2,1,-4). For their names, I will normally use small boldface letters (e.g. \(a, b, c\), ... in these notes, but \(\vec{a}, \vec{b}, \vec{c}\) on the board).

\[|a| = \sqrt{a_x^2 + a_y^2 + a_z^2}\] is the vector’s LENGTH or MAGNITUDE (\(a_x, a_y\) and \(a_z\) are the vector’s three components; sometimes I may also call them \(a_1, a_2\) and \(a_3\)).

When \(|u| = 1\), \(u\) is called a UNIT vector (representing a DIRECTION). \(e_1, e_2, e_3\) is the unit vector of the +x (+y, +z) direction, respectively.

\((0,0,0)\) is called a ZERO vector.

Multiplying every component of a vector by the same SCALAR (single) number is called SCALAR MULTIPLICATION, e.g. \(3 \cdot (2, -1, 4) = (6, -3, 12)\). Geometrically, this represents modifying the vector’s length according to the scalar’s magnitude, without changing direction [a negative value of the scalar also changes the vector’s orientation].

Addition of two vectors is the corresponding component-wise operation, e.g.:
\((3, -1, 2) + (4, 0, -3) = (7, -1, -1)\). It is clearly commutative, i.e. \(a + b \equiv b + a\) [be able to visualize this].

\[a \cdot b \equiv |a| \cdot |b| \cdot \cos \gamma\]

(a scalar result), where \(\gamma\) is the angle between the direction of \(a\) and \(b\) [anywhere from 0 to \(\pi\)]. Geometrically, this corresponds to the length of the projection of \(a\) into the direction of \(b\), multiplied by \(|b|\) (or, equivalently, reverse). It is usually computed based on
\[a \cdot b \equiv a_1b_1 + a_2b_2 + a_3b_3\]
[e.g. \((2, -3, 1) \cdot (4, 2, -3) = 8 - 6 - 3 = -1\)], and it is obviously commutative [i.e. \(a \cdot b \equiv b \cdot a\)].

To prove the equivalence of the two definitions, your textbook starts with \(|a| \cdot |b| \cdot \cos \gamma\) and reduces it to \(a_1b_1 + a_2b_2 + a_3b_3\). The crucial part of their proof is the following DISTRIBUTIVE LAW, which they don’t justify: \((a + b) \cdot c \equiv a \cdot c + b \cdot c\).

To see why it is correct, think of the two projections of \(a\) and \(b\) (individually) into the direction of \(c\), and why their sum must equal the projection of \(a + b\) into the same \(c\)-direction. \(\square\)
An alternate proof (of the original equivalence) would put the butts of \(\mathbf{a}\) and \(\mathbf{b}\) into the origin (they are free vectors, i.e. free to ‘slide’), and out of all points along \(\mathbf{b}\) [i.e. \(t(b_1, b_2, b_3)\), where \(t\) is arbitrary], find the one which is closest to the tip of \(\mathbf{a}\) (resulting in the \(\mathbf{a} \rightarrow \mathbf{b}\) projection). This leads to minimizing the corresponding distance, namely \(\sqrt{(a_1 - tb_1)^2 + (a_2 - tb_2)^2 + (a_3 - tb_3)^2}\). The smallest value is achieved with \(t_m = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{b_1^2 + b_2^2 + b_3^2}\) [by the usual procedure]. Thus, the length of this projection is \(t_m |\mathbf{b}| = \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{\sqrt{b_1^2 + b_2^2 + b_3^2}} |\mathbf{b}|\).

This must equal to \(|\mathbf{a}| \cos \gamma\), as we wanted to prove. □

▷ Cross (outer) [vector] product

(\textit{notation: } \mathbf{a} \times \mathbf{b}\) is defined as a vector whose length is \(|\mathbf{a}| \cdot |\mathbf{b}| \cdot \sin \gamma\) (i.e. the area of a parallelogram with \(\mathbf{a}\) and \(\mathbf{b}\) as two of its sides), whose direction is perpendicular (ORTHOgonAL) to each \(\mathbf{a}\) and \(\mathbf{b}\), and whose orientation is such that \(\mathbf{a}\), \(\mathbf{b}\) and \(\mathbf{a} \times \mathbf{b}\) follow the right-handed pattern \([\text{this makes the product anti-}\text{commutative, i.e. } \mathbf{a} \times \mathbf{b} = -\mathbf{b} \times \mathbf{a}]\).

One way of visualizing its construction is this: project \(\mathbf{a}\) into the plane perpendicular to \(\mathbf{b}\) (≡ the blackboard, \(\mathbf{b}\) is pointing inboard), rotate this projection by +90° (counterclockwise) and multiply the resulting vector by \(|\mathbf{b}|\).

Also note that this product is \textit{not} associative: \((\mathbf{a} \times \mathbf{b}) \times \mathbf{c} \neq \mathbf{a} \times (\mathbf{b} \times \mathbf{c})\).

The cross product is usually \textit{computed} based on the following symbolic scheme:

\[
\mathbf{a} \times \mathbf{b} = \begin{vmatrix}
\mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\
a_1 & a_2 & a_3 \\
b_1 & b_2 & b_3
\end{vmatrix} = (a_2 b_3 - a_3 b_2)\mathbf{e}_1 + (a_3 b_1 - a_1 b_3)\mathbf{e}_2 + (a_1 b_2 - a_2 b_1)\mathbf{e}_3 \equiv (a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1)
\]

[e.g. \((1, 3, -2) \times (4, -2, 1) = (-1, -9, -14)]\).

The proof that the two definitions are identical rests on the validity of the distributive law: \((\mathbf{a} + \mathbf{b}) \times \mathbf{c} \equiv \mathbf{a} \times \mathbf{c} + \mathbf{b} \times \mathbf{c}\), which can be understood by visualizing \(\mathbf{a}\) and \(\mathbf{b}\) projected into a plane perpendicular to \(\mathbf{c}\), constructing the vectors on each side of the equation and showing that they are identical. □

Optional: Another way of expressing the \(k^{th}\) component of \((\mathbf{a} \times \mathbf{b})\) is:

\[
(\mathbf{a} \times \mathbf{b})_k = \sum_{i=1}^{3} \sum_{j=1}^{3} a_i b_j \epsilon_{ijk}
\]

(for \(k = 1, 2, 3\)), where \(\epsilon_{ijk}\) [called a \textit{fully antisymmetric tensor}] changes sign when any two indices are interchanged (\(\Rightarrow \epsilon = 0\) unless \(i, j, k\) distinct) and \(\epsilon_{123} = 1\) (this defines the rest).

One can show that

\[
\sum_{k=1}^{3} \epsilon_{ijk} \epsilon_{k\ell m} = \delta_{il} \delta_{jm} - \delta_{ij} \delta_{lm}
\]

(where \(\delta_{ij} = 1\) when \(i = j\) and \(\delta_{ij} = 0\) when \(i \neq j\); this is \textit{Kronecker’s delta}).
Based on this result, one can prove several useful formulas such as, for example:

\[(a \times b) \times c = (a \cdot c)b - (b \cdot c)a\]

**Proof:** The \(m^{th}\) component of the left hand side is

\[
\sum_{i,j,k,\ell} \epsilon_{ijk}\epsilon_{k\ell m}a_i b_j c_{\ell} = \sum_{i,j,\ell} (\delta_{i\ell} \delta_{jm} - \delta_{ij} \delta_{jm})a_i b_j c_{\ell} = \sum_{\ell} (a_i b_m c_{\ell} - a_m b_i c_{\ell}) \quad [\text{the } m^{th} \text{ component of the right hand side}].
\]

and

\[(a \times b) \cdot (c \times d) = (a \cdot c)(b \cdot d) - (a \cdot d)(b \cdot c)\]

having a similar proof. \(\otimes\)

**Triple product**

of \(a\), \(b\) and \(c\) is, by definition, equal to \(a \cdot (b \times c)\).

**Computationally**, this is identical to following determinant
\[
\begin{vmatrix}
  a_1 & a_2 & a_3 \\
  b_1 & b_2 & b_3 \\
  c_1 & c_2 & c_3
\end{vmatrix},
\]

it represents the volume of the parallelepiped with \(a\), \(b\) and \(c\) being three of its sides (further multiplied by \(-1\) if the three vectors constitute a left-handed set).

This implies that \(a \cdot (b \times c) = b \cdot (c \times a) = c \cdot (a \times b) = -b \cdot (a \times c) = -c \cdot (b \times a) = a \cdot (c \times b)\) [its value does not change under cyclic permutation of the three vectors].

A useful application of the triple product is the following test: \(a\), \(b\) and \(c\) are in the same plane (co-planar) iff \(a \cdot (b \times c) = 0\).

Another one is to compute the volume of an (arbitrary) tetrahedron. Note that if you use the three vectors as sides of the tetrahedron (instead of parallelepiped), its base will be half of the parallelepiped’s, and its volume will thus be \(\frac{a_0 (b \times c)}{6}\).

In general: if you slide a (planar) base along a straight line to create a 3-D volume, this volume can be computed as the area of the base times the perpendicular height; if instead you create a ‘cone’ by running a straight line from each point of the base’s boundary to the tip of the object, the corresponding volume will be 3 times smaller.

**Proof** of the last assertion: Volume is computed by

\[
\int_0^h A(x) \, dx
\]

where \(A(x)\) is the area of the cross-section at height \(x\), and \(h\) is the total height of a 3-D object.

In our case \(A(x) = A_0 \cdot \frac{(x-h)^2}{h^2}\) where \(A_0\) is the base area [right?]. This implies that the total volume equals

\[
\frac{A_0}{h^2} \int_0^h (x-h)^2 \, dx = \frac{A_0}{h^2} \left[ \frac{(x-h)^3}{3} \right]_0^h = \frac{A_0 h}{3}. \quad \Box
\]

**Optional:** \(\triangleright\) Rotation \(\triangleright\)

of a coordinate system [to match it to someone else’s, who uses the same origin but places the axes differently]. Suppose that his coordinates of our point \((x, y, z)\) are \((x', y', z')\) – what is the relationship between the two?
What is needed is some mathematical description of the corresponding rotation (to move our coordinates to his). At first one may (incorrectly) assume that a rotation is best represented by a vector [its direction being the axis of rotation, its length being the rotation angle]. The problem with such a description is this: one of the main operations we want to correctly describe is performing composition (applying one after the other) of two or more rotations, and we need the corresponding mathematical ‘machinery’. If we use the proposed vector representation of a rotation, the only 'composition' of two vectors we learned about is taking their cross product, which does not correspond to the composition of two rotations [which, unlike the cross product, is associative].

To find the proper way of representing rotations, we first realize that a rotation is a transformation (mapping) of points, symbolically: \( r' = R(r) \) [as in Physics, we now use \( r \equiv (x, y, z) \) as a general notation for a point], This transformation is obviously linear [meaning \( R(c \cdot r) = c \cdot R(r) \) and \( R(r_1 + r_2) = R(r_1) + R(r_2) \), where \( c \) is a scalar]. We already know (from Linear Algebra) that a linear transformation of \( r \) corresponds to multiplying \( r \) (in its column form) by a \( 3 \times 3 \) matrices [say \( R \)], thus: \( r' = Rr \).

But a rotation is a special case of a linear transformation; it preserves both lengths and angles (between vectors), which implies that it also preserves our dot product, i.e. \((r'_1)^T \cdot r'_2 \equiv r'^T_1 \cdot r_2 \) (a matrix representation of the dot product) for any \( r_1 \) and \( r_2 \). This is the same as \( r'^T_1 \cdot R^T \cdot R \cdot r_2 \equiv r'^T_1 \cdot r_2 \Rightarrow R^T \cdot R \equiv \mathbb{I} \) (such matrices are called orthogonal).

All this implies that rotations must be represented by orthogonal matrices. Now in reverse: Does each orthogonal matrix represent a rotation? The answer is 'no', orthogonal matrices allow the possibility of a reflection (with respect to a plane), since it also preserves lengths and angles. To eliminate reflections (and be left with 'pure' rotations only), we have to further insist that \( \det(R) = +1 \) (and not \(-1\)).

The matrix representation enables us to 'compose' two rotations by a simple matrix multiplication of the corresponding \( R_1 \) and \( R_2 \) (in reverse order), thus: \( r' = R_2 R_1 r \). This operation is associative (even though non-commutative), in full agreement with what we already know about rotations.

Finding the orthogonal matrix which corresponds to a specific rotation is a fairly complicated procedure.

There is a recent mathematical formalism which simplifies all this (a rotation is represented by a vector), but it requires a rudimentary knowledge of QUATERNION algebra (that is why it has not become widely used yet). ⊗

▶Straight Lines and Planes▶

(later to be extended to curved lines and surfaces).

There are two ways of defining a STRAIGHT LINE:

(i) **parametric representation**, i.e. \( a + b \cdot t \) where \( a \) is an arbitrary point on the straight line and \( b \) is a vector along its direction, and \( t \) (the actual **PARAMETER**) is a scalar allowed to vary from \(-\infty\) to \(+\infty\).
(ii) by two linear equations, e.g. \[
\begin{align*}
2x + 3y - 4z &= 6 \\
x - 2y + z &= -2
\end{align*}
\] (effectively an intersection of two planes).

Neither description is unique (a headache when marking assignments).

Similarly, there are two ways of defining a plane:

(i) parametric, i.e. \( \mathbf{a} + \mathbf{b} \cdot u + \mathbf{c} \cdot v \) where \( \mathbf{a} \) is an arbitrary point in the plane, \( \mathbf{b} \) and \( \mathbf{c} \) are two nonparallel vectors within the plane, and \( u \) and \( v \) are scalar parameters varying over all possible real values

(ii) by a single linear equation, e.g. \( 2x + 3y - 4z = 6 \) [note that \( (2, 3, -4) \) is a vector perpendicular to the plane, its so called normal — to prove it substitute two distinct points into the equation and subtract, getting the dot product of the connecting vector and \( (2, 3, -4) \), always equal to zero].

Again, neither description is unique.

EXAMPLES:

1. Convert \[
\begin{align*}
3x + 7y - 4z &= 5 \\
2x - 3y + z &= -4
\end{align*}
\] to its parametric representation.

Solution: The cross product of the two normals must point along the straight line, giving us \( \mathbf{b} = (3, 7, -4) \times (2, -3, 1) = (-5, -11, -23) \). Solving the two equations with an arbitrary value of \( z \) (say \( z = 0 \)) yields \( \mathbf{a} = (-\frac{13}{23}, \frac{22}{23}, 0) \).

Answer: \( (-\frac{13}{23} - 5t, \frac{22}{23} - 11t, -23t) \).

2. Find a equation of an (infinite) cylindrical surface with \( (3 - 2t, 1 + 3t, -4t) \) as its axis, and with the radius of 5.

Solution: Let us first find an expression of the (shortest) distance from a point \( \mathbf{r} \equiv (x, y, z) \) to a straight line \( \mathbf{a} + \mathbf{b} \cdot t \) [bypassing minimization]. Visualize the vector \( \mathbf{r} - \mathbf{a} \). We know that \( |\mathbf{r} - \mathbf{a}| \) is its length, and that \( (\mathbf{r} - \mathbf{a}) \cdot \frac{\mathbf{b}}{|\mathbf{b}|} \) is the length of its projection into the straight line. By Pythagoras, the direct distance is \( \sqrt{|\mathbf{r} - \mathbf{a}|^2 - \left(\frac{\mathbf{r} - \mathbf{a}}{|\mathbf{b}|}\right)^2} = \sqrt{(x - 3)^2 + (y - 1)^2 + z^2 - \frac{(-2x + 3y - 4z + 3)^2}{29}} \) (in our case). Making this equal to 5 yields the desired equation (square it to simplify).

Answer: \( (x - 3)^2 + (y - 1)^2 + z^2 - \frac{(-2x + 3y - 4z + 3)^2}{29} = 25 \).

3. What is the (shortest) distance from \( \mathbf{r} = (6, 2, -4) \) to \( 3x - 4y + z = 7 \) [bypass minimization].

Solution: \( \mathbf{n} \cdot (\mathbf{r} - \mathbf{a}) \), where \( \mathbf{n} \) is the unit normal and \( \mathbf{a} \) is an arbitrary point of the plane [found, in this case, by setting \( x = y = 0 \Rightarrow (0, 0, 7) \)].

Answer: \( \frac{(3, -4, 1)}{\sqrt{9 + 16 + 1}} \cdot (6, 2, -11) = -\frac{1}{\sqrt{26}} \) [the minus sign establishes on which side of the plane we are].
4. Find the (shortest) distance between \( \mathbf{a}_1 + \mathbf{b}_1 \cdot t \) and \( \mathbf{a}_2 + \mathbf{b}_2 \cdot t \) [bypassing minimization, as always].

Solution: To find it, we have to move perpendicularly to both straight lines, i.e. along \( \mathbf{b}_1 \times \mathbf{b}_2 \). We also know that \( \mathbf{a}_2 - \mathbf{a}_1 \) is an arbitrary connection between the two lines. The projection of this vector into the direction of \( \mathbf{b}_1 \times \mathbf{b}_2 \) supplies (up to the sign) the answer: \( \frac{(\mathbf{a}_2 - \mathbf{a}_1) \cdot \mathbf{b}_1 \times \mathbf{b}_2}{|\mathbf{b}_1 \times \mathbf{b}_2|} \) [visualize the situation by projecting the two straight lines into the blackboard so that they look parallel – always possible].

\[ \text{Curves} \]

are defined via their \textbf{parametric representation} \( \mathbf{r}(t) \equiv [x(t), y(t), z(t)] \), where \( x(t), y(t) \) and \( z(t) \) are arbitrary (continuous) functions of \( t \) (the parameter, ranging over some interval of real numbers).

\textbf{EXAMPLE:} \( \mathbf{r}(t) = [\cos(t), \sin(t), t] \) is a \textit{helix} centered on the \( z \)-axis, whose radius (when projected into the \( x-y \) plane) equals 1, with one full loop per \( 2\pi \) of vertical distance. The same \( \mathbf{r}(t) \) can be also seen as a \textit{motion} of a point-like particle, where \( t \) represents time. Note that \( [\cos(2t), \sin(2t), 2t] \) represents a different motion (the particle is moving twice as fast), but the same curve (i.e. parametrization of a curve is far from unique).

\[ \text{Arc’s length} \]

(‘arc’ meaning a specific segment of the curve). The three-component (vector) distance travelled between time \( t \) and \( t + dt \) (\( dt \) infinitesimal) is \( \mathbf{r}(t + dt) - \mathbf{r}(t) \approx \mathbf{r}(t) + \mathbf{\hat{r}}(t) dt + \ldots - \mathbf{r}(t) = \mathbf{\hat{r}}(t) dt + \ldots \), where the dots stand for terms proportional to \( dt^2 \) and higher [these give zero contribution in the \( dt \rightarrow 0 \) limit], and \( \mathbf{\hat{r}}(t) \) represents the componentwise differentiation with respect to \( t \) (the particle’s \textit{velocity}). This converts to \( |\mathbf{\hat{r}}(t)| dt + \ldots \) in terms of the actual \textit{scalar} distance (length). Adding all these infinitesimal distances (from time \( a \) to time \( b \) – these should correspond to the arc’s end points) results in

\[
\int_a^b |\mathbf{\hat{r}}(t)| \, dt
\]

which is the desired formula for the total length.

\textbf{EXAMPLES:}

1. Consider the helix of the previous example. The length of one of its complete loops (say from \( t = 0 \) to \( t = 2\pi \)) is thus

\[
\int_0^{2\pi} \sqrt{\sin(t)^2 + \cos(t)^2 + 1} \, dt = 2\pi \sqrt{2}.
\]
2. The intersect of $x^2 + y^2 = 9$ (a cylinder) and $3x - 4y + 7z = 2$ (a plane) is an ellipse. How long is it?

Solution: First we need to parametrize it, thus: $\mathbf{r}(t) = [3 \cos(t), 3 \sin(t), \frac{2 - 9 \cos(t) + 12 \sin(t)}{7}]$
where $t \in [0, 2\pi)$.

Answer: $\int_{0}^{2\pi} |\dot{\mathbf{r}}| \, dt = \int_{0}^{2\pi} \sqrt{9 + \left(\frac{9 \sin t + 12 \cos t}{7}\right)^2} \, dt$ which is an integration we cannot carry out analytically (just to remind ourselves that this can frequently happen). Numerically (using Maple), this equals 21.062.

▷ A tangent (straight) line
to a curve, at a point $\mathbf{r}(t_0)$ [$t_0$ being a specific value of the parameter] passes through $\mathbf{r}(t_0)$, and has the direction of $\dot{\mathbf{r}}(t_0)$ [the velocity]. Its parametric representation will be thus

$\mathbf{r}(t_0) + \dot{\mathbf{r}}(t_0) \cdot u$

[where $u$ is the parameter now, just to differentiate].

EXAMPLE: Using the same helix, at $t = 0$ its tangent line is $[1, u, u]$. □

▷ When $\mathbf{r}(t)$ is seen as a motion of a particle, $\dot{\mathbf{r}}(t) \equiv \mathbf{v}(t)$ gives the particle’s (instantaneous, 3-D) velocity. $|\dot{\mathbf{r}}(t)|$ then yields its (scalar) speed [the speedometer reading]. It is convenient to rewrite $\mathbf{v}(t)$ as $|\dot{\mathbf{r}}(t)| \cdot \frac{\dot{\mathbf{r}}(t)}{|\dot{\mathbf{r}}(t)|} \equiv |\dot{\mathbf{r}}(t)| \cdot \mathbf{u}(t)$

[a product of its speed and unit direction].

The corresponding (3-D) acceleration is simply $\mathbf{a}(t) \equiv \ddot{\mathbf{r}}(t)$ [a double $t$-derivative]. It is more meaningful to decompose it into its ‘TANGENTIAL’ [the one observed on the speedometer, pushing you back into your seat] and ‘NORMAL’ [observed even at constant speeds, pushing you sideways − perpendicular to the motion] components. This is achieved by the product rule:

$\frac{d\mathbf{v}(t)}{dt} = \frac{d|\dot{\mathbf{r}}(t)|}{dt} \cdot \mathbf{u}(t) + |\dot{\mathbf{r}}(t)| \cdot \frac{d\mathbf{u}(t)}{dt}$ [tangential and normal, respectively].

$\frac{d|\dot{\mathbf{r}}(t)|}{dt}$ can be simplified to $\frac{d}{dt}\sqrt{x(t)^2 + y(t)^2 + z(t)^2} = \frac{1}{2} \cdot \frac{2x\ddot{x} + 2y\ddot{y} + 2z\ddot{z}}{\sqrt{x(t)^2 + y(t)^2 + z(t)^2}} = \frac{\mathbf{u} \cdot \ddot{\mathbf{r}}}{|\dot{\mathbf{r}}(t)|}$

$\mathbf{u} \cdot \ddot{\mathbf{r}}$ (tangential magnitude)

The normal acceleration is then most easily computed from

$\ddot{\mathbf{r}} - (\mathbf{u} \cdot \ddot{\mathbf{r}})\mathbf{u}$ (normal)

[full minus tangential]. In this form it is trivial to verify that the normal acceleration is perpendicular to $\mathbf{u}$.

EXAMPLE: For our helix at $t = 0$, the speed is $\sqrt{2}$, $\mathbf{u} = [0, \frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}]$ and $\ddot{\mathbf{r}} = [-1, 0, 0]$ ⇒ zero tangential acceleration and $[-1, 0, 0]$ normal acceleration.
When interested in the geometric properties of a curve only, it is convenient to make its parametrization unique by introducing a special parameter \( s \) (instead of \( t \)) which measures the actual length travelled along the curve, i.e.

\[
s(t) = \int_0^t |\mathbf{r}(t)| \, dt
\]

where \( \mathbf{r}(t) \) is the old parametrization.

Unfortunately, to carry out the details of such a 'reparametrization' is normally too difficult [to eliminate \( t \), we would have to solve the previous equation for \( t \) — but we don’t know how to solve general equations]. Yet, the idea of this new 'uniform' (in the sense of the corresponding motion) parameter \( s \) is still quite helpful, when we realize that the previous equation is equivalent to

\[
\frac{ds}{dt} = |\dot{\mathbf{r}}(t)|
\]

This further implies that, even though we don’t have an explicit formula for \( s(t) \), we know how to differentiate with respect to \( s \), as

\[
\frac{d}{ds} = \frac{d}{dt} \frac{ds}{dt} = \frac{d}{dt} |\dot{\mathbf{r}}(t)|
\]

Note that our old \( \mathbf{u} = \frac{\dot{\mathbf{r}}(t)}{|\dot{\mathbf{r}}(t)|} \) [the unit velocity direction] can thus be defined simply as \( \frac{d}{ds} \equiv \mathbf{r}' \) [prime will imply \( s \)-differentiation].

Using this new parameter \( s \), we now define a few interesting geometrical properties (describing a curve and its behavior in space); we will immediately 'translate' these into the \( t \)-‘language’, as we normally parametrize curves by \( t \) and not \( s \):

\> Curvature

Let us first compute \( \frac{d\mathbf{u}}{ds} \equiv \mathbf{r}'' \) which corresponds to the rate of change of the unit direction per (scalar) distance travelled. The result is a vector which is always perpendicular to \( \mathbf{u} \), as we will show shortly.

Curvature \( \kappa \) is the magnitude of this \( \mathbf{r}'' \), and corresponds, geometrically, to the reciprocal of the radius of a Tangent Circle to the curve at a point [a circle with the same \( \mathbf{r} \), \( \mathbf{r}' \) and \( \mathbf{r}'' - 6 \) independent conditions].

The main thing now is to figure out: how do we compute curvature when our curve has the usual \( t \)-parametrization? This is not too difficult, as

\[
\frac{d\mathbf{u}(s)}{ds} = \frac{\frac{d\dot{\mathbf{r}}}{dt}}{|\dot{\mathbf{r}}(t)|} = \frac{\frac{d\dot{\mathbf{r}}}{dt} \cdot \dot{\mathbf{r}}}{|\dot{\mathbf{r}}|^3} - \frac{\dot{\mathbf{r}}}{|\dot{\mathbf{r}}|^3} \cdot \left( \mathbf{u} \cdot \dot{\mathbf{r}} \right) \text{ [since } \frac{d|\dot{\mathbf{r}}|}{dt} = \frac{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}{\sqrt{\dot{x}^2 + \dot{y}^2 + \dot{z}^2}} = (\mathbf{u} \cdot \dot{\mathbf{r}}) \text{] =}
\]

\[
\frac{\dot{\mathbf{r}}(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}}) - \dot{\mathbf{r}}(\dot{\mathbf{r}} \cdot \dot{\mathbf{r}})}{|\dot{\mathbf{r}}|^4}
\]

[This is easily seen to be \( \dot{\mathbf{r}} \) perpendicular, as claimed].
To get $\kappa$, we need the corresponding magnitude:

$$\sqrt{\frac{(\ddot{r} \cdot \dddot{r})(\ddot{r} \cdot \dddot{r})^2 + (\ddot{r} \cdot \dddot{r})(\dddot{r} \cdot \dddot{r})^2 - 2(\ddot{r} \cdot \dddot{r})^2(\dddot{r} \cdot \dddot{r})}{(\dot{r} \cdot \dot{r})^4}}$$

This is the final formula for computing curvature.

EXAMPLE: For the same old helix, $(\dot{r} \cdot \dot{r}) = 2$, $(\ddot{r} \cdot \ddot{r}) = 1$, and $(\dddot{r} \cdot \dddot{r}) = 0 \Rightarrow$

$$\kappa = \sqrt{\frac{1}{2^2}} = \frac{1}{2} \text{ [the same for all points of the helix – that seems to make sense; the tangent circles all have a radius of 2].}$$

▷ A few related definitions

From what we already know $\mathbf{r}'' = \kappa \cdot \mathbf{p}$ where $\mathbf{p}$ is a unit vector we will call PRINCIPAL NORMAL, automatically orthogonal to $\mathbf{u}$ and pointing towards the tangent circle’s center. Furthermore, $\mathbf{b} = \mathbf{u} \times \mathbf{p}$ must thus be yet another unit vector, orthogonal to both $\mathbf{u}$ and $\mathbf{p}$. It is called the BINORMAL vector (perpendicular to the tangent circle’s plane).

One can show that the rate of change of $\mathbf{b}$ (per unit distance travelled), namely $\mathbf{b}'$ is a vector in the direction of $\mathbf{p}$, i.e. $\mathbf{b}' = -\tau \cdot \mathbf{p}$, where $\tau$ defines the so called TORSION (‘twist’) of the curve at the corresponding point [$\tau$ is thus either $+$ or $-$ of the corresponding magnitude, the extra minus sign is just a convention].

Note that knowing a curve’s curvature and torsion, we can ‘reconstruct’ the curve (by solving the corresponding set of differential equations), but we will not go into that.

We now derive a formula for computing $\tau$ based on the usual $\mathbf{r}(t)$-parametrization.

First: $\mathbf{b}' = \mathbf{u}' \times \mathbf{p} + \mathbf{u} \times \mathbf{p}' = 0 + \mathbf{u} \times \left(\frac{\mathbf{u}'}{\kappa}\right)'$ [since $\mathbf{u}' \equiv \kappa \mathbf{p}$].

Then: $\tau = -\mathbf{p} \cdot \mathbf{b}' = -\left(\frac{\mathbf{u}'}{\kappa}\right) \cdot \left[\mathbf{u} \times \left(\frac{\mathbf{u}'}{\kappa}\right)'\right] = -\frac{\mathbf{u}' \cdot (\mathbf{u} \times \mathbf{u}'')}{\kappa^2} = \frac{\mathbf{u} \cdot (\mathbf{u}' \times \mathbf{u}'')}{\kappa^2}$.

And finally: $\mathbf{u} = \mathbf{r}' = \dot{\mathbf{r}} \frac{dt}{ds}$, $\mathbf{u}' = \mathbf{r}'' = \dot{\mathbf{r}} \frac{d^2t}{ds^2}$ and $\mathbf{u}'' = \mathbf{r}''' = \dddot{\mathbf{r}} \left(\frac{dt}{ds}\right)^3 + 3\dddot{\mathbf{r}} \frac{dt}{ds} \frac{d^2t}{ds^2} + \dot{\mathbf{r}} \frac{d^3t}{ds^3}$.

Putting it together [and realizing that, whenever identical vectors ‘meet’ in a triple product, the result is zero], we get $\tau = \frac{\dddot{\mathbf{r}} \cdot (\ddot{\mathbf{r}} \times \dddot{\mathbf{r}})}{(\dddot{\mathbf{r}} \cdot \dddot{\mathbf{r}})(\dddot{\mathbf{r}} \cdot \dddot{\mathbf{r}}) - (\dddot{\mathbf{r}} \cdot \dddot{\mathbf{r}})^2}$

which is our final formula for computing torsion.

Both the original definition and the final formula clearly imply that a planar curve has a zero torsion (identically).

EXAMPLE: For the helix $\dddot{\mathbf{r}} = (\sin t, -\cos t, 0) \Rightarrow \tau = \frac{1}{2}$.

In the next chapter we will introduce SURFACES (two-dimensional structures in 3-D; curves are of course one-dimensional) and explore the related issues. But now we interrupt this line of development to introduce
Fields

A scalar field is just a fancy name for a function of $x$, $y$ and $z$ [i.e. to each point in space we attach a single value, say its temperature], e.g. $f(x,y,z) = \frac{x(y+3)}{z}$.

A vector field assigns, to each point in space, a vector value (i.e. three numbers rather than just one). Mathematically, this corresponds to having three functions of $x$, $y$ and $z$ which are seen as three components of a vector, thus: $g(x,y,z) \equiv [g_1(x,y,z), g_2(x,y,z), g_3(x,y,z)]$, e.g. $[xy, \frac{z^3}{x}, \frac{y(x-4)}{x}]$. Physically, this may represent a field of some force, permeating the space.

An operator is a 'prescription' which takes a field and modifies it (usually, by computing its derivatives, in which case it is called a differential operator) to return another field. To avoid further difficulties relating to differential operators, we have to assume that our fields are sufficiently 'smooth' (i.e. not only continuous, but also differentiable at each point).

The most important cases of operators (acting in 3-D space) are:

- **Gradient**

which converts a scalar field $f(x,y,z)$ into the following vector field

$$\left( \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right)$$

≡ [notation] $\nabla f(x,y,z)$. The $\nabla$-operator is usually called 'del' (sometimes 'nabla'), and has three components, $\frac{\partial}{\partial x}$, $\frac{\partial}{\partial y}$ and $\frac{\partial}{\partial z}$, i.e. it can be considered to have vector attributes.

It yields the direction of the fastest increase in $f(x,y,z)$ when starting at $(x,y,z)$; its magnitude provides the corresponding rate (per unit length). This can be seen by rewriting the generalized Taylor expansion of $f$ at $\mathbf{r}$, thus: $f(\mathbf{r}+\mathbf{h}) = f(\mathbf{r}) + \mathbf{h} \cdot \nabla f(\mathbf{r}) + \text{quadratic (in } \mathbf{h}\text{-components) and higher-order terms. When } \mathbf{h} \text{ is a unit vector, } \mathbf{h} \cdot \nabla f(\mathbf{r}) \text{ provides a so called directional derivative of } f$, i.e. the rate of its increase in the $\mathbf{h}$-direction [obviously the largest when $\mathbf{h}$ and $\nabla f$ are parallel].

An interesting geometrical application is this: $f(x,y,z) = c$ [constant] usually defines a surface (a 3-D 'contour' of $f$ – a simple extension of the $f(x,y) = c$ idea). The gradient, evaluated at a point of such a surface, is obviously normal (perpendicular) to the surface at that point.

**EXAMPLE:** Find the normal direction to $z^2 = 4(x^2 + y^2)$ [a cone] at $(1,0,2)$ [this must lie on the given surface, check].

**Solution:** $f \equiv 4(x^2 + y^2) - z^2 = 0$ defines the surface. $\nabla f = (8x, 8y, -2z)$, evaluated at $(1,0,2)$ yields $(8,0,-4)$, which is the answer. One may like to convert it to a unit vector, and spell out its orientation (either inward or outward).

**Application to Physics:** If $\mathbf{r}(t)$ represents a motion of a particle and $f(x,y,z)$ a temperature of the 3-D media in which the particle moves, $\mathbf{v} \cdot \nabla f(\mathbf{r}(t))$ is the rate of change (per unit of time) of temperature as the particle experiences it [nothing but a chain rule]. To convert this into a spacial (per unit length) rate, one would have to divide the previous expression by $|\mathbf{v}|$.  


**Divergence**

converts a vector field \( \mathbf{g}(\mathbf{r}) \) to the following scalar field:

\[
\frac{\partial g_1}{\partial x} + \frac{\partial g_2}{\partial y} + \frac{\partial g_3}{\partial z}
\]

\[\equiv [\text{symbolically}] \ \nabla \cdot \mathbf{g}(\mathbf{r}).\]

Its significance (to Physics) lies in the following interpretation: If \( \mathbf{g} \) represents some flow [the direction and rate of a motion of some continuous substance in space; the rate being established by measuring mass/sec./cm.\(^2\) through an infinitesimal area perpendicular to its direction], then the divergence tells us the rate of mass loss from an (infinitesimal) volume at each point, per volume [mass/sec./cm.\(^3\)]. This can be seen by surrounding the point by an (infinitesimal) cube, and figuring out the in/out flow through each of its sides \( [h^2g_1(x + \frac{h}{2}, y, z) \text{ is the outflow from one of them, etc.}] \).

Optional: A flow (also called flux) is usually expressed as a product of the substance’ density \( \rho(\mathbf{r}) \) [measured as mass/cm.\(^3\), obviously a scalar field] and its velocity \( \mathbf{v}(\mathbf{r}) \) [measured in cm./sec., obviously a vector field]. The equation

\[
\nabla \cdot [\rho \mathbf{v}] + \frac{\partial \rho}{\partial t} = 0
\]

then expresses the conservation-of-mass law – no mass is being lost or created (no sinks nor sources), any outflow of mass results in a corresponding reduction of density. Here we have assumed that our \( \rho \) and \( \mathbf{v} \) fields are functions of not only \( x, y \) and \( z \), but also of \( t \) (time), as often done in Physics. \( \otimes \)

**EXAMPLE:** Find \( \nabla \cdot (x^2, y^2, z^2) \). Answer: \( 2x + 2y + 2z \). \( \blacksquare \)

**Curl**

(sometimes also called rotation), applied to a vector field \( \mathbf{g} \), converts it to yet another vector field symbolically defined by \( \nabla \times \mathbf{g} \), i.e.

\[
[\frac{\partial g_3}{\partial y} - \frac{\partial g_2}{\partial z}, \frac{\partial g_1}{\partial z} - \frac{\partial g_3}{\partial x}, \frac{\partial g_2}{\partial x} - \frac{\partial g_1}{\partial y}]\]

If \( \mathbf{g} \) represents a flow, \( Curl(\mathbf{g}) \) can then be visualized by holding an imaginary paddle-wheel at each point to see how fast the wheel rotates (its axis at the fastest rotation yields the curl’s direction, the torque establishes the corresponding magnitude).

**EXAMPLE:** \( Curl(x, yz, -x^2 - z^2) = (-y, 2x, 0) \). \( \blacksquare \)

\( \triangleright \) One can easily prove the following trivial identities:

\[
Curl[\text{Grad} (f)] \equiv 0
\]

\[
Div[Curl (\mathbf{g})] = 0
\]

There are also several nontrivial identities, for illustration we mention one only:

\[
Div(\mathbf{g}_1 \times \mathbf{g}_2) = \mathbf{g}_2 \cdot Curl(\mathbf{g}_1) - \mathbf{g}_1 \cdot Curl(\mathbf{g}_2)
\]
Optional: Divergence and gradient are frequently applied, consecutively, to a scalar field \( f \), to create a new scalar field \( \text{Div}[\text{Grad}(f)] \equiv \Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2} \) (where \( \Delta \) is the so called \textbf{Laplace operator}). It measures how much the value of \( f \) (at each point) deviates from its average over some infinitesimal surface [visualize a cube] centered at the point, per the surface’s area (the exact answer is obtained in the limit, as the size of the cube approaches zero).

Optional: Curvilinear coordinates (such as, for example, the \textbf{spherical coordinates} \( r, \theta \) and \( \phi \)) is a set of three new \textit{independent} variables [replacing the old \((x, y, z)\)], each expressed as a function of \( x, y \) and \( z \), and used to locate points in (3-D) space [by inverting the transformation].

Varying only the first of the new 'coordinates' (keeping the other two fixed) results in a corresponding \textbf{coordinate curve} (the changing coordinate becomes its parameter – the old \( t \) whose \textit{unit} direction is labelled \( e_r \) (similarly for the other two), and whose ‘speed’ (i.e. distance travelled per unit change of the new coordinate) is called \( h_r \) [both \( e_r \) and \( h_r \) are functions of location; they can be easily established geometrically].

When \( e_r, e_\theta \) and \( e_\phi \) remain perpendicular to each other at every point, the new coordinate system is called \textbf{orthogonal} (the case of spherical coordinates). All our subsequent results apply to orthogonal coordinates only.

Fields can be easily transformed to new coordinates, all it takes is to express \( x, y \) and \( z \) in terms of \( r, \theta \) and \( \phi \). How do we compute \( \text{Grad}, \text{Div} \) and \( \text{Curl} \) in the new coordinates, so that they agree with the old, rectangular-coordinate results?

First of all, \( \text{Grad} \) and \( \text{Curl} \) will be expressed in terms of the new, curvilinear axes \( e_r, e_\theta \) and \( e_\phi \), instead of the original \( e_1, e_2 \) and \( e_3 \). Each component of the new, curvilinear \( \text{Grad}(f) \) should express the (instantaneous) rate of increase of \( f \), when moving along the respective \( e \), per distance travelled (let us call this distance \( d \)). Thus, when we increase the value of \( r \) and start moving along \( e_r \), we obtain:

\[
\frac{\partial f}{\partial d} = \frac{\partial f}{\partial r} \cdot \frac{\partial r}{\partial d} \equiv \frac{\partial f}{h_r},
\]

by our previous definition of the \( h \)-functions. [Similarly for \( \theta \) and \( \phi \).] The full gradient is thus

\[
\text{Grad}(f) = \frac{e_r}{h_r} \cdot \frac{\partial f}{\partial r} + \frac{e_\theta}{h_\theta} \cdot \frac{\partial f}{\partial \theta} + \frac{e_\phi}{h_\phi} \cdot \frac{\partial f}{\partial \phi}
\]

[NOTE that, for spherical coordinates, \( h_r = 1 \), \( h_\theta = r \) and \( h_\phi = r \sin \theta \).]

To get \( \text{Div}(g) \), we note that an infinitesimal volume ('near-cube') built by increasing \( r \) to \( r + dr \), \( \theta \) to \( \theta + d\theta \) and \( \phi \) to \( \phi + d\phi \) has sides of length \( h_rdr, h_\theta d\theta \) and \( h_\phi d\phi \), faces of area \( h_r h_\theta drd\phi \), \( h_r h_\phi drd\theta \) and \( h_\theta h_\phi d\theta d\phi \), and volume of size \( h_r h_\theta h_\phi drd\theta d\phi \). One can easily see that \( h_\theta h_\phi g_r d\theta d\phi \) is the total flow (flux)
through one of the sides; \( \frac{\partial}{\partial r} (h_\theta h_\varphi g_r) \ d\theta d\varphi dr \) is then the corresponding flux difference between the two opposite sides. Adding the three contributions and dividing by the total volume yields:

\[
Div(g) = \frac{1}{h_r h_\theta h_\varphi} \left[ \frac{\partial}{\partial r} (h_\theta h_\varphi g_r) + \frac{\partial}{\partial \theta} (h_r h_\varphi g_\theta) + \frac{\partial}{\partial \varphi} (h_r h_\theta g_\varphi) \right]
\]

In the case of spherical coordinates this reduces to:

\[
\frac{1}{r^2 \sin \theta} \left[ \frac{\partial}{\partial r} (r^2 \sin \theta g_r) + \frac{\partial}{\partial \theta} (r \sin \theta g_\theta) + \frac{\partial}{\partial \varphi} (r g_\varphi) \right],
\]

yielding, for the corresponding Laplacian

\[
Div[Grad(f)] = \frac{1}{r^2} \frac{\partial}{\partial r} \left( r^2 \frac{\partial f}{\partial r} \right) + \frac{1}{r^2 \sin \theta} \frac{\partial}{\partial \theta} \left( \sin \theta \frac{\partial f}{\partial \theta} \right) + \frac{1}{r^2 \sin^2 \theta} \frac{\partial^2 f}{\partial \varphi^2}
\]

Understanding this is essential when computing (quantum mechanically) the energy levels (eigenvalues) of Hydrogen atom.

Similarly one can derive a formula for \( Curl(g) \), which we do not quote here (see your textbook). \( \otimes \)
Chapter 8 FUNCTIONS IN 3-D — INTEGRATION

Line Integrals
are of two types:

- Scalar (Type I) Integrals

where we are given a (scalar) function \( f(x, y, z) \) and a curve \( \mathbf{r}(t) \), and need to integrate \( f \) over an arc of the curve (which now assumes the rôle of the \( x \)-axis). All it takes is to add the areas of the individual 'rectangles' of base \( |\dot{\mathbf{r}}| \, dt \) and 'height' \( f[\mathbf{r}(t)] \), ending up with

\[
\int_{a}^{b} f[\mathbf{r}(t)] \cdot |\dot{\mathbf{r}}(t)| \, dt \quad \text{(LI)}
\]

which is just an ordinary (scalar) integral of a single variable \( t \). Note that the result must be independent of the actual curve parametrization.

In this context we should mention that all our curves are piece-wise smooth, i.e. continuous, and consisting of one or more differentiable pieces (e.g. a square).

This kind of integration can used for (spacial) averaging of the \( f \)-values (over a segment of a curve). All we have to do is to divide the above integral by the arc’s length \( \int_{a}^{b} |\dot{\mathbf{r}}(t)| \, dt \):

\[
\bar{f}_{sp} = \frac{\int_{a}^{b} f[\mathbf{r}(t)] \cdot |\dot{\mathbf{r}}(t)| \, dt}{\int_{a}^{b} |\dot{\mathbf{r}}(t)| \, dt}
\]

To average in time (taking \( \mathbf{r}(t) \) to be a motion of a particle) one would do

\[
\bar{f}_{tm} = \frac{\int_{a}^{b} f[\mathbf{r}(t)] \, dt}{b - a}
\]

instead.

The symbolic notation for this integral is

\[
\int_{C} f(\mathbf{r}) \, ds
\]

\( s \) being the special unique parameter which corresponds to the 'distance travelled', and \( C \) stands for a specific segment of a curve. To evaluate this integral, we normally use a convenient (arbitrary) parametrization of the curve (the result must be the same), and carry out the integration in terms of \( t \), using (LI).
Two other possible applications are:

1. **Center of mass** of a wire-like object of uniform mass density:

\[
\begin{bmatrix}
\int_c x \, ds & \int_c y \, ds & \int_c z \, ds \\
\int_c ds & \int_c ds & \int_c ds \\
\end{bmatrix}
\]

[the denominator is the total length \( L \)].

2. **Moment of inertia** of any such an object:

\[
\frac{M}{L} \int_c d^2 \cdot ds
\]

where \( d(x, y, z) \) is distance from the axis of rotation. [Angular acceleration is torque divided by moment of inertia].

**EXAMPLES:**

- Evaluate \( \int_C (x^2 + y^2 + z^2) \, ds \) where \( C \equiv (\cos t, \sin t, 3t) \) with \( t \in (0, 2\pi) \) [one loop of a helix].

  **Solution:**
  \[
  \int_0^{2\pi} (\cos^2 t + \sin^2 t + 9t^2) \sqrt{(\sin t)^2 + (\cos t)^2 + 9} \, dt = \sqrt{10} \int_0^{2\pi} (1 + 18t^2 + 81t^4) \, dt = \sqrt{10} \left[ t + 6t^3 + \frac{81t^5}{5} \right]_0^{2\pi} = 5.0639 \times 10^5.
  \]

- Find the center of mass of a half circle (the circumference only) of radius \( a \).

  **Solution:** \( r(t) = [a \cos t, a \sin t, 0] \Rightarrow \int_C y \, ds = a^2 \int_0^\pi \sin t \, dt = 2a^2. \)

  **Answer:** The center of mass is at \([0, \frac{2a^2}{\pi a}, 0] = [0, 0.63662a, 0]\).

- Find the moment of inertia of a circle (circumference) of mass \( M \) and radius \( a \) with respect to an axis passing through its center and two of its points.

  **Solution:** Using \( r(t) = [a \cos t, a \sin t, 0] \) and \( y \) as the axis, we get \( \int_C (x^2 + z^2) \, ds = a^3 \int_0^{2\pi} \cos^2 t \, dt = \pi a^3. \)

  **Answer:** \( \frac{M}{2\pi a} \cdot \pi a^3 = \frac{Ma^2}{2}. \)

**Vector (Type II) Integrals**

Here, we are given a vector function \( \mathbf{g}(x, y, z) \) [i.e. effectively three functions \( g_1, g_2 \) and \( g_3 \)] which represents a force on a point particle at \( (x, y, z) \), and a curve \( \mathbf{r}(t) \) which represents the particle’s ‘motion’. We know (from Physics) that, when the particle is moved by an infinitesimal amount \( d\mathbf{r} \), the energy it extracts from the field equals \( \mathbf{g} \cdot d\mathbf{r} \) [when negative, the magnitude is the amount of work needed...
to make it move]. This is independent of the actual speed at which the move is made.

The total energy thus extracted (or, with a minus sign, the work needed) when a particle moves over a segment $C$ is, symbolically,

$$\int_C \mathbf{g}(\mathbf{r}) \cdot d\mathbf{r}$$

[$\int g_1dx + g_2dy + g_3dz$ is an alternate notation] and can be computed by parametrizing the curve (any way we like – the result is independent of the parametrization, i.e. the actual motion of the particle) and finding

$$\int_{a}^{b} \mathbf{g}[\mathbf{r}(t)] \cdot \dot{\mathbf{r}}(t) \, dt \quad \text{(LII)}$$

EXAMPLE: Evaluate $\int_C (5z, xy, x^2z) \cdot d\mathbf{r}$ where $C \equiv (t, t, t^2)$, $t \in (0, 1)$.

Solution: $\int_{0}^{1} (5t^2, t^2, t^4) \cdot (1, 1, 2t) \, dt = \int_{0}^{1} (6t^2 + 2t^5) \, dt = \frac{7}{3} = 2.3333$.  

Note that, in general, the integral is path dependent, i.e. connecting the same two points by a different curve results in two different answers.

EXAMPLE: Compute the same $\int_C (5z, xy, x^2z) \cdot d\mathbf{r}$, where now $C \equiv (t, t, t)$, $t \in (0, 1)$.

Solution: $\int_{0}^{1} (5t, t^2, t^3) \cdot (1, 1, 1) \, dt = \int_{0}^{1} (5t + t^2 + t^3) \, dt = \frac{37}{12} = 3.0833$.  

Could there be a special type of vector fields to make all such vector integrals path independent?

Path Independent

The answer is yes, this happens for any $\mathbf{g}$ which can be written as

$$\nabla f(x, y, z)$$

[a gradient of a scalar field, which is called the corresponding potential; $\mathbf{g}$ is then called a conservative vector field].

Proof: $\int_C ( \nabla f) \cdot d\mathbf{r} = \int_C ( \nabla f[\mathbf{r}(t)]) \cdot \dot{\mathbf{r}}(t) \, dt = \left[\text{chain rule}\right] \int_{a}^{b} \frac{df[\mathbf{r}(t)]}{dt} \, dt = f[\mathbf{r}(b)] - f[\mathbf{r}(a)]$.  

But how can we establish whether a given $\mathbf{g}$ is conservative? Easily, the sufficient and necessary condition is

$$\text{Curl}(\mathbf{g}) \equiv 0$$
Proof: \( \mathbf{g} = \nabla f \) clearly implies that \( \text{Curl}(\mathbf{g}) \equiv 0 \).

Now the reverse: Given such a \( \mathbf{g} \), we construct (as discussed in the subsequent example) \( f = \int g_1 \, dx + \int g_2 \, dy - \int \left( \int \frac{\partial g_1}{\partial y} \, dx \right) \, dy + \int g_3 \, dz - \int \left( \int \frac{\partial g_2}{\partial z} \, dy \right) \, dz \).

This implies: \( \frac{\partial f}{\partial x} = g_1 + \int \frac{\partial g_1}{\partial y} \, dy - \int \frac{\partial g_1}{\partial z} \, dz - \int \left( \int \frac{\partial^2 g_1}{\partial y \partial z} \, dy \right) \, dz \), \( \frac{\partial f}{\partial y} = g_2 + \int \frac{\partial g_2}{\partial z} \, dz - \int \left( \int \frac{\partial^2 g_2}{\partial y \partial z} \, dy \right) \, dz \), \( \frac{\partial f}{\partial z} = g_3 - \int \int \frac{\partial^2 g_1}{\partial y^2} \, dy \) \( \equiv g_1 \).

Similarly, we can show \( \frac{\partial f}{\partial y} = g_2 \) and \( \frac{\partial f}{\partial z} = g_3 \).

Note that when \( \mathbf{g} \) is conservative, all we need to specify is the starting and final point of the arc (how you connect them is irrelevant, as long as you avoid an occasional singularity). We can then use the following notation:

\[
\int_{a}^{b} \mathbf{g}(r) \bullet dr
\]

which gives you a strong hint that \( \mathbf{g} \) is conservative (the notation would not make sense otherwise).

**EXAMPLE:** Evaluate \( \int_{(0,0,0)}^{(1,\pi/2,2)} 2xyz^2 \, dx + [x^2z^2 + z \cos(yz)] \, dy + [2x^2yz + y \cos(yz)] \, dz \).

**Solution:** This is what we used to call 'exact differential form', extended to three independent variables. We solve it by integrating \( g_1 \) with respect to \( x \) [calling the result \( f_1 \)], adding \( g_2 - \frac{\partial f_1}{\partial y} \) integrated with respect to \( y \) [call the overall answer \( f_2 \)], then adding the \( z \) integral of \( g_3 - \frac{\partial f_2}{\partial z} \), to get the final \( f \). In our case, this yields \( x^2yz^2 \) for \( f_1 \), \( x^2yz^2 + \sin(yz) \) for \( f_2 \equiv f \), as nothing is added in the last step. Thus \( f(x,y,z) = x^2yz^2 + \sin(yz) \) [check].

**Answer:** \( f(1,\pi/2,2) - f(0,0,0) = 1 + \pi = 4.1416. \)

**Optional:** We mention in passing that, similarly, \( \text{Div}(\mathbf{g}) = 0 \Leftrightarrow \) there is a vector field \( \mathbf{h} \) say such that \( \mathbf{g} \equiv \text{Curl}(\mathbf{h}) \) \( \|\mathbf{g}\| \) is then called PURELY ROTATIONAL. Any vector field \( \mathbf{g} \) can be written as \( \text{Grad}(f) + \text{Curl}(\mathbf{h}) \), i.e. decomposed into its conservative and purely rotational part. \( \otimes \)

**Double integrals**

Can be evaluated by two consecutive (univariate) integrations, the first with respect to \( x \), over its conditional range given \( y \), the second with respect to \( y \), over its marginal range (or the other way round, the two answers must agree).

**EXAMPLES:**

- To integrate over the \( \left\{ \begin{array}{l} x > 0 \\ y > 0 \\ x + y < 1 \end{array} \right\} \) triangle, we first do \( \int_{0}^{1-y} \ldots \, dx \) followed by \( \int_{0}^{1-x} \ldots \, dy \) (or \( \int_{0}^{1} \ldots \, dx \) followed by \( \int_{0}^{1-y} \ldots \, dy \)).
• To integrate over \(0 < y < \frac{1}{x}\), where \(1 < x < 3\), we can do either \(\int_{0}^{3} \int_{\frac{1}{x}}^{\frac{1}{y}} dy \, dx\)

or \(\int_{\frac{1}{x}}^{\frac{1}{y}} \int_{1}^{3} dx \, dy\) [only a graph of the region can reveal why it is so].

\[
\begin{align*}
\iint_{x^2 + y^2 < 1} y^2 \, dx \, dy &= \int_{-1}^{1} \left( \int_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} y^2 \, dx \right) \, dy = \int_{-1}^{1} 2y^2 \sqrt{1-y^2} \, dy = \\
&= \left[ \frac{1}{4} \arcsin y + \frac{1}{4} y \sqrt{1-y^2} - 12y(1-y^2)^{\frac{3}{2}} \right]_{y=-1}^{1} = \frac{\pi}{4}. \quad \blacksquare
\end{align*}
\]

The last of these double integrals can be simplified by introducing

\[\text{Polar Coordinates}\]

(effectively a change of variables, from the old \(x, y\), to a new pair of \(r, \varphi\) by:

\[
\begin{align*}
x &= r \cos \varphi \\
y &= r \sin \varphi
\end{align*}
\]

One has to remember that \(dx \, dy\) of the double integration must be replaced by \(dr \, d\varphi\), further multiplied by the **Jacobian** of the transformation, namely the absolute value of

\[
\left| \begin{array}{ccc}
\frac{\partial x}{\partial r} & \frac{\partial x}{\partial \varphi} \\
\frac{\partial y}{\partial r} & \frac{\partial y}{\partial \varphi}
\end{array} \right| = r.
\]

**EXAMPLE:** \(\iint_{x^2 + y^2 < 1} y^2 \, dx \, dy\) where \(\mathcal{R}\) is a square with corners at \((0, 1), (1, 0), (0, -1)\) and \((-1, 0)\).

Introducing \(u, v\) by \(x = u + v\) and \(y = u - v\), we will cover the same square with \(-\frac{1}{2} < u < \frac{1}{2}\) and \(-\frac{1}{2} < v < \frac{1}{2}\). Furthermore, the Jacobian of this transformation equals to 2.

\[
\begin{align*}
\text{Solution: } \int_{-\frac{1}{2}}^{\frac{1}{2}} \int_{-\frac{1}{2}}^{\frac{1}{2}} (u - v)^2 \, du \, dv &= 2 \int_{-\frac{1}{2}}^{\frac{1}{2}} \left[ \frac{u^3}{3} - 2uv^2 + uv \right]_{u=-\frac{1}{2}}^{\frac{1}{2}} \, dv = 2 \int_{-\frac{1}{2}}^{\frac{1}{2}} \left( \frac{1}{12} + v^2 \right) \, dv = 2 \left( \frac{1}{12} + \frac{1}{12} \right) = \frac{1}{3}. \quad \blacksquare
\end{align*}
\]
An important **special case** is integrating a **constant**, say \( c \), which can often be done geometrically. i.e.
\[
\int_{\mathcal{R}} c \, dx \, dy = c \cdot \text{Area}(\mathcal{R})
\]

△ Applications △

of two-dimensional integrals to geometry and physics:

▷ An **area**

of a 2-D region \( \mathcal{R} \) is computed by
\[
\int_{\mathcal{R}} \int_{\mathcal{R}} dx \, dy
\]

▷ **Center of mass**

of a 2-D object (lamina) is computed by
\[
\frac{\int \int x \rho(x, y) \, dx \, dy}{\int \int \rho(x, y) \, dx \, dy}
\]

[x component] and
\[
\frac{\int \int y \rho(x, y) \, dx \, dy}{\int \int \rho(x, y) \, dx \, dy}
\]
[y component], where \( \rho(x, y) \) is the corresponding mass density. When the object is of uniform density \( (\rho \equiv \text{const.}) \), the formulas simplify to
\[
\frac{\int \int x \, dx \, dy}{\int \int dx \, dy}
\]
and
\[
\frac{\int \int y \, dx \, dy}{\int \int dx \, dy}
\]

▷ **Moment of inertia**

with respect to some axis (this is needed when computing angular acceleration as torque/moment-of-inertia):
\[
\int \int d(x, y)^2 \cdot \rho(x, y) \, dx \, dy
\]

where \( d(x, y) \) is the (perpendicular) distance of \((x, y)\) from the axis [when the axis is \( x \), \( d \equiv y \) and vice versa; when the axis is \( z \), \( d = \sqrt{x^2 + y^2} \)].

▷ **3-D volume**
\[
\int \int h(x, y) \, dx \, dy
\]
where \( h(x, y) \) is the object’s ‘thickness’ (height) at \((x, y)\).

**EXAMPLES:**

1. Find the center of mass of a half disk of radius \( R \) and uniform mass density.
   
   **Solution:** We position the object in the upper half plane with its center at the origin, and use polar coordinates to evaluate:
   
   \[
   \frac{1}{\frac{4R}{3\pi}} = \int_0^\pi \int_0^R r \sin \varphi \cdot r \, dr \, d\varphi = \frac{R^3}{R} (\cos 0 - \cos \pi) = \frac{4R}{3\pi} = 0.42441R \]
   
   [its \( y \) component]. From symmetry, its \( x \) component must be equal to zero.

2. Find the volume of a cone with circular base of radius \( R \) and height \( H \).
   
   We do this in polar coordinates where the formula for \( h(r, \varphi) \) simplifies to \( H \cdot \frac{R-r}{R} \).
   
   **Answer:** \( \frac{R}{H} \int_0^{2\pi} \int_0^R (R - r) \cdot r \, dr \, d\varphi = \frac{2\pi R}{H} \cdot [R^2 - r^3]_{r=0}^R = \frac{\pi R^2 H}{3} \) (check).

3. Find the volume of a sphere of radius \( R \).
   
   **Solution:** Introducing polar coordinates in \( x, y \) the \( z \)-thickness is \( h(x, y) = 2\sqrt{R^2 - r^2} \) [Pythagoras]. Integrating this over the sphere’s \( x, y \) projection (a circle of radius \( R \)) yields \( \int_0^{2\pi} \int_0^R \sqrt{R^2 - r^2} \cdot r \, dr \, d\varphi = 4\pi \left[ -\frac{1}{3} (R^2 - r^2)^{\frac{3}{2}} \right]_{r=0}^R = \frac{4}{3}\pi R^3 \) (check).

4. Find the volume of the (solid) cylinder \( x^2 + z^2 < 1 \) cut along \( y = 0 \) and \( z = y \) [i.e. \( 0 < y < z \)].
   
   **Solution:** Its \( x, z \) projection is a half-circle \( x^2 + z^2 < 1 \) with \( z > 0 \), its thickness along \( y \) is \( h(x, z) = z \). Replacing \( x \) and \( z \) by polar coordinates, we can readily integrate \( \int_0^\pi \int_0^R \sin \varphi \cdot r \, dr \, d\varphi = \frac{1}{3} \cdot [-\cos \varphi]_{\varphi=0}^{\pi} = \frac{2}{3} \). There are two alternate ways of computing the volume, integrating the \( z \)-thickness over the \((x, y)\) projection, or the \( x \)-thickness over \( dy \, dz \) [try both of them].

5. Find the volume of the 3-D region defined by \( x^2 + y^2 < 1 \) and \( y^2 + z^2 < 1 \) [the common part of two cylinders crossing each other at the right angle].
   
   **Solution:** The \((x, y)\) projection of the region is describe by \( x^2 + y^2 < 1 \) (now a circle, not a cylinder), the corresponding \( z \)-thickness is \( h(x, y) = 2\sqrt{1 - y^2} \).
   
   **Answer:** \( \frac{2\pi}{3} \int_0^1 \sqrt{1 - r^2} \, dr \cdot d\varphi = \frac{2\pi}{3} \int_0^1 \left[ 1 - r^2 \sin^2 \varphi \right]_{r=0}^1 d\varphi = \frac{2\pi}{3} \int_0^\pi \frac{2(1 - |\cos \varphi|^3)}{3 \sin^3 \varphi} \, d\varphi = \int_0^\frac{\pi}{2} \frac{2(1 - |\cos \varphi|^3)}{3 \sin^3 \varphi} \, d\varphi + \int_0^\frac{\pi}{2} \frac{2(1 + \cos^3 \varphi)}{3 \sin^3 \varphi} \, d\varphi = \frac{16}{3} \). [The integration is quite tricky, later on we learn how to deal with it more efficiently].
An alternate way is to use the \((x, z)\) projection (a unit square, divided by its two diagonals into four sections of identical volume), and then integrate over one of these sections (say the right-most) the corresponding \(y\)-thickness
\[
h(x, z) = 2\sqrt{1 - x^2}, \text{ thus: } 2 \int_0^1 \sqrt{1 - x^2} \int_x^0 dz \, dx = 2 \int_0^1 \sqrt{1 - x^2} \cdot 2x \, dx = \frac{-4}{3} (1 - x^2)^{3/2} \bigg|_{x=0}^{x=1} = \frac{4}{3}.\]
The total volume is four times bigger (check). The integration was now a lot easier.

In these type of questions, it is important to first identify each side of the 3-D object (and the corresponding equation), and each of its edges (described by two equations). To project a specific edge into, say, the \((x, y)\) plane, one must eliminate \(z\) from one of the two equations and substitute into the other (getting a single \(x-y\) equation).

**Surfaces in 3-D**

There are two ways of defining a 2-D surface:

1. By an **equation**: \(f(x, y, z) = c\) \([c\ being a constant]\).

2. **Parametrically**: \(r(u, v) \equiv [x(u, v), y(u, v), z(u, v)]\) \([\text{three arbitrary functions of two parameters } u \text{ and } v; \text{ restricting these to a 2-D region selects a section of the surface}].\)

**EXAMPLES**:

- Parametrize a **sphere** of radius \(a\).

  **Answer**: \(r(u, v) = [a \sin v \cos u, a \sin v \sin u, a \cos v]\) where \(0 \leq u < 2\pi\) and \(0 \leq v \leq \pi\) \([\text{later on we introduce the so called spherical coordinates in almost the same manner – they are usually called } r, \theta \text{ and } \varphi \text{ rather than } a, v \text{ and } u].\) The curves we get by fixing \(v\) and varying \(u\) (or vice versa) are called 'coordinate' curves \([\text{latitude circles and longitude half-circles in this case}].\)

- Identify \(r(u, v) = [u \cos v, u \sin v, u].\)

  **Answer**: a 45° cone centered on \(z\).

- Parametrize the **cylinder** \(x^2 + y^2 = a^2\).

  **Solution**: \(r(u, v) = [a \cos u, a \sin u, v].\)

- Identify \([u \cos v, u \sin v, u^2].\)

  **Answer**: A paraboloid centered on \(+z\).
Surface integrals

Let us consider a specific parametrization of a surface. It is obvious that $\frac{\partial \mathbf{r}}{\partial u}$ [componentwise operation, keeping $v$ fixed] is a tangent direction to the corresponding coordinate curve and consequently tangent to the surface itself. Similarly, so is $\frac{\partial \mathbf{r}}{\partial v}$ (note that these two don’t have to be orthogonal). Constructing the corresponding tangent plane is then quite trivial.

Consequently, $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ yields a direction normal (perpendicular) to the surface, and its magnitude $|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}|$, multiplied by $dudv$, provides the area of the corresponding (infinitesimal) parallelogram, obtained by increasing $u$ by $du$ and $v$ by $dv$ [$\frac{\partial \mathbf{r}}{\partial u} du$ and $\frac{\partial \mathbf{r}}{\partial v} dv$ being its two sides]. This can be seen from:

$$|\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}| = |\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial u}| dudv \equiv dA$$

Since $|\mathbf{a} \times \mathbf{b}|^2 = |\mathbf{a}|^2 |\mathbf{b}|^2 \sin^2 \gamma = |\mathbf{a}|^2 |\mathbf{b}|^2 (1 - \cos^2 \gamma) = |\mathbf{a}|^2 |\mathbf{b}|^2 - (\mathbf{a} \cdot \mathbf{b})^2$, we can simplify it to

$$dA = \sqrt{\left|\frac{\partial \mathbf{r}}{\partial u}\right|^2 \left|\frac{\partial \mathbf{r}}{\partial v}\right|^2 - \left(\frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial v}\right)^2} \, dudv$$

which is more convenient computationally (bypassing the cross product).

To find an area of a whole surface (or its section), we need to 'add' the contributions from all these parallelograms, thus:

$$\text{Area} = \iint_{\mathcal{R}} dA = \iint_{\mathcal{S}} \sqrt{\left|\frac{\partial \mathbf{r}}{\partial u}\right|^2 \left|\frac{\partial \mathbf{r}}{\partial v}\right|^2 - \left(\frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial v}\right)^2} \, dudv$$

where $\mathcal{R}$ is the $(u, v)$ region needed to cover the (section of the) surface $\mathcal{S}$. Needless to say, the answer must be the same, regardless of the parametrization.

EXAMPLES:

1. Find the tangent plane to the ellipsoid $3x^2 + 2y^2 + z^2 = 20$ at $(1, 2, 3)$.

Solution: First one can easily check that the point is on the ellipsoid (just in case). We can parametrize the upper half of the ellipsoid (which is sufficient in this case) by $\mathbf{r}(u, v) = (u, v, \sqrt{20 - 3u^2 - 2v^2})$. Then $\frac{\partial \mathbf{r}}{\partial u} = (1, 0, -\frac{3u}{\sqrt{20 - 3u^2 - 2v^2}}) = (1, 0, -1)$ and $\frac{\partial \mathbf{r}}{\partial v} = (0, 1, -\frac{2v}{\sqrt{20 - 3u^2 - 2v^2}}) = (0, 1, -\frac{4}{3})$. The corresponding cross product $(1, 0, -1) \times (0, 1, -\frac{4}{3}) = (1, \frac{4}{3}, 1)$ yields the tangent plane’s normal; we also know that the plane has to pass through $(1, 2, 3)$.

Answer: $3x + 4y + 3z = 20$. 
2. Find the area of a surface of a sphere of radius $a$.

Solution: Using the $\mathbf{r}(u,v) = (a \sin v \cos u, a \sin v \sin u, a \cos v)$ parametrization, we get: $\frac{\partial \mathbf{r}}{\partial u} = (-a \sin v \sin u, a \sin v \cos u, 0)$ and $\frac{\partial \mathbf{r}}{\partial v} = (a \cos v \cos u, a \cos v \sin u, -a \sin v)$ $\Rightarrow dA \equiv a^{2} |\sin v| \, du \, dv$.

Answer: $A = 2 \pi a^{2}$.

3. Find the surface area of a torus (donut) of dough-radius equal to $b$ and hole-radius equal to $a - b$.

Solution: We make $z$ its axis, and $[0, a + b \cos v, b \sin v]$ its cross section with the $(y, z)$ plane. The full parametrization is then: $\mathbf{r}(u,v) = [(a + b \cos v) \cos u, (a + b \cos v) \sin u, b \sin v]$, where both $u$ and $v$ vary from 0 to $2\pi$. This yields $\frac{\partial \mathbf{r}}{\partial u} = [-(a + b \cos v) \sin u, (a + b \cos v) \cos u, 0]$, $\frac{\partial \mathbf{r}}{\partial v} = [-b \sin v \cos u, b \sin v \sin u, -b \cos v]$ $\Rightarrow dA \equiv b(a + b \cos v)$.

Answer: $b \int_{0}^{2\pi} \int_{0}^{2\pi} (a + \cos v) \, dv \, du = 4\pi^{2} ab$.

Computing areas is just a special case of a

\[\text{Surface Integral of Type I} \]

(‘SCALAR’ type). In general, we can integrate any scalar function $f(x,y,z)$ over a surface $S$ [symbolic notation $\iint_{S} f(x,y,z) \, dA$ by parametrizing the surface and computing

$\iiint_{\mathcal{R}} f[\mathbf{r}(u,v)] \cdot \sqrt{\left(\frac{\partial \mathbf{r}}{\partial u}\right)^{2} + \left(\frac{\partial \mathbf{r}}{\partial v}\right)^{2} - \left(\frac{\partial \mathbf{r}}{\partial u} \cdot \frac{\partial \mathbf{r}}{\partial v}\right)^{2}} \, du \, dv$

[the answer is independent of parametrization].

When divided by the corresponding surface area, this represents the \textbf{average} of $f(x,y,z)$ over $S$.

\textbf{Other applications} to Physics are:

1. Moment of inertia of a shell-like structure (LAMINA) of surface density $\rho(x,y,z)$:

$\iiint_{S} d^{2} \cdot \rho \cdot dA$

where $d(x,y,z)$ is the distance from the rotation axis. For a lamina of uniform density, $\rho = \frac{M}{A}$ (total mass over total area).

2. Center of mass

$\left[ \iint_{S} x \cdot \rho \cdot dA, \iint_{S} y \cdot \rho \cdot dA, \iint_{S} z \cdot \rho \cdot dA \right]$ $\left[ \iint_{S} \rho \cdot dA, \iint_{S} \rho \cdot dA, \iint_{S} \rho \cdot dA \right]$

($\rho$ cancels out when constant, i.e. uniform mass density). Note that $\iint_{S} \rho \cdot dA$ is the total mass. $\Box$
EXAMPLE: Find the moment of inertia of a spherical shell of radius $a$ and total mass $M$ (uniformly distributed) with respect to an axis going through its center.

Solution: 'Borrowing' the parametrization (and $dA$) from the previous Example 2, and using $z$ as the axis, we get

$$a^2 \sin v \, du \, dv = 2\pi \rho a^4 \left[ \cos^3 v - \cos v \right]_{v=0}^\pi = 2\pi \rho \frac{M}{4\pi a^2} a^4 \cdot \frac{4}{3} = \frac{2}{3} Ma^2. \blacksquare$$

Surface integrals of Type II

('vector' type): When integrating a vector field $\mathbf{g}(x, y, z)$ [representing some stationary flow] over an orientable (having two sides) surface $\mathcal{S}$, we are usually interested in computing the total flow (flux) through this surface, in a chosen direction.

The flow through an 'infinitesimal' area [our parallelogram] of the surface is given by the dot product

$$\mathbf{g} \cdot \mathbf{n} \, dA$$

where $\mathbf{n}$ is a unit direction normal [perpendicular] to the area, since the flow is obviously proportional to the area’s size $dA$, to the magnitude of $\mathbf{g}$ (the flow’s speed), and to the cosine of the $\mathbf{n}$-$\mathbf{g}$ angle.

'Adding' these, one gets

$$\iint_{\mathcal{S}} \mathbf{g} \cdot \mathbf{n} \, dA \equiv \iint_{\mathcal{S}} \mathbf{g} \cdot d\mathbf{A} \equiv \iint_{\mathcal{S}} (g_1 \, dz \, dy + g_2 \, dx \, dz + g_3 \, dx \, dy)$$

introducing two more alternate, symbolic notations (I usually use the middle one).

We can convert this to a regular double-integral (in $u$ and $v$), by parametrizing the surface [different parametrizations must give the same correct answer] and replacing $\mathbf{n} \, dA$ by $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ [having both the correct area and direction], getting:

$$\iint_{\mathcal{R}} \mathbf{g}[\mathbf{r}(u, v)] \cdot \left[ \frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right] \, du \, dv$$

where $\mathcal{R}$ is the $(u, v)$ region corresponding to $\mathcal{S}$. Note that $\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}$ does not necessarily have the correct (originally prescribed) orientation; when that happens, we fix it by reversing the sign of the result.

EXAMPLES (to simplify our notation, we use $\frac{\partial \mathbf{r}}{\partial u} \equiv \mathbf{r}_u$ and $\frac{\partial \mathbf{r}}{\partial v} \equiv \mathbf{r}_v$):

1. Evaluate $\iint_{\mathcal{S}} (x, y, z - 3) \cdot d\mathbf{A}$ where $\mathcal{S}$ is the upper (i.e. $z > 0$) half of the $x^2 + y^2 + z^2 = 9$ sphere, oriented upwards.

Solution: Here we can bypass spherical coordinates (why?) and use instead $\mathbf{r}(u, v) = \left[ u, v, \sqrt{9 - u^2 - v^2} \right]$ with $u^2 + v^2 < 9$ [defining the two-dimensional region $\mathcal{R}$ over which we integrate]. Furthermore, $\mathbf{r}_u = [1, 0, -\frac{u}{\sqrt{9 - u^2 - v^2}}]$ and
\( \mathbf{r}_v = [0, 1, -\frac{v}{\sqrt{9 - u^2 - v^2}}] \Rightarrow \mathbf{r}_u \times \mathbf{r}_v = [\frac{u}{\sqrt{9 - u^2 - v^2}}, \frac{v}{\sqrt{9 - u^2 - v^2}}, 1] \) [correct orientation!] \( \Rightarrow \mathbf{g} \cdot (\mathbf{r}_u \times \mathbf{r}_v) = \frac{2^3}{2 \pi} \int_0^3 \int_0^9 (\frac{9}{\sqrt{9 - r^2}} - 3) \, r \, dr \, d\varphi = 27 \pi. \)

2. Evaluate \( \iint_S (yz, xz, xy) \cdot dA \) where \( S \) is the full \( x^2 + y^2 + z^2 = 1 \) sphere oriented outwards. Using the usual parametrization: \( \mathbf{r}(u, v) = (\cos u \sin v, \sin u \sin v, \cos v) \Rightarrow \mathbf{r}_u = (-\sin u \sin v, \cos u \sin v, 0), \mathbf{r}_v = (\cos u \cos v, \sin u \cos v, -\sin v) \) and \( \mathbf{r}_u \times \mathbf{r}_v = (-\cos u \sin^2 v, -\sin u \sin^2 v, -\sin v \cos v) \) [wrong orientation, reverse its sign!], we get \( \mathbf{g} \cdot (-\mathbf{r}_u \times \mathbf{r}_v) = 3 \cos u \sin u \sin^3 v \cos v. \)

**Answer:** \( 3 \int_0^{2\pi} \sin u \cos u \, du \times \int_0^{\pi} \sin^3 v \cos v \, dv = 0 \)

Shortly we learn a shortcut for evaluating Type II integrals over a closed surface which will make the last example trivial. But first we need to discuss

'Volume' integrals

which are only of the scalar type (there is no natural direction to associate with an infinitesimal volume, say a cube; contrast this with the tangent direction for a curve and the normal direction for a surface).

The other difference is that the 3-D integration can be carried out directly in terms of \( x, y \) and \( z \), which are in a sense direct 'parameters' of the corresponding 3-D region (sometimes called 'volume'). This is not to say that we can not try different (more convenient) ways of 'parametrizing' it, but this will now be referred to as a 'change of variables' (or introducing generalized coordinates).

The most typical example of these are the spherical coordinates \( r, \theta \) and \( \varphi \) (to simplify integrating over a sphere). When using spherical coordinates, \( dx \, dy \, dz \) (\( \equiv dV \)) needs to be replaced by \( dr \, d\theta \, d\varphi \) multiplied by the Jacobian of the transformation, namely:

\[
\begin{vmatrix}
\sin \theta \cos \varphi & r \cos \theta \cos \varphi & -r \sin \theta \sin \varphi \\
\sin \theta \sin \varphi & r \cos \theta \sin \varphi & r \sin \theta \cos \varphi \\
\cos \theta & -r \sin \theta & 0 \\
\end{vmatrix} = r^2 \sin \theta
\]

[this expression can be derived and understood geometrically].

Similarly to double integration of a constant, some triple integrals can be also evaluated 'geometrically', by

\[
\iiint_V c \, dV = c \cdot \text{Volume}(V)
\]

whenever the 3-D region is of a simple enough shape, and we remember a formula for the corresponding volume.
Possible Applications

of volume integrals include computing the actual volume of a 3-D body

\[ V = \iiint_V dV \]

averaging a scalar function \( f(x, y, z) \) over a 3-D region

\[ \iiint_V f(x, y, z) dV \]

computing the center of mass of a 3-D object of mass density \( \rho(x, y, z) \) [it cancels out when constant]

\[
\begin{bmatrix}
\iiint_V x \rho(x, y, z) dV \\
\iiint_V y \rho(x, y, z) dV \\
\iiint_V z \rho(x, y, z) dV
\end{bmatrix} = \begin{bmatrix}
\iiint_V \rho(x, y, z) dV \\
\iiint_V \rho(x, y, z) dV \\
\iiint_V \rho(x, y, z) dV
\end{bmatrix}
\]

and computing the corresponding moment of inertia

\[ \iiint_V d^2 \rho dV \]

where \( d(x, y, z) \) is distance from the rotational axis, and \( \rho \equiv \frac{M}{V} \) when the mass density is uniform.

EXAMPLE: Find the moment of inertia of a uniform sphere of radius \( a \) with an axis going through its center.

Solution: 

\[
\frac{M}{\frac{4}{3}\pi a^3} \iiint_V (x^2 + y^2) dV = \frac{M}{\frac{4}{3}\pi a^3} 2\pi \int_0^a \int_0^{2\pi} r^2 \sin^2 \theta \cdot r^2 \sin \theta \, dr \, d\theta \, d\varphi = \frac{M}{\frac{4}{3}\pi a^3} \cdot \frac{2}{5} a^5.
\]

\[
\left[ \cos^3 \theta - \cos \theta \right]_{\theta=0}^{\pi} \cdot 2\pi = \frac{2}{5} Ma^2.
\]

There is an interesting and useful relationship between a Type II integral over a closed (outward oriented) surface \( S_c \), and a volume integral over the 3-D region \( V \) enclosed by this \( S_c \), called

Gauss Theorem

\[ \int \int \int_{S_c} \mathbf{g} \cdot d\mathbf{A} \equiv \iiint_V \text{Div}(\mathbf{g}) \, dV \]

[\text{Div}(\mathbf{g}) \text{ must have no singularities throughout } V].

Indication of Proof: We have already seen this to be true for an infinitesimal volume \( dx \, dy \, dz \) when we introduced divergence of a vector field. When the contributions of all these infinitesimal volumes are added together (to build the surface integral on the left hand side of our formula) the adjacent-side flows cancel out and we are left with the overall surface only; adding the divergences (each multiplied by the corresponding infinitesimal volume) results in the right-hand-side integral. \( \square \)
EXEMPLARY:

1. The integral of Example 2 from the previous section thus becomes quite trivial, as $\text{Div} ([yz, xz, xy]) \equiv 0$.

2. Evaluate $\iint_S (x^3, x^2y, x^2z) \cdot \mathbf{n} \, dA$, where $S$ is the surface of $\{ x^2 + y^2 < a^2 \}
0 < z < b$ (a cylinder of radius $a$ and height $b$), oriented outwards.

Solution: Using the Gauss theorem, we get $\iiint_{x^2+y^2<a^2} 5x^2 \, dV = 5b \iint_{x^2+y^2<a^2} x^2 \, dx \, dy = \frac{5b}{4} a^4 \pi$.

Let us verify this by recomputing the original surface integral directly (note that now we have to deal with three distinct surfaces: the top disk, the bottom disk, and the actual cylindrical walls): The top can be parametrized by $r(u, v) = [u, v, b]$, contributing $\iint_{u^2+v^2<a^2} [u^3, u^2v, u^2b] \cdot (0, 0, 1) \, du \, dv = \frac{b^2 a^4}{4} \pi$. The bottom is parametrized by $r(u, v) = [u, v, 0]$, contributing minus (because of the wrong orientation) $\iint_{u^2+v^2<a^2} [u^3, u^2v, 0] \cdot (0, 0, 1) \, du \, dv = 0$. Finally, the sides are parametrized by $r(u, v) = [a \cos u, a \sin u, v]$, contributing $\iint_{0<u<2\pi, 0<v<b} [a^3 \cos^3 u, a^3 \cos^2 u \sin u, a^2 v \cos^2 u] \cdot [a \cos u, a \sin u, 0] \, du \, dv = a^4 \int_0^{2\pi} \int_0^b \cos^2 u \, du \, dv = a^4 b \pi$. Adding the three contributions gives $\frac{5}{4} a^4 b \pi$ [check].

Similarly, there is an interesting relationship between the Type II line integral over a closed curve $C_d$ and a Type II surface integral over any surface $S$ having $C_d$ as its boundary, called

\[ \mathbf{\text{Stokes' Theorem}} \]

\[ \iint_S \text{Curl}(\mathbf{g}) \cdot \mathbf{n} \, dA \equiv \oint_{C_d} \mathbf{g} \cdot d\mathbf{r} \]

where the orientation of $C_d$ and that of $\mathbf{n} \, dA$ follow the right-handed pattern. [When $C_d$ and $S$ lie in the $(x, y)$ plane, this is known as the Green's Theorem].

Indication of Proof (which is, this time, a lot more complicated):

We parametrize $S$ and then, using the corresponding coordinate lines, we divide $S$ into many infinitesimal parallelograms and evaluate $\oint_{C_d} \mathbf{g} \cdot d\mathbf{r}$ for each of these. When adding these together, the contributions of any two adjacent sides cancel out, and we end up with the integral on the right hand side of the Stokes' formula.

On the other hand, the contribution of each of these integrals can be approximated (the approximation becomes exact in the appropriate limit) by the
difference in $g$ between two opposite sides of the parallelogram $(\frac{\partial g}{\partial u} du$ and $\frac{\partial g}{\partial v} dv$) dot-multiplied by the vector representation of the two sides $(\frac{\partial r}{\partial u} du$ and $\frac{\partial r}{\partial v} dv$, respectively). These two are then subtracted (since one runs with, and the other one against, the counterclockwise orientation of the boundary) to get

$$\oint g \cdot dr \simeq (\frac{\partial g}{\partial u} \cdot \frac{\partial r}{\partial v} - \frac{\partial g}{\partial v} \cdot \frac{\partial r}{\partial u}) du dv$$

Using the chain rule for expanding $\frac{\partial r}{\partial u}$ and $\frac{\partial r}{\partial v}$, and expressing the dot products in terms of individual components, we get $\sum_{i,j=1}^{3} (\frac{\partial g_i}{\partial r_i} \cdot \frac{\partial r_j}{\partial u} \cdot \frac{\partial r_k}{\partial v} \cdot (\delta_{ik} \delta_{jk} - \delta_{ij} \delta_{jk}) = \sum_{i,j,k,l,m=1}^{3} \frac{\partial g_i}{\partial r_l} \cdot \frac{\partial r_j}{\partial u} \cdot \frac{\partial r_k}{\partial v} \cdot \frac{\partial r_m}{\partial u} \cdot \epsilon_{ijk} \epsilon_{klm} = \text{Curl}(g) \cdot (\frac{\partial r}{\partial u} \times \frac{\partial r}{\partial v})$. These, when added together, result in the left hand side of our formula. □

EXAMPLE: Evaluate $\int f(y, xz^3, -zy^3) \cdot dr$, where $C_d$ is defined by \[ \begin{align*}
    x^2 + y^2 &= 4 \\
    z &= -3
\end{align*} \]
clockwise when viewed from the top.

Solution: Using Stokes’ Theorem we replace this integral by $\iint_S [\text{Curl}(\mathbf{g}) \cdot \mathbf{n}] dA$, where $S$ is the corresponding (flat) disk. Parametrizing $S$ by $\mathbf{r}(u, v) \equiv [u, v, -3] \Rightarrow \mathbf{n} dA = [0, 0, 1] du dv$, this converts to $\iint_{u^2+v^2<4} (-28) du dv = -28 \cdot 4\pi = -112\pi$ [note that we did not need to know the first two components of $\text{Curl}(\mathbf{g})$ in this case, i.e. it pays to do the $\mathbf{n} dA$ first].

We will verify the answer by performing the original line integral, directly: $\mathbf{r}(t) = [2 \cos t, 2 \sin t, -3]$ is the parametrization of $C_d$, which converts the integral to $\int [2 \sin t, -54 \cos t, 24 \sin^3 t] \cdot [-2 \sin t, 2 \cos t, 0] dt = \int_{0}^{2\pi} (-4 \sin^2 t - 108 \cos^2 t) dt = -112\pi$ [almost equally easily]. □

Unless $\text{Curl}(\mathbf{g}) \equiv 0$, the computational simplification achieved by applying the Stokes’ theorem is very limited (a far cry from the Gauss theorem). One exception is when $C_d$ is a ‘broken' planar curve (consisting of several segments), as we can trade one surface integral for several line integrals.

Review exercises

1. Find the area of the following (truncated) paraboloid: \[ \begin{align*}
    z &= x^2 + y^2 \\
    z &< b
\end{align*} \]

Solution: Parametrize: $\mathbf{r} = [u, v, u^2+v^2] \Rightarrow \mathbf{r}_u = [1, 0, 2u]$ and $\mathbf{r}_v = [0, 1, 2v] \Rightarrow dA = \sqrt{(1 + 4u^2)(1 + 4v^2) - 16u^2v^2} du dv = \sqrt{1 + 4(u^2 + v^2)} du dv$. We need

$$\iint_{u^2+v^2<b} dA = \text{going polar} \int_{0}^{2\pi} \int_{0}^{\sqrt{b}} \sqrt{1 + 4r^2} \cdot r dr d\varphi = 2\pi \left[ \frac{1}{12} (1 + 4r^2)^{\frac{3}{2}} \right]_{r=0}^{\sqrt{b}} = \frac{\pi}{6} \left[ (1 + 4b)^{\frac{3}{2}} - 1 \right].$$
2. Evaluate \( \iint_S [y, 2, xz] \cdot \mathbf{n} \, dA \), where \( S \) is defined by \( \begin{cases} y = x^2 \\ 0 < x < 2 \end{cases} \) and \( \mathbf{n} \) is pointing in the direction of \(-y\).

**Solution:** \( \mathbf{r}(u, v) = [u, u^2, v] \Rightarrow \mathbf{r}_u \times \mathbf{r}_v = [1, 2u, 0] \times [0, 0, 1] = [2u, -1, 0] \) (correct orientation). The integral thus converts to \( \int_0^3 \int_2^0 [2u, -1, 0] \) \( du \, dv = \frac{3}{2} \int_0^2 (2u^3 - 2) \) \( du \, dv = 3 \left[ \frac{u^4}{2} - 2u \right]_{u=0}^{u=3} = 12. \)

3. Find \( \iint_S [x^2, 0, 3y^2] \cdot \mathbf{n} \, dA \), where \( S \) is the \( \begin{cases} x > 0 \\ y > 0 \end{cases} \) portion of the \( x+y+z = 1 \) plane, and \( \mathbf{n} \) is pointing upwards.

**Solution:** \( \mathbf{r} = [u, v, 1-u-v] \Rightarrow \mathbf{r}_u \times \mathbf{r}_v = [1, 0, -1] \times [0, 1, -1] = [1, 1, 1] \) (correct orientation). \( \int_0^1 \int_0^{1-v} [u^2, 0, 3uv] \) \( \cdot [1, 1, 1] \) \( du \, dv = \frac{1}{12} \int_0^1 (1-v)^3 + 3(1-v)^2 v^3 \) \( du \) \( = \frac{1}{12} \).

4. Parametrize a circle of radius \( \rho = 5 \), centered on \( \mathbf{a} = [1, -2, 4] \), and normal to \( \mathbf{n} = [2, 0, -3] \).

**Solution:** In general, a circle is parametrized by: \( \mathbf{r}(t) = \mathbf{a} + \rho \mathbf{m}_1 \cos t + \rho \mathbf{m}_2 \sin t \), where \( \mathbf{m}_1 \) and \( \mathbf{m}_2 \) are unit vectors perpendicular to \( \mathbf{n} \) and to each other. They can be found by taking the cross product of \( \mathbf{n} \) and an arbitrary vector, then taking the cross product of the resulting vector and \( \mathbf{n} \), and normalizing both, thus: \( [2, 0, -3] \times [1, 0, 0] = [0, -3, 0] \) and \( [0, -3, 0] \times [2, 0, -3] = [9, 0, 6] \Rightarrow \mathbf{m}_1 = [0, -3, 0] = 3 = [0, -1, 0] \) and \( \mathbf{m}_2 = [9, 0, 6] \div \sqrt{9^2 + 0^2} = \left[ \frac{3}{\sqrt{13}}, 0, \frac{2}{\sqrt{13}} \right] \).

**Answer:** \( \mathbf{r}(t) = [1 + \frac{15}{\sqrt{13}} \sin t, -2 - 5 \cos t, 4 + \frac{10}{\sqrt{13}} \sin t] \) where \( 0 \leq t < 2\pi \). Subsidiary: To parametrize the corresponding disk: \( \mathbf{r}(u, v) = [1 + \frac{3u}{\sqrt{13}} \sin u, -2 - v \cos u, 4 + \frac{2v}{\sqrt{13}} \sin u] \) where \( 0 \leq u < 2\pi \) and \( 0 \leq v < 5 \).

5. Find the moment of inertia (with respect to the \( z \) axis) of a shell-like torus (parametrized earlier) of uniform mass density and total mass \( M \).

**Solution:** Recall that \( \mathbf{r}(u, v) = [(a + b \cos v) \cos u, (a + b \cos v) \sin u, b \sin v] \Rightarrow dA = b(a + b \cos v) \) \( du \, dv \) [done earlier] and \( d^2 = (a + b \cos v)^2 \Rightarrow \rho \int_0^{2\pi} \int_0^2 (a + b \cos v)^2 \) \( du \, dv = \rho b^2 \pi \int_0^{2\pi} (a^3 + 3a^2b \cos v + 3ab^2 \cos^2 v + b^3 \cos^3 v) \) \( dv = \frac{M}{4\pi ab} b^2 \pi [2\pi a^3 + 3ab^2 \pi] = M (a^2 + \frac{3}{2} b^2). \)

6. Repeat with a solid torus.
Solution: We replace \( r \) by \([(a + r \cos v) \cos u, (a + r \cos v) \sin u, r \sin v]\), where the new variables \( u, v \) and \( r \) \( (0 \leq r < b) \) can be also seen as orthogonal coordinates. For any orthogonal coordinates it is easy to find the Jacobian, geometrically, by \( dx \, dy \, dz \rightarrow r \, dv \cdot (a + r \cos v) \, du \, dv \, dr \) \( \Rightarrow \rho \int \int \int (a + r \cos v)^2 r (a + r \cos v) \, du \, dv \, dr = \rho 2\pi^2 \int_0^b \int_0^1 \int_0^{2\pi} r (2a^3 + 3ar^2) \, dr = \rho 2\pi^2 (a^3b^2 + \frac{3}{4}ab^4). \)

Similarly, the total volume is \( \int \int \int r (a + r \cos v) \, du \, dv \, dr = (2\pi)^2 a^2 b^2. \)

Answer: \( \frac{M}{2\pi a^2 b^2} 2\pi^2 (a^3b^2 + \frac{3}{4}ab^4) = M (a^2 + \frac{3}{4}b^2). \)

An alternate approach would introduce polar coordinates in the \((x, y)\)-plane, use \( 2\sqrt{b^2 - (r-a)^2} \) for the \( z \)-thickness and \( r^2 \) for \( d^2 \), leading to \( r \int \int \int r^2 \cdot r \, dr \, d\varphi = ... \) [verify that this leads to the same answer].

7. Consider the following solid \( \begin{cases} x > 0 \\ y > 0 \\ z > 0 \\ x + y + z < 1 \end{cases} \) of uniform density. Find:

(a) Center of mass.

Solution: To find its \( x \)-component we need to divide \( \int \int \int x \, dx \, dy \, dz = \int \int_0^1 \int_0^{1-z} \int_0^{1-y-z} x \, dy \, dz \cdot \int \int \int (1-y-z)^2 \, dy \, dz = \frac{1}{24} \) by the volume \( \int \int \int \, dx \, dy \, dz = \int_0^1 \int_0^{1-z} \int_0^{1-y-z} (1-y-z) \, dy \, dz = \frac{1}{6} \).

Answer: \( [\frac{1}{4}, \frac{1}{4}, \frac{1}{4}] \), as the \( y \) and \( z \)-components must have the same value as the \( x \)-component [obvious from symmetry].

(b) Moment of inertia with respect to \( [t, t, t] \) (the axis).

Solution: To find \( d^2 \) we project \( [x, y, z] \) into \([\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}]\) (unit direction of the axis), getting \( [x, y, z] \cdot [\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}] = \frac{x+y+z}{\sqrt{3}}. \) By Pythagoras, \( d^2 = x^2 + y^2 + z^2 - \left[\frac{x+y+z}{\sqrt{3}}\right]^2. \)

Answer: \( M \int \int \int x^2 + y^2 + z^2 - \frac{(x+y+z)^2}{3} \, dx \, dy \, dz = 4M \int \int \int \left[ x^2 + y^2 + z^2 - xy - xz - yz \right] \, dx \, dy \, dz = 12M \int \int \int \left[ x^2 - xz \right] \, dx \, dy \, dz \) [due to symmetry] =
116

\[ 12M \int_0^1 \int_0^{1-z} \left[ \frac{(1-y-z)^3}{3} - \frac{(1-y-z)^2}{2} z \right] dy \, dz = M \int_0^1 [(1-z)^4 - 2(1-z)^3 \, z] \, dz = M \left[ -\frac{(1-z)^5}{5} + 2\frac{(1-z)^4}{4} z + 2\frac{(1-z)^5}{20} \right]_0^1 = \frac{M}{10}. \]

8. A container is made of a spherical shell of radius 1 and height \( h \). Find:

(a) The shell’s surface area.

Solution: \( \mathbf{r}(u, v) = [u, v, -\sqrt{1-u^2-v^2}] \Rightarrow \mathbf{r}_u = [1, 0, \sqrt{1-u^2-v^2}] \) and \( \mathbf{r}_v = [0, 1, -\sqrt{1-u^2-v^2}] \Rightarrow dA = \sqrt{(1+\frac{u^2}{1-u^2-v^2})(1+\frac{v^2}{1-u^2-v^2}) - \frac{u^2 v^2}{(1-u^2-v^2)^2}} \, du \, dv \equiv \sqrt{\frac{1}{1-u^2-v^2}} \, du \, dv.

Answer: \( \iint_{u^2+v^2<h(2-h)} \frac{du \, dv}{\sqrt{1-u^2-v^2}} = \int_0^1 \int_0^{\sqrt{h(2-h)}} \frac{r \, dr \, d\varphi}{\sqrt{1-r^2}} \, d\varphi = 2\pi \left[-\sqrt{1-r^2}\right]_{r=0}^{\sqrt{h(2-h)}} = 2\pi \left[1 - \sqrt{1-h(2-h)}\right] = 2\pi h. \)

(b) The container’s volume:

Solution: Since the \( z \)-thickness (depth) equals \( \sqrt{1-x^2-y^2} - (1-h) \), all we need is \( \iint_{x^2+y^2<h(2-h)} \left[ \sqrt{1-x^2-y^2} - (1-h) \right] \, dx \, dy = \int_0^1 \int_0^{\sqrt{h(2-h)}} \left[ \sqrt{1-r^2} - 1 + h \right] \cdot r \, dr \, d\varphi = 2\pi \left[-\frac{1}{3} (1-r^2)^{3/2} - (1-h) \frac{r^2}{2} \right]_{r=0}^{\sqrt{h(2-h)}} \right]_{r=0}^{\sqrt{h(2-h)}} = 2\pi \left[-\frac{1}{3} (1-h)^3 - (1-h) \frac{h(2-h)}{2} + \frac{1}{3} \right] = \pi h^2 \left(1 - \frac{h}{3}\right). \)

9. Evaluate \( \oint_C [(x+y) \, dx + (2x-z) \, dy + (y+z) \, dz] \), where \( C \) is the closed curve consisting of three straight-line segments connecting \([2,0,0]\) to \([0,3,0]\), that to \([0,0,6]\), and back to \([2,0,0]\).

Solution: Applying the Stokes’ Theorem, which enables us to trade three line integrals (the three segments would require individual parametrization) for one surface integral, we first compute \( \text{Curl}(\mathbf{g}) = [2, 0, 1] \), then \( \mathbf{r}(u, v) = [u, v, 6-3u-2v] \) (note that \( \frac{2}{3} + \frac{1}{3} + \frac{2}{3} = 1 \) is the equation of the corresponding plane) \Rightarrow \mathbf{r}_u \times \mathbf{r}_v = [1, 0, -3] \times [0, 1, -2] = [3, 2, 1] \) which has the correct orientation.

Answer: \( \iint_{u>0, \, v>0, \, 3u+2v<6} [3, 2, 1] \, du \, dv = 7 \times \text{Area} = 7 \times \frac{2\times 3}{2} = 21 \)

[Verify by computing the line integral (broken onto three parts) directly].

10. Evaluate \( \oint_C [yz \, dx + xz \, dy + xy \, dz] \), where \( C \) is the intersection of \( x^2 + 9y^2 = 9 \) and \( z = 1 + y^2 \) oriented counterclockwise when viewed from above (in terms of \( z \)).

Solution: Applying the same Stokes’ Theorem, we get \( \text{Curl}(\mathbf{g}) \equiv [0, 0, 0] \).
Answer: 0.

We will verify this by evaluating the line integral directly: \( \mathbf{r}(t) = [3 \cos t, \sin t, 1 + \sin^2 t] \) parametrizes the curve \((0 < t < 2\pi) \Rightarrow \int_0^{2\pi} [(1 + \sin^2 t) \sin t, 3(1 + \sin^2 t) \cos t, 3 \sin t \cos t] \cdot [-3 \sin t, \cos t, 2 \sin t \cos t] \, dt = \int_0^{2\pi} [-3 \sin^2 t (1 + \sin^2 t) + 3(1 + \sin^2 t)(1 - \sin^2 t) + 6 \sin^2 t (1 - \sin^2 t)] \, dt = \int_0^{2\pi} (3 + 3 \sin^2 t - 12 \sin^4 t) \, dt = 2\pi \times (3 + 3 \times \frac{1}{2} - 12 \times \frac{3}{8}) = 0 \) [check].

Note that \( \int_0^{2\pi} \sin^2 n \, dt = \int_0^{2\pi} \cos^2 n \, dt = 2\pi \times \frac{1}{2} \times \frac{3}{4} \times \frac{5}{6} \times \frac{7}{8} \times \ldots \times \frac{2n-1}{2n} \).

11. In Physics we learned that the gravitational force of a 'solid' (i.e. 3-D) body exerted on a point-like particle at \( \mathbf{R} \equiv [X, Y, Z] \) is given by

\[
\mu \iiint_{\mathcal{V}} \rho(\mathbf{r}) \frac{\mathbf{r} - \mathbf{R}}{|\mathbf{r} - \mathbf{R}|^3} \, dV
\]

where \( \mu \) is a constant, \( \rho \) is the body’s mass density, and \( \mathcal{V} \) is its 'volume' (i.e. 3-D extent). [Here we are integrating a vector field in the componentwise (scalar) sense, i.e. these are effectively three volume integrals, not one].

Prove that, when the body is spherical (of radius \( a \)) and \( \rho \) is a function of \( r \) only (placing the coordinate origin at the body’s center), this force equals

\[
\mu M \cdot \frac{\mathbf{R}}{|\mathbf{R}|^3}
\]

where \( M \) is the body’s total mass.

Solution: First we notice that \( \frac{\mathbf{r} - \mathbf{R}}{|\mathbf{r} - \mathbf{R}|} \equiv \nabla_{\mathbf{R}} \frac{1}{|\mathbf{r} - \mathbf{R}|} \), where \( \nabla_{\mathbf{R}} \equiv \left[ \frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right] \).

This implies that \( \mu \iiint_{\mathcal{V}} \rho(\mathbf{r}) \frac{\mathbf{r} - \mathbf{R}}{|\mathbf{r} - \mathbf{R}|} \, dV \equiv \mu \nabla_{\mathbf{R}} \iiint_{\mathcal{V}} \rho(\mathbf{r}) \frac{1}{|\mathbf{r} - \mathbf{R}|} \, dV \)

leading to a lot easier integration (also, now we need one, not three integrals).

Evaluating \( \iiint_{\mathcal{V}} \rho(\mathbf{r}) \frac{\mathbf{r} - \mathbf{R}}{|\mathbf{r} - \mathbf{R}|} \, dV \) (the so called gravitational potential) in spherical coordinates yields \( \frac{a}{0} \frac{\rho(r)}{0} \int_0^{2\pi} \int_0^\pi \frac{r^2 \sin \theta \, d\varphi \, d\theta \, dr}{\sqrt{r^2 + R^2 - 2rr \cos \theta}} \) [note that \(|\mathbf{r} - \mathbf{R}| = \sqrt{x^2 + y^2 + (z - R)^2} \), where we have conveniently chosen the direction of \( \mathbf{R} \) (instead of the usual \( z \)) to correspond to \( \theta = 0 \)]. This further equals \( \frac{2\pi}{R} \int_0^a \rho(r) \cdot r \cdot [(R + r) - (R - r)] \, dr = \frac{4\pi}{R} \int_0^a \rho(r) r^2 \, dr = \frac{M}{R} \).
This proves our assertion, as \( \nabla R \frac{1}{R} = \nabla R \frac{1}{\sqrt{x^2+y^2+z^2}} = \)
\[
\left[ \frac{-x}{(x^2+y^2+z^2)^{\frac{3}{2}}}, \frac{-y}{(x^2+y^2+z^2)^{\frac{3}{2}}}, \frac{-z}{(x^2+y^2+z^2)^{\frac{3}{2}}} \right] = \frac{-R}{R^3}.
\]
Now try to prove the original statement directly (bypassing the potential), you should not find it too difficult.

12. **Optional**: Using Gauss Theorem (somehow indirectly, because of the singularity at \( R = r \), but this need not concern us here), one can show that

\[
\iiint_{\mathcal{V}} \text{Div}_R \left( \frac{r-R}{|r-R|^3} \right) \, dV_R = -4\pi
\]

for any \( \mathcal{V} \) containing \( r \), and equals 0 otherwise [note that the variable of both the integration, and the divergence operator, is \( R \), whereas \( r \) is considered a fixed parameter]. This implies that \( \text{Div}_R \left( \frac{r-R}{|r-R|^3} \right) \) [as a function of \( R \)] must be equal to zero everywhere except at \( R = r \), where its value becomes minus infinity (this can be easily verified by direct differentiation). This infinite 'blip' of its value contributes an exact, finite amount of \(-4\pi\) when the function is integrated. We can change this to +1 by a simple division, thus:

\[
\frac{1}{-4\pi} \text{Div}_R \left( \frac{r-R}{|r-R|^3} \right) \equiv \delta^{(3)}(R - r)
\]

defining the so called (3-D) **Dirac’s Delta Function**. Its basic property is

\[
\iiint_{\mathcal{V}} f(R) \cdot \delta^{(3)}(R - r) \, dV_R = f(r)
\]

We can now understand why

\[
\mathbf{F}(R) = \mu \iiint_{\mathcal{V}} \rho(r) \frac{r-R}{|r-R|^3} \, dV
\]

(the gravitational force of the previous example) implies

\[
\text{Div}_R \left( \mathbf{F}(R) \right) = -4\pi \mu \rho(R)
\]

When studying partial differential equations, one learns that the last equality also implies the previous one, and there are thus two equivalent ways of expressing the same law of Physics [in the so called **integral** and **differential** form, respectively]. This is essential for understanding Maxwell equations and their experimental basis. \( \otimes \)
Part III

COMPLEX ANALYSIS
Chapter 9  COMPLEX FUNCTIONS – DIFFERENTIATION

Preliminaries
We already know how to add, subtract, multiply [e.g. \((4 + 3i)(2 - 5i) = 8 + 6i - 20i + 15 = 23 - 14i\)] and divide [e.g. \(\frac{4+3i}{2-5i} = \frac{(4+3i)(2+5i)}{(2-5i)(2+5i)} = \frac{-7+26i}{29} = -\frac{7}{29} + \frac{26}{29}i\)] complex numbers. In this chapter, we will learn how to evaluate most of the usual functions using a complex argument (getting, in general, a complex answer). We will also investigate the issue of taking a derivative of any such function.

Basic Definitions

We reserve the letter \(z = x + iy\) for a complex number (soon to become a complex variable), where \(x = \text{Re}(z)\) is its REAL PART and \(y = \text{Im}(z)\) is its (purely) IMAGINARY PART [these are already two simple examples of functions of \(z\)].

Similarly, \(\bar{z} \equiv x - iy\) (some books use \(z^*\)) is the complex CONJUGATE of \(z\) [yet another function of \(z\)].

Proof: \((x_1 + iy_1)(x_2 + iy_2) = x_1x_2 - y_1y_2 - i(x_1y_2 + x_2y_1)\) and \((x_1 - iy_1)(x_2 + iy_2) = x_1x_2 - y_1y_2 - i(x_1y_2 + x_2y_1)\), which agree.

and

\[
\frac{z_1}{z_2} = \frac{\bar{z}_1}{\bar{z}_2}
\]

[proof similar]. Also note that \(z\bar{z} = x^2 + y^2\).

Geometrically, complex numbers are often represented as points of the \(x\)-\(y\) plane, leading to their so called polar representation:

\[
r = |z| \equiv \sqrt{x^2 + y^2}
\]

[the MAGNITUDE] and

\[
\theta = \pm \arctan(y/x)
\]

[the ARGUMENT, where the sign is chosen according to the quadrant of \(\theta\)]. The value of \(\theta\) is usually chosen from the \((-\pi, \pi]\) interval (the so called PRINCIPAL VALUE of the argument), but it has obviously infinitely many potential values \([\theta \pm 2\pi k; \text{where } k \text{ is an integer}]\). Conversely,

\[
z = r(\cos \theta + i \sin \theta)
\]

Using this representation, one can easily show that the product \(z_1 \cdot z_2 = r_1r_2(\cos \theta_1 + i \sin \theta_1)(\cos \theta_2 + i \sin \theta_2) = r_1r_2(\cos \theta_1 \cos \theta_2 - \sin \theta_1 \sin \theta_2) + i r_1r_2(\sin \theta_1 \cos \theta_2 + \sin \theta_2 \cos \theta_1) = r_1r_2[\cos(\theta_1 + \theta_2) + i \sin(\theta_1 + \theta_2)]\)
and similarly, the ratio $\frac{z_1}{z_2} =$

$$\frac{r_1}{r_2}[\cos(\theta_1 - \theta_2) + i \sin(\theta_1 - \theta_2)]$$

The first of these two formulas can be extended to any number of factors, further implying that an integer power of a complex number can be computed from

$$z^n = r^n[\cos(n\theta) + i \sin(n\theta)]$$

This also enables us to derive formulas of the following type:

$$\cos 5\theta + i \sin 5\theta = (\cos \theta + i \sin \theta)^5 = \cos^5 \theta - 10 \cos^3 \theta \sin^2 \theta + 5 \cos \theta \sin^4 \theta + i(5 \cos^4 \theta \sin \theta - 10 \cos^2 \theta \sin^3 \theta + \sin^5 \theta).$$

**EXAMPLES:** Find the region (of complex plane) which corresponds to:

1. $|z| \leq 1$.
   Solution: $\sqrt{x^2 + y^2} \leq 1$, i.e. the unit disk centered on $(0, 0)$.

2. $|z - 1| + |z + 1| = 3$.
   Solution: Square $|z - 1| = 3 - |z + 1|$ to get $(z - 1)(\bar{z} - 1) = 9 - 6|z + 1| + (z + 1)(\bar{z} + 1) \Leftrightarrow 6|z + 1| = 9 + 2(z + \bar{z})$. Square again getting: $36[(x + 1)^2 + y^2] = 81 + 72x + 16x^2 \Leftrightarrow 20x^2 + 36y^2 = 45$ [ellipse centered on $(0, 0)$].

3. $0 < \text{Im} \left(\frac{1}{z}\right) < 1$.
   Solution: Since $\text{Im} \left(\frac{1}{z}\right) = \text{Im} \left(\frac{x - iy}{x^2 + y^2}\right) = \frac{-y}{x^2 + y^2}$, we get $0 < -y < x^2 + y^2$, or $y < 0$ and $x^2 + (y + \frac{1}{2})^2 > \frac{1}{4}$, i.e. a set of points below the x-axis and outside the disk of radius $\frac{1}{2}$ centered on $(0, -\frac{1}{2})$.

**Introducing complex functions**

Any expression involving $z$ defines a complex function, e.g., $f(z) = z^2 + 3z$. In general, any such function will have complex values and can be thus expressed in terms of two (real) function, the real part of $f(z)$ and its (purely) imaginary part. These are usually called $u(x, y)$ and $v(x, y)$ respectively, each being a function of $x$ and $y$ (real arguments), i.e.

$$f(z) \equiv u(x, y) + i v(x, y)$$

**EXAMPLE:** $f(z) \equiv z^2 + 3z = x^2 + 2ixy - y^2 + 3x + 3iy = (x^2 - y^2 + 3x) + i(2xy + 3y) \equiv u(x, y) + iv(x, y)$. ■

**Derivative**

of a complex function is a fairly difficult concept, even though, its **definition** is seemingly the same as in the real case, namely

$$f'(z) = \lim_{\Delta \to 0} \frac{f(z + \Delta) - f(z)}{\Delta}$$
where $\Delta$ can approach zero from any (complex) direction. And only when all these limits agree, the function is called **differentiable** (at $z$), the value of the resulting derivative equal to this common limit.

Is there a simple way to establish that a given function is differentiable [we don’t want to compare infinitely many limits]? The answer is yes, the two real functions $u$ and $v$ must meet the following, so called **Cauchy-Riemann conditions**:

\[
\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}
\]

**Proof:** $f(z + \Delta) = u(x + \Delta_x, y + \Delta_y) + iv(x + \Delta_x, y + \Delta_y)$ where $\Delta = \Delta_x + i\Delta_y$. This can be expanded (generalized Taylor) as $u(x, y) + \frac{\partial u}{\partial x} \Delta_x + \frac{\partial u}{\partial y} \Delta_y + ... + iv(x, y) + i\frac{\partial v}{\partial x} \Delta_x + i\frac{\partial v}{\partial y} \Delta_y + ... \Rightarrow$

\[
\frac{f(z + \Delta) - f(z)}{\Delta} \approx \frac{\frac{\partial u}{\partial x} \Delta_x + \frac{\partial u}{\partial y} \Delta_y + i\frac{\partial v}{\partial x} \Delta_x + i\frac{\partial v}{\partial y} \Delta_y}{\Delta} = \frac{\Delta_x + i\Delta_y}{\Delta_x + i\Delta_y}
\]

Furthermore [we know from real analysis], all limits will agree when the 'horizontal' limit $\lim_{\Delta_x \to 0}$ and the 'vertical' limit $\lim_{\Delta_y \to 0}$ do. This implies that $\lim_{\Delta_x \to 0} \frac{\partial u}{\partial x} \Delta_x + \frac{\partial u}{\partial y} \Delta_y = -\frac{\partial u}{\partial y} - i\frac{\partial u}{\partial y}$, from which the Cauchy-Riemann conditions easily follow. \(\square\)

Note that

\[
f'(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} \equiv \frac{\partial v}{\partial y} - i\frac{\partial u}{\partial y}
\]

when the function is differentiable.

**EXAMPLE:**

\begin{itemize}
  \item $f(z) = z^2 = (x + iy)^2 = x^2 - y^2 + 2ixy$. Find $f'(z)$ [first check whether it exists].
  \end{itemize}

Solution (checking C-R): $\frac{\partial}{\partial x}(x^2 - y^2) \equiv \frac{\partial}{\partial y}(2xy) \sqrt{\text{ and } \frac{\partial}{\partial x}(2xy) \equiv -\frac{\partial}{\partial y}(x^2 - y^2)}$.

Answer: $f'(z)$ is then equal to $\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = 2x + 2iy \equiv 2z$. Note that we are getting the same answer as if $z$ were real. \(\blacksquare\)

Along these lines, one can show that in general all **polynomial** functions are differentiable, and that the corresponding derivative can be obtained by applying the usual $(z^n)' = nz^{n-1}$ rule.

This follows from the fact that, when $f_1$ and $f_2$ are differentiable, so is $f_1 + f_2$ [quite trivial to prove] and $f_1 \cdot f_2 \equiv (u_1u_2 - v_1v_2) + i(u_1v_2 + u_2v_1)$.

**Proof** (of the latter): $\frac{\partial u_1}{\partial x}u_2 + u_1\frac{\partial u_2}{\partial x} - \frac{\partial u_2}{\partial y}v_2 - v_1\frac{\partial v_2}{\partial x} - \frac{\partial u_1}{\partial y}v_2 - v_1\frac{\partial v_2}{\partial y} = \frac{\partial u_1}{\partial x}v_2 + u_1\frac{\partial v_2}{\partial x} + \frac{\partial u_2}{\partial x}u_2 + v_1\frac{\partial u_2}{\partial y} \Rightarrow \text{ the product rule still applies}. \(\square\)
Similarly, if \( f \) is differentiable, so is \( \frac{1}{f} \equiv \frac{u}{u^2 + v^2} - i\frac{v}{u^2 + v^2} \).

**Proof:** \( \frac{\partial u}{\partial x} (u^2 + v^2) - 2u(\frac{\partial u}{\partial x} u + \frac{\partial v}{\partial x} v) = -\frac{\partial v}{\partial y}(u^2 + v^2) + 2v(\frac{\partial u}{\partial y} u + \frac{\partial v}{\partial y} v) \) and \( \frac{\partial u}{\partial y} (u^2 + v^2) - 2u(\frac{\partial u}{\partial y} u + \frac{\partial v}{\partial y} v) = \frac{\partial v}{\partial x}(u^2 + v^2) - 2v(\frac{\partial u}{\partial x} u + \frac{\partial v}{\partial x} v) \) [each divided by \((u^2 + v^2)\)] ⇒ the quotient rule still applies.

And finally a **composition** \( [f_1(f_2(z))] \equiv u_1(u_2, v_2) + iv_1(u_2, v_2) \), sometimes denoted \( f_1 \circ f_2 \) of two differentiable functions is also differentiable.

**Proof:** \( \frac{\partial u_1}{\partial x} \frac{\partial u_2}{\partial x} + \frac{\partial u_1}{\partial y} \frac{\partial u_2}{\partial y} = \frac{\partial v_1}{\partial x} \frac{\partial v_2}{\partial y} + \frac{\partial v_1}{\partial y} \frac{\partial v_2}{\partial x} \) and \( \frac{\partial u_1}{\partial x} + \frac{\partial u_1}{\partial y} \frac{\partial v_2}{\partial y} = \frac{\partial u_1}{\partial x} \frac{\partial u_2}{\partial x} + \frac{\partial u_1}{\partial y} \frac{\partial v_2}{\partial x} \) where both \( u_1 \) and \( v_1 \) have \((u_2, v_2)\) as arguments ⇒ the chain rule still applies.

**In summary:**

All rational expressions (in \( z \)) are differentiable (everywhere, except when dividing by zero – the so called singularities), and the old differentiation formulas still apply (after the \( x \to z \) replacement). This provides us with a huge collection of differentiable functions (later to be extended further).

The natural question to ask now is: Are there any complex functions which are **not differentiable**? The answer is yes, aplenty as well.

**EXAMPLE:**

- \( f(z) = \bar{z} \equiv x - iy \). The first C-R condition requires \( \frac{\partial x}{\partial x} \equiv \frac{\partial (-y)}{\partial y} \) which is obviously not met. This function is **nowhere** differentiable.

Similarly one can verify that \(|z|, Re(z)\) and \(Im(z)\) are **nowhere differentiable** (since they have zero \(v(x, y)\) component). Thus, any expression involving any of these function is nowhere differentiable in consequence.

If a function is differentiable at a point, **and also at all points of its (open)** neighborhood, the function is called **analytic** at that point \(⇒\) a function can be differentiable at a single, isolated point, but it can be analytic only in some open region. This subtle distinction is not going to have much impact on us, we will simply take ’analytic’ as another name for ’differentiable’.

One interesting implication of C-R is that \( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0 \) and \( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0 \) (proof is trivial). Functions which meet the corresponding partial differential equation are called **harmonic**. This means that both \( u(x, y) \) and \( v(x, y) \) of an analytic function are harmonic, and reverse: given a harmonic function, we can make it a \( u(x, y) \) [or \( v(x, y) \)] of an analytic function by deriving the corresponding \( v(x, y) \) [or \( u(x, y) \)].

**EXAMPLE:**

Given \( u(x, y) = e^x \cos(y) \), find the corresponding \( f(z) \).

**Solution:** First we can easily check that the function is harmonic. Then we compute \( \frac{\partial \bar{z}}{\partial x} \equiv -\frac{\partial u}{\partial y} = e^x \sin(y) \) and \( \frac{\partial \bar{z}}{\partial y} \equiv \frac{\partial u}{\partial x} = e^x \cos(y) \), from which \( v(x, y) \) follows [by the procedure of solving exact differential equation]: \( v(x, y) = e^x \sin(y) + c_1(x) = e^x \sin(y) + c_2(y) \). Making these ’compatible’ we get \( v(x, y) = e^x \sin(y) \). Thus \( f(z) = e^x (\cos(y) + i\sin(y)) \equiv e^z \cdot e^{iy} = e^z \).
We have thus proved that the **exponential function** \( e^z \equiv \exp(z) \) is also analytic, with a derivative given by \((e^z)' = e^z\) [which follows easily from our last example].

\[ e^z \equiv \exp(z) \]

**Roots of \( z \)**

The function \( f(z) = \sqrt[2]{2} \equiv z^{1/2} \), where \( n \) is an integer, can in general have \( n \) possible (distinct) values, all given by

\[ \sqrt[n]{r}(\cos \frac{\theta}{n} + i \sin \frac{\theta}{n}) \]

(depending on the choice of \( \theta \)). We normally select its **principal** value, to make the answer unique.

This (principal-value) function is **analytic** everywhere except 0 and the negative \( x \)-axis (due to the discontinuity of \( \theta \) when crossing \(-x\)). The old

\[ (z^{1/2})' = \frac{1}{2}z^{-1/2} \]

formula still applies, in the analytic region.

We now return to our

**Exponential Function** \( f(z) = e^z \)

It is **periodic** in the following sense: \( f(z \pm 2k\pi i) \equiv f(z) \). The complex strip \(-\pi < y \leq \pi\) is called its **fundamental region** (everywhere else, the values of \( e^z \) are just repeated). Note that now the function can have negative values, e.g. \( e^{i\pi} = \cos \pi + i \sin \pi = -1 \). Its derivative equals to \( e^z \), as already mentioned.

**Proof:** In the previous example we saw that \( e^z \) was analytic everywhere. Its derivative when \( z \) is real is \( e^x \). The only way to extend this to an analytic expression is to add \( iy \) to \( x \) (making it \( z \)), as any other combination of \( x \) and \( y \) would not be analytic. (This argument applies to any analytic function, which means that we can always use the old formulas for differentiation, just replacing \( x \) by \( z \)). \( \Box \)

The corresponding **inverse function**

(to \( e^z \)) is \( w(z) \equiv \ln z \), a solution to \( z = e^w \) or, more explicitly, to

\[ z = e^{u+iv} = e^u(\cos v + i \sin v) \]

where \( w \equiv u + iv \). To solve this equation for \( u \) and \( v \), we must express \( z \) in its polar form, thus: \( r(\cos \theta + i \sin \theta) = e^u(\cos v + i \sin v) \Rightarrow \)

\[ u = \ln r \]

(this is the usual, **real** logarithm) and

\[ v = \theta \]
ln z is thus a multivalued function of z; to fix that, we define its principal value by taking $-\pi < \theta \leq \pi$ [in which case we call the function $Ln(z)$]. It is analytic everywhere except at 0 and negative real values [its derivative is given by the old $\frac{dLn(z)}{dz} = \frac{1}{z}$]. Since, in this manner, one can take a logarithm of any complex number, we can now find $Ln(-1) = Ln(1 \cdot e^{i\pi})$ [express the number in its polar form] $= 0 + i\pi$ (purely imaginary).

Using the $Ln$ function, we can now define

\[ f(z) \equiv z^a \]

where $a$ is also complex. This equals to $(e^{Ln(z)})^a = e^{aLn(z)}$ which is well (and uniquely) defined [in terms of its principal value – one could also define the corresponding multivalued function].

This function is thus analytic everywhere except the negative x axis and 0 [when $a$ is an integer, we need to exclude only 0 when $a < 0$, and nothing when $a > 0$]. Its derivative is of course the old $(z^a)' = az^{a-1}$.

Using this definition we can compute $i^i = e^{i(i\pi/2)} = e^{-\pi/2} = 0.20788$ [real!].

Similarly, one can also define the usual

\[ \sin z \equiv \frac{1}{2i}(e^{iz} - e^{-iz}) \]

and

\[ \cos z \equiv \frac{1}{2}(e^{iz} + e^{-iz}) \]

and the corresponding inverse functions $\arcsin(z)$ and $\arccos(z)$. Since these are of lesser importance to us, we are skipping the respective sections of your textbook.

**Chapter summary**

Complex differentiation is trivial for expressions contains $z$ only; we differentiate them as if $z$ were real.

This similarity of real and complex differentiation is a nontrivial (certainly not an automatic) consequence of the algebra of complex numbers; it is also the main reason why extending Calculus to complex numbers is so fruitful (as we will see in the next chapter).

As soon as we find $\bar{z}$, $|z|$, $Re(z)$ or $Im(z)$ in a definition of a complex function, the function is nowhere differentiable.
Chapter 10  COMPLEX FUNCTIONS – INTEGRATION

Similar to differentiation, complex integration of analytic functions will be shown to follow the formulas of real integration. But there is an extra bit of good news: the so called contour integration will make complex integration even easier, so that many real integrals can be simplified by going complex.

**Definition**

of a complex integral is similar to that of a line integral in a plane, with a small but essential modification: instead of taking the dot product of \( f \) and \( dr \), the complex function \( f(z) \) and the (two-component) infinitesimal element \( dz \) are multiplied using complex algebra, resulting in a complex (i.e. two-component) answer, thus:

\[
\int_C f(z) \, dz \equiv \int_C (u + iv)(dx + idy) = \int_C (udx - vdy) + i \int_C (vdx +udy)
\]

where \( C \) is some complex curve.

**Note** that this definition does not require \( f(z) \) to be analytic; we can thus integrate all (not just analytic) complex function.

The actual integration can be carried out by parametrizing \( C \) and performing the implied single-variable (\( dt \)) integration (we need two of them now).

**EXAMPLES:**

Evaluate:

1. \( \int_{C} \frac{dz}{z} \), where \( C \) is the unit circle centered at 0, traversed counterclockwise.

   **Solution:** Parametrize \( z \) by \( z = e^{it} \) where \( 0 \leq t < 2\pi \) [this form is more convenient than the more explicit but equivalent \( z = \cos t + i \sin t \)]. Since \( \frac{dz}{dt} = ie^{it} \), we can replace \( dz \) by \( ie^{it} \, dt \) [and \( \frac{1}{z} \) by \( e^{-it} \)].

   **Answer:** \( \int_{0}^{2\pi} e^{-it} \cdot ie^{it} \, dt = i \int_{0}^{2\pi} dt = 2\pi i. \)

   [Using the more explicit form of \( z \), we would have to struggle with \( \int_{0}^{2\pi} \frac{-\sin t + i \cos t}{\cos t + i \sin t} \, dt \) to get the same answer].

2. \( \int_{C} \text{Re}(z) \, dz \), where \( C \) is the straight-line segment from 0 to 1 + i.

   **Solution:** \( z = t + it \) with \( t \in (0, 1) \Rightarrow \text{Re}(z) = t \) and \( dz = (1 + i) \, dt. \)

   **Answer:** \( \int_{0}^{1} t(1 + i) \, dt = (1 + i) \left[ \frac{t^2}{2} \right]_{t=0}^{1} = \frac{1}{2} + \frac{i}{2}. \)
3. \( \int_{C} (z - z_0)^m dz \), where \( m \) is an integer (of either sign), \( z_0 \) is a complex constant, and \( C \) is a counterclockwise circle of radius \( \rho > 0 \) centered at \( z_0 \) (this is an extension of Example 1).

Solution: \( z = z_0 + \rho e^{it} \) with \( t \in (0, 2\pi) \) \( \Rightarrow \int_{0}^{2\pi} \rho^m e^{imt} \cdot \rho e^{it} dt = \rho^{m+1} \int_{0}^{2\pi} e^{i(m+1)t} dt = \rho^{m+1} \int_{0}^{2\pi} (\cos((m+1)t) + i \sin((m+1)t)) dt = 0 \), with one important exception: when \( m = -1 \), we get \( \rho(2\pi) \cdot \int_{0}^{2\pi} (\cos 0 + i \sin 0) dt = 2\pi i \) [remember this result, it is of special importance].

Integrating analytic functions

An analytic function \( f(z) \) can be integrated by first finding its anti-derivative (as if \( z \) were real, i.e. all of the old formulas and techniques still apply), then evaluating it at the first and the last point of \( C \) (each being a complex number), and finally subtracting the former from the latter (the same old procedure).

Proof: \( \int_{C} (u+iv)(dx+idy) = \int_{C} (udx-vdy) + i \int_{C} (vdx+udy) \). The last two integrals are, effectively, line integrals in a plane (with \( r \equiv [u, -v] \) and \( r \equiv [v, u] \), respectively). They are both path independent iff the C-R conditions are met. Furthermore, when integrated via their 'potentials', they yield

\[
g(x, y)|^{(x_1, y_1)}_{(x_0, y_0)} + i \ h(x, y)|^{(x_1, y_1)}_{(x_0, y_0)}
\]

where \( g \) and \( h \) clearly meet the C-R conditions as well [we use 0 and 1 as indices of the first and last point of \( C \), respectively]. We can thus write the result as \( F(z)|_{z_0}^{z_1} \) where \( F \) is analytic, and must agree with the usual antiderivative when \( z \) is real. The only way to extend a real function \( F(x) \) to an analytic function is by \( x \to z \). \( \square \)

EXAMPLE: Find \( \int_{C} z^2 dz \), where \( C \) is a straight line segment from 0 to 2 + i.

Solution: \( \int_{C} z^2 dz \) \( = \frac{z^3}{3} \bigg|_{z=0}^{2+i} = \frac{(2+i)^3}{3} = \frac{8+3\times4i+3\times2^2+i^3}{3} = \frac{2}{3} + \frac{11}{3} i \) [note that the result is path independent]. \( \blacksquare \)

Integration is thus very simple for fully analytic functions (i.e. analytic everywhere, they are also called entire functions). Things get a bit more interesting when the function has one or more

\( \blacktriangleright \) Singularities \( \blacktriangleright \)

EXAMPLE: Integrating \( \int_{C} \frac{dz}{z} \), where \( C \) is a curve starting at \(-1-i\) and ending at \(1+i\) yields two different results \([-\pi i \text{ and } \pi i \], respectively\] depending on whether we pass to the left or to the right of 0 [we have to avoid 0, since the function is singular there].
Thus, we have to conclude that the anti-derivative technique \( \int_C \frac{dz}{z} = Ln(z)|_{z=-1-i}^{1+i} \) returns the correct answer only when \( C \) does not cross the \(-x\) axis (the so called "cut"), since \( Ln(z) \) is not analytic there.

To correctly evaluate the other integral (when \( C \) passes to the left of \( 0 \)), we must define our own \( \ln(z) \) function whose cut follows \(+x\) (rather than \(-x\)). This simply means selecting \( \theta \) from the \([0, 2\pi)\) interval, rather than the 'principal' \((-\pi, \pi]\). Using this function, \( \ln(z)|_{z=-1+i}^{1+i} \) returns the correct answer (for all paths to the left of \( 0 \)).

**In general:** \( \oint_C f(z) \, dz \) is **path-independent** as long as \( C \) is modified **without** crossing any singularity of \( f(z) \).

This implies:

When \( C \) is closed

\[
\oint_C f(z) \, dz \equiv 0
\]

provided that there are **no singularities** of \( f(z) \) **inside** \( C \).

\( \oint \) is a **notation** we use when integrating over a closed curve, this is also called **contour integration**.

Now, the **main question** is: How do we evaluate \( \oint_C f(z) \, dz \) with **some singularities** inside a closed curve \( C \)?

This leads to a very important technique of the so called **Contour integration**

We are in a good position to figure out the answer:

Each such 'contour' \( (C) \) can be subdivided into as many pieces (also closed curves) as there are singularities (with one singularity in each).

Each of these pieces can be then continuously modified (without crossing any singularity) until its singularity is encircled. This will not change the value of the original integral, which has now become a sum of several 'circle' integrals.

And the value of each of these can be computed by expanding \( f(z) \) at \( z_0 \) [the corresponding singular point], thus:

\[
f(z) = \ldots + \frac{a_{-3}}{(z-z_0)^3} + \frac{a_{-2}}{(z-z_0)^2} + \frac{a_{-1}}{z-z_0} + a_0 + a_1(z-z_0) + a_2(z-z_0)^2 + \ldots
\]

[the so called **Laurent series**, where \( a_i \) are constant coefficients], and then integrating term by term.

We already know that only the \( \frac{a_{-1}}{z-z_0} \) term will contribute a nonzero value of \( 2\pi i a_{-1} \). The \( a_{-1} \) coefficient of the Laurent expansion is thus of a rather special importance. It is called the **residue** of \( f(z) \) at \( z_0 \), and is usually denoted by \( \text{Res}(f) \).

\( \text{Res}(f) \) is **important** and will be used in various contexts.

**The final result:** \( \oint_C f(z) \, dz \) is equal to \( 2\pi i \), multiplied by the **sum of residues** of all singular points of \( f(z) \) **inside** \( C \) [no actual integration is thus necessary].
A special (but very common and important) case is

\[ f(z) = \frac{g(z)}{(z - z_0)^m} \]

where \( g(z) \) is analytic at \( z_0 \) [i.e. the singularity at \( z_0 \) is due to an explicit division by \( (z - z_0)^m \)]. We can thus expand \( g(z) \) at \( z_0 \) in a regular (Taylor) manner, divide the result by \( (z - z_0)^m \), and clearly see that

\[ \text{Res}(f) = \frac{g^{(m-1)}(z_0)}{(m-1)!} \]

EXAMPLES:

1. The residue of \( e^z \) at \( z = 0 \) is \( (e^z)|_{z=0} = 1 \). We can thus easily evaluate \( \oint_C e^{z^2} \, dz \) [where \( C \) is any contour encircling 0 counterclockwise] as \( 2\pi i \).

2. Find \( \oint_C \frac{1+z^2}{z^2-1} \, dz \), where \( C \) is (counterclockwise):

   (a) The unit circle centered at 1.
   Solution: The only singularities of \( \frac{1+z^2}{(z+1)(z-1)} \) are at \( z = -1 \) (residue equal to \( \frac{1+(-1)^2}{-1-1} = -1 \)) and at \( z = 1 \) (residue equal to \( \frac{1+1^2}{1+1} = 1 \)).
   Answer: \( 2\pi i \).

   (b) The unit circle centered at \( -1 \).
   Answer: \( 2\pi i \times (-1) = -2\pi i \).

   (c) The unit circle centered at \( i \).
   Answer: 0 (no singularity is inside this \( C \)).

   (d) The circle of radius 3, centered at 0.
   Answer: \( 2\pi i \times (1 - 1) = 0 \) [both singularities are inside this \( C \)].

3. Identify the singularities (and find the corresponding residues) of \( \frac{z^2-1}{z^2+1} \).
   Solution: \( \frac{z^2-1}{(z+1)(z-1)} \) has singularities at \( z = -i \) (residue: \( \frac{(-i)^2-1}{-i-1} = -i \)) and at \( z = i \) (residue: \( \frac{i^2-1}{i+1} = i \)).

4. Same for \( \frac{z^2+1}{(4z-1)^2} \).
   Solution: \( \left[ \frac{z^2+1}{16} \right] = \frac{1}{32} \).

5. Same for \( \frac{(z+4)^3}{z^4+5z^3+6z^2} \).
   Solution: \( \frac{(z+4)^3}{z^4(z+2)(z+3)} \) has singularities at \( z = -2 \) (residue: \( \frac{(-2+4)^3}{(-2)^2} = 2 \)), at \( z = -3 \) (residue: \( \frac{(-3+4)^3}{(-3)^2} = -\frac{1}{9} \)) and at \( z = 0 \) (residue: \( \left[ \frac{(z+4)^3}{z^2+5z+6} \right] \) at \( z = 0 > 3 \times 4^2 \times 5 \times 4^3 - \frac{8}{9} \)).
The general case of
\[ f(z) = \frac{g(z)}{h(z)} \]
where \( h(z_0) = 0 \). The corresponding residue (at \( z = z_0 \)) equals
\[
\lim_{z \to z_0} \left[ \frac{f(z)(z - z_0)^m}{(m - 1)!} \right]^{(m-1)}
\]
where \( m \) is the order of the \( z_0 \) root. If this can not be established in advance, one has to try \( m = 1, m = 2, m = 3, \ldots \), until the limit is finite (a larger value of \( m \) would still yield the correct answer, but with a lot more effort).

**EXAMPLES:**

Find the residue of

1. \( \frac{e^z}{\cos \pi z} \) at \( z = \frac{1}{2} \).

   Solution: \( \lim_{z \to \frac{1}{2}} z \cdot \frac{e^z}{\cos \pi z} \) equals, by L'Hopital rule, \( \left. \frac{e^z + (z - \frac{1}{2})2ze^z}{-\pi \sin \pi z} \right|_{z = \frac{1}{2}} = -\frac{e^{\frac{1}{2}}}{\pi} = -0.40872 \).

2. \( \frac{2z + 1}{(1 - e^z)^2} \) at \( z = 0 \).

   Solution: This looks like a second-order singularity, so we first find \( \left( \frac{z^2(2z + 1)}{(1 - e^z)^2} \right)' = 2\left( \frac{z^2}{1 - e^z} \right)^2 + 2 \cdot \frac{z}{1 - e^z} \cdot \frac{1 - e^z + ze^z}{(1 - e^z)^2} \cdot (1 + 2z) \) and then take the \( z \to 0 \) limit. Using L'Hopital rule gives (individually) \( \frac{z}{1 - e^z} \to \frac{1}{-e} \) at \( z = 0 \) and \( \frac{1 - e^z + ze^z}{(1 - e^z)^2} \to \frac{-2e^z}{2(e - 2e^z)} \) at \( z = 0 \).

   Answer: \( 2(-1)^2 + 2 \times (-1) \times \frac{1}{2} \times 1 = 1 \).

   An alternate, and in many cases easier, approach is to directly expand:
   \[
   \frac{2}{(z - \frac{1}{2})^2} = (1 + 2z) \cdot \frac{1}{2} \cdot (1 + \frac{2}{3} + \frac{2}{6} + \ldots) = (1 + 2z) \cdot \frac{1}{2} \cdot (1 - \frac{2}{3} + \frac{3}{4} + \ldots) = \frac{1}{z^2} + \frac{1}{z} - \frac{10}{12} + \ldots
   \]

3. \( \frac{1}{1 - \cos z} \) at \( z = 0 \).

   Solution: Since \( \cos z = 1 - \frac{z^2}{2} + \frac{z^4}{24} - \ldots \) we can deduce that the singularity is of second order. The residue is thus \( \left( \frac{z^2}{1 - \cos z} \right)' = \frac{2z(1 - \cos z) + z^2 \sin z}{(1 - \cos z)^2} \), after we take the \( z \to 0 \) limit. This requires applying L'Hopital rule four times (relatively easy if we differentiate and substitute at the same time), resulting in 0 for the numerator, and \( 2 \times \binom{3}{1} \times \cos 0 \times \cos 0 = 6 \) for the denominator (we need to know that it is non-zero).

   Answer: 0.

   Alternately, we expand \( \left( \frac{z^2}{2} - \frac{z^4}{24} - \ldots \right)^{-1} = \frac{2}{z^2} \left( 1 - \frac{z^2}{12} + \ldots \right)^{-1} = \frac{2}{z^2} \left( 1 + \frac{z^2}{12} + \ldots \right) = \frac{1}{z^2} + \frac{1}{6} + \ldots \), yielding the same result.
4. \( \frac{1}{1 + z^3} \) at \( z_a = \frac{1}{2} + i \frac{\sqrt{3}}{2} \).

Solution: Even though this is a rational function, we may now prefer to do \( \lim_{z \to z_a} \left( \frac{z - z_a}{1 + z^3} \right) = [L'Hopital] \left| \frac{1}{3z^2} \right|_{z=z_a} = -\frac{2}{3} \) [since \( z_a^3 = -1 \)] = \(-\frac{1}{6} - i \frac{\sqrt{3}}{6} \).

5. \( \frac{1}{e^s - 1 - z} \) at \( z = 0 \).

Solution: Here we would rather expand: \( \frac{1}{z} \left( 1 + \frac{z}{3} + ... \right)^{-1} = \frac{2}{z} \cdot (1 - \frac{z}{3} + ...) \Rightarrow \)

Answer: \(-\frac{2}{3} \) [coefficient of \( \frac{1}{z} \)].

Applications
to evaluating real integrals of several special types:

\[
\text{Rational Functions of } \sin t \text{ and/or } \cos t
\]

integrated over a full-period interval, i.e. from \( c \) to \( c + 2\pi \) where \( c \) is any real number (usually equal to 0).

We introduce \( z = e^{it} (\Rightarrow \sin t = \frac{e^{iz} - e^{-iz}}{2i}, \cos t = \frac{e^{iz} + e^{-iz}}{2} \text{ and } dt = \frac{dz}{iz} ) \) and integrate the corresponding complex function over \( C_0 \) [the counterclockwise unit circle centered at 0] via contour integration (i.e. by finding all singularities inside this circle and adding their residues \( \times 2\pi i \); note that the final answer must be real).

EXAMPLES:

1. \( \frac{2\pi}{5 - 3 \cos t} \)

\[
\int_0^{2\pi} \frac{dt}{5 - 3 \cos t} = \frac{2\pi}{C_0} \int_0^{2\pi} \frac{dz}{iz(5 - 3 \frac{z - 1}{2})} = \frac{2\pi}{C_0} \int_0^{2\pi} \frac{dz}{i(z - 3)(z - 1)} = -2\pi \times \frac{1}{3 - 3} \times 2\pi i = \frac{8\pi}{3} \text{ [the only singularity inside } C_0 \text{ is at } z = \frac{1}{3}, \text{ the corresponding residue].}
\]

Note that \( \int_0^{4\pi} \frac{dt}{5 - 3 \cos t} \) would be simply twice as large. Similarly, \( \int_0^{\pi} \frac{dt}{5 - 3 \cos t} \) would yield half the value, because \( \frac{1}{5 - 3 \cos t} \) is an even function of \( t \).

2. \( \frac{1 + \sin \theta}{3 + \cos \theta} \)

\[
\int_0^{\pi} \frac{d\theta}{3 + \cos \theta} = \frac{1 + \frac{z - 1}{2i}}{\frac{3 + \frac{z - 1}{2}}{i}} \cdot \frac{dz}{i} = -\frac{2\pi}{C_0} \frac{2i z - 1}{(z + 3 + \sqrt{8})(z + 3 - \sqrt{8})} dz.
\]

The singularities are at \( z = 0 \) [residue: \(-1\)] and \( z = -3 + \sqrt{8} \) [residue: \( \frac{(3 + \sqrt{8})^2 + 2i(-3 + \sqrt{8}) - 1}{(-3 + \sqrt{8}) \times 2\sqrt{8}} = 1 + \frac{i}{\sqrt{8}} \)].

Answer: \(-\frac{i}{\sqrt{8}} \times 2\pi i = \frac{\pi}{2\sqrt{8}} \).
3. \[ \int_0^\pi \frac{\sin^2 t - 2 \cos t}{2 + \cos t} dt = \frac{1}{2} \int_0^{2\pi} \frac{\sin^2 t - 2 \cos t}{2 + \cos t} dt \] [even function] =

\[ \frac{1}{2} \oint_{C_0} \frac{(z-z^{-1})^2}{2 + z + z^{-1}} \cdot \frac{dz}{iz} = -\frac{1}{4i} \oint_{C_0} \frac{z^4 + 4z^3 - 2z^2 + 4z + 1}{z^2(z^2 + 4 + 1)} dz. \]

The singularities are at \( z = 0 \) (residue: \( z = 0 \)), and at \( z = -2 + \sqrt{3} \) (residue: \( z^2(2z + 4 + 1) \)) =

\[-\frac{z^2 - 2z^2 - z^2}{z^2 \times 2\sqrt{3}} = -\frac{2}{\sqrt{3}}, \text{ where } z_s \equiv -2 + \sqrt{3}; \text{ note that } z_s^2 + 4z_s \equiv -1 \text{ and } 4z_s + 1 \equiv -2.\]

Answer: \( -\frac{1}{4i} \times \frac{2}{\sqrt{3}} \times 2\pi i = \frac{\pi}{\sqrt{3}}. \]

4. \[ \int_0^{\pi} \frac{dx}{a + \cos^2 x} \text{ where } a > 0. \]

Solution: \[ \int_{-\pi}^{\pi} \frac{dx}{-a + (e^{ix} + e^{-ix})^2} = \int_{-\pi}^{\pi} \frac{dz}{iz} = \int_{C_0} \frac{dz}{4az^2 + (1 + z^2)^2}. \]

The singularities are the roots of \( z^4 + (2 + 4a)z^2 + 1 = 0 \) namely: \( z^2 = -(1 + 2a) \pm 2\sqrt{a + a^2} \) [the minus sign puts us outside the unit circle, the plus sign results in a negative value between 0 and -1, the proof is simple].

The roots we need are thus \( z_{1,2} = \pm i \sqrt{1 + 2a - 2\sqrt{a + a^2}}, \text{ the corresponding factorization of the function’s denominator results in } \]

\[ (z - z_1)(z - z_2)(z^2 + 1 + 2a + 2\sqrt{a + a^2}). \]

The residues are thus \( \frac{z_1}{(z_1 - z_2) \cdot 4\sqrt{a + a^2}} \) and \( \frac{z_2}{(z_2 - z_1) \cdot 4\sqrt{a + a^2}} \) The sum is \( \frac{1}{4\sqrt{a + a^2}} \), multiplied by \( 2\pi i \times \frac{2}{i} \) equals \( \frac{\pi}{\sqrt{a + a^2}}. \)

►Rational Function of \( x \)◄

(the denominator must be a polynomial of at least two degrees higher than that in the numerator), integrated from \(-\infty \) to \( \infty \).

We replace any such integral by a complex integral over \( z (\equiv x) \), from \(-R \) to \( R \), extended by the half circle \( Re^{it} \) \( t \in (0, \pi) \) to make the contour closed. One can show that, in the \( R \to \infty \) limit, the half circle’s contribution tends to 0, and one thus obtains the correct answer to the original integral.

**Proof:** One can easily show that the magnitude of a complex integral cannot exceed the maximum magnitude of the integrand, multiplied by the length of \( C \). In our case, the length of the half circle is \( \pi R \), and the function’s magnitude has an upper bound (based on the triangular inequality) equal to a polynomial in \( R \) over another polynomial of at least two degrees higher. The product of the \( \pi R \) length and this upper bound tends to 0 as \( R \to \infty \).
Algorithm: To find the value of the original integral, one thus has to replace $x$ by $z$ in the integrand, find all its singularities in the upper half plane ($y > 0$), add their residues, and multiply by $2\pi i$.

Note that the answer must be real.

EXAMPLES:

Evaluate:

1. $\int_{-\infty}^{\infty} \frac{dx}{1+x^2}$.

Solution: $\frac{\sqrt{z}}{z}$ are the function’s singularities, $\pm \frac{1+i}{\sqrt{2}}$ being in the upper half plane. The corresponding residues are found from $\lim_{z \to z_s} \left( \frac{\sqrt{z}}{z} \right) = [\text{by L'Hospital}] \frac{1}{4z_s^2} \equiv -\frac{i}{2}$ [since $z_s^4 \equiv -1$], where $z_s$ is either one of the two singularities. Substituting $z_s = \frac{1+i}{\sqrt{2}}$ this yields $-\frac{1+i}{4\sqrt{2}}$, substituting $z_s = \frac{-1+i}{\sqrt{2}}$ we get $\frac{1-i}{4\sqrt{2}}$.

Answer: $\left(-\frac{1+i}{4\sqrt{2}} + \frac{1-i}{4\sqrt{2}}\right) \times 2\pi i = \frac{\pi}{\sqrt{2}}$.

2. $\int_{-\infty}^{\infty} \frac{dz}{1+z^2}$.

Solution: The relevant singularities are $\frac{\sqrt{z}}{z} = \pm \frac{\sqrt{3}+i}{2}$ and $i$. The corresponding residues are given by $\lim_{z \to z_s} \left( \frac{\sqrt{z}}{z} \right) = \frac{-1}{6z_s^2} = -\frac{i}{2}$ [since $z_s^6 \equiv -1$]. The sum of the three residues is thus $-\frac{2+i}{6} = -\frac{i}{3}$.

Answer: $-\frac{i}{3} \times 2\pi i = \frac{2}{3}\pi$.

3. $\int_{0}^{\infty} \frac{dx}{(1+x^2)^{\pi}}$.

Solution: This equals to $\frac{1}{2} \int_{-\infty}^{\infty} \frac{dx}{(1+x^2)^{\pi}}$ since the integrand is even. Furthermore, $\frac{1}{(z-i)^3(z+i)^3}$ has only one $y > 0$ singularity ($z_s = i$), which is of the third order.

The corresponding residue is thus $\frac{1}{2} \left( \frac{1}{(z+i)^3} \right)_{z=i} = 6(z+i)^{-5}|_{z=i} = \frac{6}{(2i)^5} = \frac{3}{16i}$.

Answer: $\frac{3}{16i} \times 2\pi i = \frac{3}{8}\pi$.

4. $\int_{0}^{\infty} \frac{1+x^2}{1+x^4} \, dx$.

Solution: $= \frac{1}{2} \int_{-\infty}^{\infty} \frac{1+x^2}{1+x^4} \, dx$ [even]. The relevant singularities are at $\pm \frac{1+i}{\sqrt{2}}$, the corresponding residues are given by $\lim_{z \to z_s} \left( \frac{\sqrt{z}}{z} \right) = \frac{1+z^2}{4z^3} = -\frac{1+z^2}{4} \cdot z_s$ [as $z_s^4 \equiv -1$]. Their sum equals $-\frac{1}{4} \cdot \frac{1+i}{\sqrt{2}} - \frac{1}{4} \cdot \frac{1-(-1+i)^2}{\sqrt{2}} = -\frac{(1+i)^2}{4\sqrt{2}} - \frac{(1-i)(-1+i)}{4\sqrt{2}} = -\frac{2i}{4\sqrt{2}} - \frac{2i}{4\sqrt{2}} = -\frac{i}{\sqrt{2}}$. 
Answer: \( \frac{1}{2} \cdot \frac{-i}{\sqrt{2}} \cdot 2\pi i = \frac{\pi}{\sqrt{2}}. \)

We now consider integrating, from \(-\infty\) to \(\infty\), the same type of rational expression as in the previous case, further

\[ \text{Multiplied by } \sin kx \text{ or } \cos kx. \]

Now we replace \( \sin kx \) (or \( \cos kx \)) by \( e^{ikz} \), find the value of the corresponding integral in the same manner as before, and then take the imaginary (or real) part of the answer.

EXAMPLES:

Find the value of:

1. \[ \int_{-\infty}^{\infty} \frac{\sin 2x}{x^2 + x + 1} \, dx. \]

Solution: \( = Im \int_{-\infty}^{\infty} \frac{e^{2iz}}{z^2 + z + 1} \, dz. \) The last integral can be evaluated by adding the \( y > 0 \) residues of the integrand and multiplying their sum by \( 2\pi i. \) The only contributing singularity is at \( z_s = -\frac{1}{2} + i\frac{\sqrt{3}}{2}, \) the corresponding residue equals \( \frac{e^{2iz_s}}{(z_s - \bar{z}_s)} = \frac{e^{-\sqrt{3}(\cos 1 - i \sin 1)}}{i\sqrt{3}}. \) Multiplying this by \( 2\pi i \) and keeping the imaginary part only results in \(-\frac{2\pi}{\sqrt{3}} e^{-\sqrt{3}} \sin 1 = -0.54006. \)

2. Similarly, \[ \int_{-\infty}^{\infty} \frac{\cos 2x}{(x^2 + 4)^2} \, dx = Re \int_{-\infty}^{\infty} \frac{e^{2iz}}{(z^2 + 4)^2} \, dz, \] with \( z_s = 2i \) being (the only relevant) second-order singularity, having a residue of \( \left( \frac{e^{2iz}}{(z + 2i)^2} \right)' \bigg|_{z=2i} = \frac{2ie^{2iz}(z + 2i)^2 - 2(z + 2i)e^{2iz}}{(z + 2i)^4} \bigg|_{z=2i} = \frac{2i(4i)^2 - 2(4i)}{(4i)^4} \cdot e^{-4} = -\frac{5i}{32} \cdot e^{-4}. \) Multiplying this by \( 2\pi i \) and keeping the real part of the result yields \( 0.017981. \)

EXAMPLES:

1. \( I \equiv \int_{-\infty}^{\infty} \frac{e^{mx}}{1 + e^z} \, dx \) where \( 0 < m < 1. \)

Solution. We make \( x \) complex \( (x \to z) \) and integrate the same function over the following contour \([\text{a collection of four straight-line segments which we call } C_1, C_2, C_3 \text{ and } C_4]: -R \to R \text{ (real), } R \to R + 2\pi i, R + 2\pi i \to -R + 2\pi i, \text{ and } -R + 2\pi i \to -R. \) One can show that, on \( C_3, \) the integrand equals \( \frac{e^{mx}}{1 + e^z} \equiv e^{2\pi mi} \cdot \frac{e^{mx}}{1 + e^z}. \) Since \( \left| \frac{e^{mx}}{1 + e^z} \right| \leq \frac{e^{mR}}{e^R - 1} \) on \( C_2 \) and \( \left| \frac{e^{mx}}{1 + e^z} \right| \leq \frac{e^{-mR}}{1 - e^{-R}} \) on \( C_3, \) their contributions disappear in the \( R \to \infty \) limit. In the same limit, the
contribution of $C_1$ yields $I$, and that of $C_3$ results in $-e^{2\pi mi} \cdot I$ [since we are going backwards].

The contour integral has only one singularity at $z = \pi i$, with the residue equal to 
\[ \lim_{z \to \pi i} \left( \frac{(z - \pi i) e^{\pi z}}{1 + e^z} \right) = \frac{e^{\pi \pi i}}{e^\pi i} = -e^{\pi mi}. \] Its value is thus $-2\pi i e^{\pi mi}$; the value of our $I$ must be this, divided by $1 - e^{2\pi mi}$, i.e. 
\[ \frac{-2\pi i e^{\pi mi}}{1 - e^{2\pi mi}} \equiv \frac{2\pi i}{e^{\pi mi} - e^{-\pi mi}} = \frac{\pi}{\sin(m\pi)}. \]

2. $I \equiv \int_0^\infty \frac{x^{p-1}}{1 + x} \, dx$ where $0 < p < 1$. [Note that this integral can be converted to the previous one by a $x = e^u$ substitution, but we will pretend not to notice].

Solution: This time we use $C_1$: $ri$ to $R + ri$ (straight line), $C_2$: $R + ri$ to $R - ri$ (nearly a full circle centered on 0), $C_3$: $R - ri$ to $-ri$ (straight line), and $C_4$: $-ri$ to $ri$ (a semicircle centered at 0). Since $|\frac{z^{p-1}}{1 + z}| \leq \frac{r^{p-1}}{1-r}$ on $C_2$, this contribution disappears in the $R \to \infty$ limit, since $|\frac{z^{p-1}}{1 + z}| \leq \frac{r^{p-1}}{1-r}$ on $C_4$, ditto for the $r \to 0$ limit.

Secondly, on $C_3$, 
\[ \frac{z^{p-1}}{1 + z} = \frac{e^{(p-1)ln(x+2\pi i)}}{1 + x} \equiv e^{2\pi pi} \cdot \frac{x^{p-1}}{1 + x} \] [in the $r \to 0$ limit], so its contribution is $-e^{2\pi pi} \times I$. The only singularity inside the contour is at $z = -1$, with the residue of $(-1)^{p-1} = e^{i\pi(p-1)}$.

Answer: 
\[ \frac{-2\pi i e^{i\pi(p-1)}}{1 - e^{2\pi pi}} = \frac{-2\pi i e^{i\pi}}{1 - e^{2\pi pi}} = \frac{\pi}{\sin(p\pi)}. \]

3. Compute $\int_0^\infty \cos(x^2) \, dx$ and $\int_0^\infty \sin(x^2) \, dx$ as the real and imaginary part of $\int_0^\infty e^{ix^2} \, dx$.

Solution: Our segments are now: $C_1$ [0 to $R$, a straight line], $C_2$ [$Re^{it}$ with $0 \leq t \leq 1$, an eighth of a circle], and $C_3$ [$Re^{it}$/4 to 0, a straight line].

On $C_2$ 
\[ \int_{C_2} e^{iz^2} \, dz \leq \int_{C_2} |e^{iz^2}| \, |dz| = \int_0^1 e^{-R^2 \sin^2 \frac{\pi t}{4}} R \frac{\pi}{4} \, dt \leq \frac{\pi}{4} R \int_0^1 e^{-R^2 t} \, dt = \frac{\pi R}{4} e^{-R^2 t} \bigg|_{t=0}^{t=1} \leq \frac{\pi}{4} \cdot \frac{1}{R} \to 0 \] as $R \to \infty$.

Since our integrand has no singularities, the contributions of $C_1$ and $C_3$ must be identical, with opposite signs (to cancel out). Parametrizing $C_3$ by $z = t(1 + i)$ with $0 \leq t \leq \infty$ (taking the $R \to \infty$ limit at the same time) results in $(1 + i) \int_0^\infty e^{-2t^2} \, dt$ [as $z^2 = 2it^2$, and $dz = (1 + i) \, dt = (1 + i) \cdot \sqrt{2}$. This implies that $\int_0^\infty \cos(x^2) \, dx = \int_0^\infty \sin(x^2) \, dx = \frac{\sqrt{2\pi}}{4}$.]