Costas J. Papachristou

MATHEMATICAL HANDBOOK

2024

CONTENTS

MATHEMATICAL FORMULAS AND PROPERTIES

Trigonometric formulas

$\sin^2 A + \cos^2 A = 1$; $\tan x = \dfrac{\sin x}{\cos x}$; $\cot x = \dfrac{\cos x}{\sin x} = \dfrac{1}{\tan x}$

$\cos^2 x = \dfrac{1}{1+\tan^2 x}$; $\sin^2 x = \dfrac{1}{1+\cot^2 x} = \dfrac{\tan^2 x}{1+\tan^2 x}$

$\sin(A \pm B) = \sin A \cos B \pm \cos A \sin B$

$\cos(A \pm B) = \cos A \cos B \mp \sin A \sin B$

$\tan(A \pm B) = \dfrac{\tan A \pm \tan B}{1 \mp \tan A \tan B}$, $\cot(A \pm B) = \dfrac{\cot A \cot B \mp 1}{\cot B \pm \cot A}$

$\sin 2A = 2 \sin A \cos A$

$\cos 2A = \cos^2 A - \sin^2 A = 2\cos^2 A - 1 = 1 - 2\sin^2 A$

$\tan 2A = \dfrac{2\tan A}{1-\tan^2 A}$, $\cot 2A = \dfrac{\cot^2 A - 1}{2\cot A}$

$$\sin A + \sin B = 2\sin\frac{A+B}{2}\cos\frac{A-B}{2}$$

$$\sin A - \sin B = 2\sin\frac{A-B}{2}\cos\frac{A+B}{2}$$

$$\cos A + \cos B = 2\cos\frac{A+B}{2}\cos\frac{A-B}{2}$$

$$\cos A - \cos B = 2\sin\frac{A+B}{2}\sin\frac{B-A}{2}$$

$$\sin A \sin B = \frac{1}{2}[\cos(A-B) - \cos(A+B)]$$

$$\cos A \cos B = \frac{1}{2}[\cos(A+B) + \cos(A-B)]$$

$$\sin A \cos B = \frac{1}{2}[\sin(A+B) + \sin(A-B)]$$

$\sin(-A) = -\sin A$, $\cos(-A) = \cos A$

$\tan(-A) = -\tan A$, $\cot(-A) = -\cot A$

$\sin(\dfrac{\pi}{2} \pm A) = \cos A$, $\cos(\dfrac{\pi}{2} \pm A) = \mp \sin A$

$\sin(\pi \pm A) = \mp \sin A$, $\cos(\pi \pm A) = -\cos A$

	sin	cos	tan	cot
0	0	1	0	∞
$\pi/6$ = 30°	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{3}}{3}$	$\sqrt{3}$
$\pi/4$ = 45°	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1	1
$\pi/3$ = 60°	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\sqrt{3}$	$\frac{\sqrt{3}}{3}$
$\pi/2$ = 90°	1	0	∞	0
π = 180°	0	–1	0	∞

Basic trigonometric equations

$$\sin x = \sin\alpha \;\Rightarrow\; \begin{cases} x=\alpha+2k\pi \\ x=(2k+1)\pi-\alpha \end{cases} \qquad (k=0,\pm1,\pm2,\cdots)$$

$$\cos x = \cos\alpha \;\Rightarrow\; \begin{cases} x=\alpha+2k\pi \\ x=2k\pi-\alpha \end{cases} \qquad (k=0,\pm1,\pm2,\cdots)$$

$$\tan x = \tan\alpha \;\Rightarrow\; x=\alpha+k\pi \qquad (k=0,\pm1,\pm2,\cdots)$$

$$\cot x = \cot\alpha \;\Rightarrow\; x=\alpha+k\pi \qquad (k=0,\pm1,\pm2,\cdots)$$

$$\sin x = -\sin\alpha \;\Rightarrow\; \begin{cases} x=2k\pi-\alpha \\ x=\alpha+(2k+1)\pi \end{cases} \qquad (k=0,\pm1,\pm2,\cdots)$$

$$\cos x = -\cos\alpha \;\Rightarrow\; \begin{cases} x=(2k+1)\pi-\alpha \\ x=\alpha+(2k+1)\pi \end{cases} \qquad (k=0,\pm1,\pm2,\cdots)$$

Hyperbolic functions

$$\cosh x = \frac{e^x+e^{-x}}{2} \quad ; \quad \sinh x = \frac{e^x-e^{-x}}{2} \quad ; \quad \tanh x = \frac{\sinh x}{\cosh x} = \frac{e^x-e^{-x}}{e^x+e^{-x}} = \frac{1}{\coth x}$$

$$\cosh^2 x - \sinh^2 x = 1$$

$$\cosh(-x) = \cosh x \quad , \quad \sinh(-x) = -\sinh x$$

Power formulas

$(a\pm b)^2 = a^2 \pm 2ab + b^2$

$(a\pm b)^3 = a^3 \pm 3a^2b + 3ab^2 \pm b^3$

$a^2 - b^2 = (a+b)(a-b)$

$a^3 \pm b^3 = (a\pm b)(a^2 \mp ab + b^2)$

$$(a+b)^n = a^n + na^{n-1}b + \frac{n(n-1)}{2!}a^{n-2}b^2 + \frac{n(n-1)(n-2)}{3!}a^{n-3}b^3 + \cdots + b^n \quad (n=1,2,3,\cdots)$$

Quadratic equation: $ax^2 + bx + c = 0$

Call $D = b^2 - 4ac$ (discriminant)

Roots: $x = \dfrac{-b \pm \sqrt{D}}{2a}$

Roots are real and distinct if $D>0$; real and equal if $D=0$; complex conjugate if $D<0$.

Geometric formulas

A= area or surface area ; V= volume ; P= perimeter

Parallelogram of base b and altitude h : $A=bh$

Triangle of base b and altitude h : $A= (1/2)bh$

Trapezoid of altitude h and parallel sides a and b : $A= (1/2)(a+b)h$

Circle of radius r : $P=2\pi r$, $A=\pi r^2$

Ellipse of semi-major axis a and semi-minor axis b : $A=\pi ab$

Parallelepiped of base area A and height h : $V=Ah$

Cylindroid of base area A and height h : $V=Ah$

Sphere of radius r : $A=4\pi r^2$, $V= (4/3)\pi r^3$

Circular cone of radius r and height h : $V= (1/3)\pi r^2 h$

Properties of inequalities

$$a < b \;\; and \;\; b < c \;\Rightarrow\; a < c$$

$$a \geq b \;\; and \;\; b \geq a \;\Rightarrow\; a = b$$

$$a < b \;\Rightarrow\; -a > -b$$

$$0 < a < b \;\Rightarrow\; \frac{1}{a} > \frac{1}{b}$$

$$a < b \;\; and \;\; c \leq d \;\Rightarrow\; a + c < b + d$$

$$0 < a < b \;\; and \;\; 0 < c \leq d \;\Rightarrow\; ac < bd$$

$$0 < a < 1 \;\Rightarrow\; a > a^2 > a^3 > \cdots \;, \quad a^n < 1 \,, \quad \sqrt[n]{a} < 1$$

$$a > 1 \;\Rightarrow\; a < a^2 < a^3 < \cdots \;, \quad a^n > 1 \,, \quad \sqrt[n]{a} > 1$$

$$0 < a < b \;\Rightarrow\; a^n < b^n \,, \quad \sqrt[n]{a} < \sqrt[n]{b}$$

Properties of proportions

Assume that $\dfrac{\alpha}{\beta} = \dfrac{\gamma}{\delta} = \kappa$. Then,

$$\alpha\delta = \beta\gamma \;, \qquad\qquad \frac{\alpha \pm \gamma}{\beta \pm \delta} = \kappa$$

$$\frac{\alpha \pm \beta}{\beta} = \frac{\gamma \pm \delta}{\delta} \;, \qquad\qquad \frac{\alpha}{\beta \pm \alpha} = \frac{\gamma}{\delta \pm \gamma}$$

Properties of absolute values of real numbers

$$|a| = a, \quad if \quad a \geq 0$$
$$\quad = -a, \quad if \quad a < 0$$

$$|a| \geq 0$$

$$|-a| = |a|$$

$$|a|^2 = a^2$$

$$\sqrt{a^2} = |a|$$

$$|x| \leq \varepsilon \quad \Leftrightarrow \quad -\varepsilon \leq x \leq \varepsilon \quad (\varepsilon > 0)$$

$$|x| \geq a > 0 \quad \Leftrightarrow \quad x \geq a \quad or \quad x \leq -a$$

$$\big||a| - |b|\big| \leq |a \pm b| \leq |a| + |b|$$

$$|a \cdot b| = |a||b|$$

$$|a^k| = |a|^k \quad (k \in Z)$$

$$\left|\frac{a}{b}\right| = \frac{|a|}{|b|} \quad (b \neq 0)$$

Properties of powers and logarithms

$$x^0 = 1 \qquad (x \neq 0)$$

$$x^\alpha x^\beta = x^{\alpha+\beta}$$

$$\frac{x^\alpha}{x^\beta} = x^{\alpha-\beta}$$

$$\frac{1}{x^\alpha} = x^{-\alpha}$$

$$\left(x^\alpha\right)^\beta = x^{\alpha\beta}$$

$$(xy)^\alpha = x^\alpha y^\alpha \quad ; \qquad \left(\frac{x}{y}\right)^\alpha = \frac{x^\alpha}{y^\alpha}$$

$$\ln 1 = 0$$

$$\ln\left(e^\alpha\right) = \alpha \quad (\alpha \in \mathbb{R}) \quad , \qquad e^{\ln\alpha} = \alpha \quad (\alpha \in \mathbb{R}^+)$$

$$\ln(\alpha\beta) = \ln\alpha + \ln\beta$$

$$\ln\left(\frac{\alpha}{\beta}\right) = \ln\alpha - \ln\beta = -\ln\left(\frac{\beta}{\alpha}\right)$$

$$\ln\left(\frac{1}{\alpha}\right) = -\ln\alpha$$

$$\ln\left(\alpha^k\right) = k\ln\alpha \quad (k \in \mathbb{R})$$

Derivatives and integrals of elementary functions

$(c)' = 0 \quad (c = const.)$

$(\sin x)' = \cos x$

$(\arcsin x)' = \dfrac{1}{\sqrt{1-x^2}}$

$(x^\alpha)' = \alpha x^{\alpha-1} \quad (\alpha \in R)$

$(\cos x)' = -\sin x$

$(\arccos x)' = -\dfrac{1}{\sqrt{1-x^2}}$

$(e^x)' = e^x$

$(\tan x)' = \dfrac{1}{\cos^2 x}$

$(\arctan x)' = \dfrac{1}{1+x^2}$

$(\ln x)' = \dfrac{1}{x} \quad (x > 0)$

$(\cot x)' = -\dfrac{1}{\sin^2 x}$

$(\operatorname{arc}\cot x)' = -\dfrac{1}{1+x^2}$

$(\sinh x)' = \cosh x$

$(\cosh x)' = \sinh x$

$\int dx = x + C$; $\int x^a dx = \dfrac{x^{a+1}}{a+1} + C \quad (a \neq -1)$

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

$$\int e^x dx = e^x + C$$

$\int \cos x\, dx = \sin x + C$; $\int \sin x\, dx = -\cos x + C$

$\int \dfrac{dx}{\cos^2 x} = \tan x + C$; $\int \dfrac{dx}{\sin^2 x} = -\cot x + C$

$$\int \frac{dx}{\sqrt{1-x^2}} = \arcsin x + C$$

$$\int \frac{dx}{1+x^2} = \arctan x + C$$

$$\int \frac{dx}{x^2-1} = \frac{1}{2}\ln\left|\frac{x-1}{x+1}\right| + C$$

$$\int \frac{dx}{\sqrt{x^2 \pm 1}} = \ln\left(x + \sqrt{x^2 \pm 1}\right) + C$$

COMPLEX NUMBERS

Consider the equation $x^2+1=0$. This has no solution for real x. For this reason we extend the set of numbers beyond the real numbers by defining the *imaginary unit number* i by

$$i^2=-1 \quad \text{or, symbolically,} \quad i=\sqrt{-1} \; .$$

Then, the solution of the above-given equation is $x=\pm i$.

Given the *real* numbers x and y, we define the *complex number*

$$z=x+iy \; .$$

This is often represented as an ordered pair

$$z=x+iy \equiv (x,y) \; .$$

The number $x=\mathrm{Re}\, z$ is the *real part* of z while $y=\mathrm{Im}\, z$ is the *imaginary part* of z. In particular, the value $z=0$ corresponds to $x=0$ and $y=0$. In general, if $y=0$, then z is a *real* number.

Given a complex number $z=x+iy$, the number

$$\bar{z}=x-iy$$

is called the *complex conjugate* of z (the symbol z^* is also used for the complex conjugate). Furthermore, the *real* quantity

$$|z| = (x^2+y^2)^{1/2}$$

is called the *modulus* (or absolute value) of z. We notice that

$$|z|=|\bar{z}| \; .$$

Example: If $z=3+2i$, then $\bar{z}=3-2i$ and $|z|=|\bar{z}|=\sqrt{13}$.

Exercise: Show that, if $z=\bar{z}$, then z is *real*, and conversely.

Exercise: Show that, if $z=x+iy$, then

$$\mathrm{Re}\, z = x = \frac{z+\bar{z}}{2} \;, \qquad \mathrm{Im}\, z = y = \frac{z-\bar{z}}{2i} \; .$$

Consider the complex numbers $z_1 = x_1 + i\,y_1$, $z_2 = x_2 + i\,y_2$. As we can show, their sum and their difference are given by

$$z_1 + z_2 = (x_1 + x_2) + i\,(y_1 + y_2) \,,$$

$$z_1 - z_2 = (x_1 - x_2) + i\,(y_1 - y_2) \,.$$

Exercise: Show that, if $z_1 = z_2$, then $x_1 = x_2$ and $y_1 = y_2$.

Taking into account that $i^2 = -1$, we find the product of z_1 and z_2 to be

$$z_1 z_2 = (x_1 x_2 - y_1 y_2) + i\,(x_1 y_2 + x_2 y_1) \,.$$

In particular, for $z_1 = z = x + i\,y$ and $z_2 = \bar{z} = x - i\,y$, we have:

$$z\,\bar{z} = x^2 + y^2 = |\,z\,|^2 \,.$$

To evaluate the quotient z_1 / z_2 $(z_2 \neq 0)$ we apply the following trick:

$$\frac{z_1}{z_2} = \frac{z_1 \bar{z}_2}{z_2 \bar{z}_2} = \frac{z_1 \bar{z}_2}{|\,z_2\,|^2} = \frac{(x_1 + i\,y_1)(x_2 - i\,y_2)}{x_2^{\,2} + y_2^{\,2}} = \frac{x_1 x_2 + y_1 y_2}{x_2^{\,2} + y_2^{\,2}} + i\,\frac{x_2 y_1 - x_1 y_2}{x_2^{\,2} + y_2^{\,2}} \;.$$

In particular, for $z = x + i\,y$,

$$\frac{1}{z} = \frac{\bar{z}}{z\,\bar{z}} = \frac{\bar{z}}{|\,z\,|^2} = \frac{x - i\,y}{x^2 + y^2} = \frac{x}{x^2 + y^2} - i\,\frac{y}{x^2 + y^2} \;.$$

Properties:

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2 \;, \qquad \overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2$$

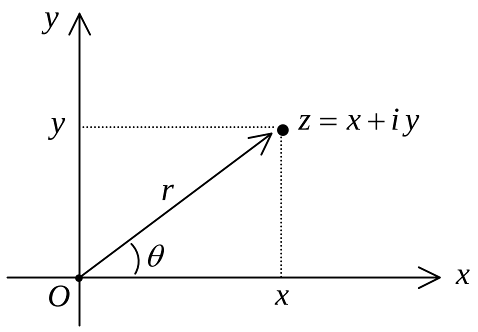
$$\overline{z_1 z_2} = \bar{z}_1\,\bar{z}_2 \;, \qquad \overline{\left(\frac{z_1}{z_2}\right)} = \frac{\bar{z}_1}{\bar{z}_2}$$

$$|\,\bar{z}\,| = |\,z\,| \;, \qquad z\,\bar{z} = |\,z\,|^2 \;, \qquad |\,z_1 z_2\,| = |\,z_1\,|\,|z_2\,|$$

$$|\,z^n\,| = |\,z\,|^n \;, \qquad \left|\frac{z_1}{z_2}\right| = \frac{|\,z_1\,|}{|\,z_2\,|}$$

Exercise: Given the complex numbers $z_1 = 3 - 2\,i$ and $z_2 = -2 + i$, evaluate the quantities $|\,z_1 \pm z_2\,|$, $\bar{z}_1 z_2$ and $\overline{z_1 / z_2}$.

Polar form of a complex number



A complex number $z = x + i\,y \equiv (x, y)$ corresponds to a point of the *x*-*y* plane. It may also be represented by a vector joining the origin *O* of the axes of the complex plane with this point. The quantities *x* and *y* are the Cartesian coordinates of the point, or, the orthogonal components of the corresponding vector. We observe that

$$x = r\cos\theta \ , \quad y = r\sin\theta$$

where

$$r = |z| = (x^2 + y^2)^{1/2} \quad \text{and} \quad \tan\theta = \frac{y}{x} \ .$$

Thus, we can write

$$\boxed{z = x + i\,y = r\,(\cos\theta + i\sin\theta)}$$

The above expression represents the *polar form* of z. Note that

$$\bar{z} = r(\cos\theta - i\sin\theta) \ .$$

Let $z_1 = r_1\,(\cos\theta_1 + i\sin\theta_1)$ and $z_2 = r_2\,(\cos\theta_2 + i\sin\theta_2)$ be two complex numbers. As can be shown,

$$z_1 z_2 = r_1 r_2\,[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)] \ ,$$

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}\,[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)] \ .$$

In particular, the inverse of a complex number $z = r\,(\cos\theta + i\sin\theta)$ is written

$$z^{-1} = \frac{1}{z} = \frac{1}{r}(\cos\theta - i\sin\theta) = \frac{1}{r}\,[\cos(-\theta) + i\sin(-\theta)] \ .$$

Exercise: By using polar forms, show analytically that $z\,z^{-1} = 1$.

Exponential form of a complex number

We introduce the notation

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

(this notation is not arbitrary but has a deeper meaning that reveals itself within the context of the theory of analytic functions). Note that

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

Also,

$$|e^{i\theta}| = |e^{-i\theta}| = \cos^2\theta + \sin^2\theta = 1 .$$

Exercise: Show that

$$e^{-i\theta} = 1/e^{i\theta} = \overline{e^{i\theta}} .$$

Also show that

$$\cos\theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad , \quad \sin\theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} \ .$$

The complex number $z = r(\cos\theta + i\sin\theta)$, where $r = |z|$, may now be expressed as follows:

$$\boxed{z = r\,e^{i\theta}}$$

It can be shown that

$$z^{-1} = \frac{1}{z} = \frac{1}{r}\,e^{-i\theta} = \frac{1}{r}\,e^{i(-\theta)} , \qquad \bar{z} = r\,e^{-i\theta}$$

$$z_1 z_2 = r_1 r_2\,e^{i(\theta_1+\theta_2)} , \qquad \frac{z_1}{z_2} = \frac{r_1}{r_2}\,e^{i(\theta_1-\theta_2)}$$

where $z_1 = r_1 e^{i\theta_1}$, $z_2 = r_2 e^{i\theta_2}$.

Example: The complex number $z = \sqrt{2} - i\sqrt{2}$, with $|z| = r = 2$, is written

$$z = 2\left(\frac{\sqrt{2}}{2} - i\,\frac{\sqrt{2}}{2}\right) = 2\left[\cos\left(-\frac{\pi}{4}\right) + i\sin\left(-\frac{\pi}{4}\right)\right] = 2\,e^{i(-\pi/4)} = 2\,e^{-i\pi/4} .$$

Powers and roots of complex numbers

Let $z = r\,(\cos\theta + i\sin\theta) = r\,e^{i\theta}$ be a complex number, where $r = |z|$. It can be proven that

$$z^n = r^n e^{in\theta} = r^n (\cos n\theta + i\sin n\theta) \qquad (n = 0, \pm 1, \pm 2, \cdots) \ .$$

In particular, for $z = \cos\theta + i\sin\theta = e^{i\theta}$ $(r=1)$ we find the *de Moivre formula*

$$(\cos\theta + i\sin\theta)^n = (\cos n\theta + i\sin n\theta) \ .$$

Note also that, for $z \neq 0$, we have that $z^0 = 1$ and $z^{-n} = 1/z^n$.

Given a complex number $z = r\,(\cos\theta + i\sin\theta)$, where $r = |z|$, an *nth root of* z is any complex number c satisfying the equation $c^n = z$. We write $c = \sqrt[n]{z}$. An nth root of a complex number admits n different values given by the formula

$$c_k = \sqrt[n]{r}\left(\cos\frac{\theta + 2k\pi}{n} + i\sin\frac{\theta + 2k\pi}{n}\right), \quad k = 0, 1, 2, \cdots, (n-1) \ .$$

Example: Let $z = 1$. We seek the 4th roots of unity, i.e., the complex numbers c satisfying the equation $c^4 = 1$. We write

$$z = 1\,(\cos 0 + i\sin 0) \qquad (\text{that is, } r = 1,\ \theta = 0) \ .$$

Then,

$$c_k = \cos\frac{2k\pi}{4} + i\sin\frac{2k\pi}{4} = \cos\frac{k\pi}{2} + i\sin\frac{k\pi}{2} \ , \quad k = 0, 1, 2, 3 \ .$$

We find:

$$c_0 = 1 \ , \quad c_1 = i \ , \quad c_2 = -1 \ , \quad c_3 = -i \ .$$

Example: Let $z = i$. We seek the square roots of i, that is, the complex numbers c satisfying the equation $c^2 = i$. We have:

$$z = 1\,[\cos(\pi/2) + i\sin(\pi/2)] \qquad (\text{that is, } r = 1,\ \theta = \pi/2) \ ;$$

$$c_k = \cos\frac{(\pi/2) + 2k\pi}{2} + i\sin\frac{(\pi/2) + 2k\pi}{2} \ , \quad k = 0, 1 \ ;$$

$$c_0 = \cos(\pi/4) + i\sin(\pi/4) = \frac{\sqrt{2}}{2}\,(1+i) \ ,$$

$$c_1 = \cos(5\pi/4) + i\sin(5\pi/4) = -\frac{\sqrt{2}}{2}\,(1+i) \ .$$

ALGEBRA: SOME BASIC CONCEPTS

Sets

Subset:	$X \subseteq Y \Leftrightarrow (x \in X \Rightarrow x \in Y)$; $X = Y \Leftrightarrow X \subseteq Y$ *and* $Y \subseteq X$
Proper subset:	$X \subset Y \Leftrightarrow X \subseteq Y$ *and* $X \neq Y$
Union of sets:	$X \cup Y = \{ x / x \in X$ *or* $x \in Y \}$
Intersection of sets:	$X \cap Y = \{ x / x \in X$ *and* $x \in Y \}$
Disjoint sets:	$X \cap Y = \varnothing$
Difference of sets:	$X - Y = \{ x / x \in X$ *and* $x \notin Y \}$
Complement of a subset:	$X \supset Y$; $X \setminus Y = X - Y$
Cartesian product:	$X \times Y = \{(x, y) / x \in X$ *and* $y \in Y \}$
Mapping:	$f : X \to Y$; $(x \in X) \to y = f(x) \in Y$
Domain / range of f:	$D(f) = X$, $R(f) = f(X) = \{ f(x) / x \in X \} \subseteq Y$; f is *defined* in X and *has values* in Y ; $y = f(x)$ is the *image* of x under f
Composite mapping:	$f : X \to Y$, $g : Y \to Z$; $f \circ g : X \to Z$; $(x \in X) \to g(f(x)) \in Z$
Injective (1-1) mapping:	$f(x_1) = f(x_2) \Leftrightarrow x_1 = x_2$, *or* $x_1 \neq x_2 \Leftrightarrow f(x_1) \neq f(x_2)$
Surjective (onto) mapping:	$f(X) = Y$
Bijective mapping:	f *is both injective and surjective* $\Rightarrow$ *invertible*
Identity mapping:	$f_{id} : X \to X$; $f_{id}(x) = x$, $\forall x \in X$
Internal operation on X:	$X \times X \to X$; $[(x, y) \in X \times X] \to z \in X$
External operation on X:	$A \times X \to X$; $[(a, x) \in A \times X] \to y = a \cdot x \in X$

Groups

A *group* is a set G, together with an internal operation $G \times G \to G$; $(x, y) \to z = x \cdot y$, such that:

1. The operation is *associative*: $x \cdot (y \cdot z) = (x \cdot y) \cdot z$
2. $\exists\, e \in G$ (*identity*) : $x \cdot e = e \cdot x = x$, $\forall\, x \in G$
3. $\forall\, x \in G$, $\exists\, x^{-1} \in G$ (*inverse*): $x^{-1} \cdot x = x \cdot x^{-1} = e$

A group G is *abelian* or *commutative* if $x \cdot y = y \cdot x$, $\forall\, x, y \in G$.

A subset $S \subseteq G$ is a *subgroup* of G if S is itself a group (clearly, then, S contains the identity e of G, as well as the inverse of every element of S).

Vector space over R

Let $V=\{\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, ...\}$, and let $a, b, c, ... \in R$. Consider an internal operation + and an external operation $\cdot$ on V :

$$+ : \; V \times V \to V \;\; ; \quad \boldsymbol{x}+\boldsymbol{y} = \boldsymbol{z}$$
$$\cdot : \; R \times V \to V \;\; ; \quad a \cdot \boldsymbol{x} = \boldsymbol{y}$$

Then, V is a *vector space over R* iff

1. V is a commutative group with respect to + . The identity element is denoted $\mathbf{0}$, while the inverse of $\boldsymbol{x}$ is denoted $-\boldsymbol{x}$.
2. The operation $\cdot$ satisfies the following:
   $a \cdot (b \cdot \boldsymbol{x}) = (ab) \cdot \boldsymbol{x}$
   $(a+b) \cdot \boldsymbol{x} = a \cdot \boldsymbol{x} + b \cdot \boldsymbol{x}$
   $a \cdot (\boldsymbol{x}+\boldsymbol{y}) = a \cdot \boldsymbol{x} + a \cdot \boldsymbol{y}$
   $1 \cdot \boldsymbol{x} = \boldsymbol{x}$, $\quad 0 \cdot \boldsymbol{x} = \mathbf{0}$

A set $\{\boldsymbol{x}_1 , \boldsymbol{x}_2 , ... , \boldsymbol{x}_k\}$ of elements of V is *linearly independent* iff the equation[1]

$$c_1 \boldsymbol{x}_1 + c_2 \boldsymbol{x}_2 + ... + c_k \boldsymbol{x}_k = 0$$

can only be satisfied for $c_1 = c_2 = ... = c_k = 0$; otherwise, the set is *linearly dependent*. The *dimension* $\dim V$ of V is the largest number of vectors in V that constitute a linearly independent set. If $\dim V = n$, then any system $\{\boldsymbol{e}_1 , \boldsymbol{e}_2 , ... , \boldsymbol{e}_n\}$ of n linearly independent elements is a *basis* for V, and any $\boldsymbol{x} \in V$ can be uniquely expressed as $\boldsymbol{x} = c_1 \boldsymbol{e}_1 + c_2 \boldsymbol{e}_2 + ... + c_n \boldsymbol{e}_n$.

A subset $S \subseteq V$ is a *subspace* of V if S is itself a vector space under the operations (+) and ($\cdot$). In particular, the sum $\boldsymbol{x}+\boldsymbol{y}$ of any two elements of S, as well as the scalar multiple $a\boldsymbol{x}$ and the inverse $-\boldsymbol{x}$ of any element $\boldsymbol{x}$ of S, must belong to S. Clearly, this set must contain the identity $\mathbf{0}$ of V. If S is a subspace of V, then $\dim S \le \dim V$. In particular, S coincides with V iff $\dim S = \dim V$.

---

[1] The symbol ($\cdot$) will often be omitted in the sequel.

Functionals

A *functional* $\boldsymbol{\omega}$ on a vector space V is a mapping $\boldsymbol{\omega}: V \to R$; $(\boldsymbol{x}\in V) \to t = \boldsymbol{\omega}(\boldsymbol{x})\in R$. The functional $\boldsymbol{\omega}$ is *linear* if $\boldsymbol{\omega}(a\cdot\boldsymbol{x}+b\cdot\boldsymbol{y})=a\cdot\boldsymbol{\omega}(\boldsymbol{x})+b\cdot\boldsymbol{\omega}(\boldsymbol{y})$. The collection of all linear functionals on V is called the *dual space* of V, denoted V^*. It is itself a vector space over R, and $\dim V^* = \dim V$.

Algebras

A *real algebra* A is a vector space over R equipped with a binary operation $(\cdot\,|\,\cdot) : A \times A \to A$; $(\boldsymbol{x}\,|\,\boldsymbol{y}) = \boldsymbol{z}$, such that, for $a, b \in R$,

$$(a\cdot\boldsymbol{x} + b\cdot\boldsymbol{y}\,|\,\boldsymbol{z}) = a\cdot(\boldsymbol{x}\,|\,\boldsymbol{z}) + b\cdot(\boldsymbol{y}\,|\,\boldsymbol{z})$$
$$(\boldsymbol{x}\,|\,a\cdot\boldsymbol{y} + b\cdot\boldsymbol{z}) = a\cdot(\boldsymbol{x}\,|\,\boldsymbol{y}) + b\cdot(\boldsymbol{x}\,|\,\boldsymbol{z})$$

An algebra is *commutative* if, for any two elements $\boldsymbol{x}$, $\boldsymbol{y}$, $(\boldsymbol{x}\,|\,\boldsymbol{y}) = (\boldsymbol{y}\,|\,\boldsymbol{x})$; it is *associative* if, for any $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$, $(\boldsymbol{x}\,|\,(\boldsymbol{y}\,|\,\boldsymbol{z})) = ((\boldsymbol{x}\,|\,\boldsymbol{y})\,|\,\boldsymbol{z})$.

Example: The set $\Lambda^0(R^n)$ of all functions on R^n is a commutative, associative algebra. The multiplication operation $(\cdot\,|\,\cdot) : \Lambda^0(R^n)\times \Lambda^0(R^n)\to \Lambda^0(R^n)$ is defined by

$$(f\,|\,g)(x^1, \ldots, x^n) = f(x^1, \ldots, x^n)\, g(x^1, \ldots, x^n) .$$

Example: The set of all $n\times n$ matrices is an associative, non-commutative algebra. The binary operation $(\cdot\,|\,\cdot)$ is matrix multiplication.

A subspace S of A is a *subalgebra* of A if S is itself an algebra under the same binary operation $(\cdot\,|\,\cdot)$. In particular, S must be closed under this operation; i.e., $(\boldsymbol{x}\,|\,\boldsymbol{y})\in S$ for any $\boldsymbol{x}, \boldsymbol{y}$ in S. We write: $S\subset A$.

A subalgebra $S\subset A$ is an *ideal* of A iff $(\boldsymbol{x}\,|\,\boldsymbol{y})\in S$ and $(\boldsymbol{y}\,|\,\boldsymbol{x})\in S$, for any $\boldsymbol{x}\in S$, $\boldsymbol{y}\in A$.

Modules

Note first that R is an associative, commutative algebra under the usual operations of addition and multiplication. Thus, a vector space over R is a vector space over an associative, commutative algebra. More generally, a *module M over A* is a vector space over an *associative* but (generally) *non-commutative* algebra. In particular, the external operation $(\cdot)$ on M is defined by

$$\cdot : A \times M \to M ; \quad a\cdot\boldsymbol{x} = \boldsymbol{y} \quad (a\in A ;\ \boldsymbol{x}, \boldsymbol{y} \in M) .$$

Example: The collection of all n-dimensional column matrices, with A taken to be the algebra of $n\times n$ matrices, and with matrix multiplication as the external operation.

Vector fields

A *vector field* $\boldsymbol{V}$ on R^n is a map from a domain of R^n into R^n :

$$\boldsymbol{V} : R^n \supseteq U \to R^n \ ; \quad [\boldsymbol{x} \equiv (x^1,..., x^n) \in U \ ] \to \boldsymbol{V}(\boldsymbol{x}) \equiv (V^1(x^k),...,V^n(x^k)) \in R^n \ .$$

The vector $\boldsymbol{x}$ represents a point in U, with coordinates $(x^1,..., x^n)$. The functions $V^i(x^k)$ $(i=1,...,n)$ are the *components* of $\boldsymbol{V}$ in the coordinate system (x^k).

Given two vector fields $\boldsymbol{U}$ and $\boldsymbol{V}$, we can construct a new vector field $\boldsymbol{W}=\boldsymbol{U}+\boldsymbol{V}$ such that $\boldsymbol{W}(\boldsymbol{x})=\boldsymbol{U}(\boldsymbol{x})+\boldsymbol{V}(\boldsymbol{x})$. The components of $\boldsymbol{W}$ are the sums of the respective components of $\boldsymbol{U}$ and $\boldsymbol{V}$.

Given a vector field $\boldsymbol{V}$ and a constant $a \in R$, we can construct a new vector field $\boldsymbol{Z}=a\boldsymbol{V}$ such that $\boldsymbol{Z}(\boldsymbol{x})= a\boldsymbol{V}(\boldsymbol{x})$. The components of $\boldsymbol{Z}$ are scalar multiples (by a) of those of $\boldsymbol{V}$.

It follows from the above that *the collection of all vector fields on R^n is a vector space over R* .

More generally, given a vector field $\boldsymbol{V}$ and a function $f \in \Lambda^0(R^n)$, we can construct a new vector field $\boldsymbol{Z}= f\,\boldsymbol{V}$ such that $\boldsymbol{Z}(\boldsymbol{x})= f(\boldsymbol{x})\boldsymbol{V}(\boldsymbol{x})$. Given that $\Lambda^0(R^n)$ is an associative algebra, we conclude that *the collection of all vector fields on R^n is a module over* $\Lambda^0(R^n)$ (in this particular case, the algebra $\Lambda^0(R^n)$ is commutative).

A note on linear independence:

Let $\{\boldsymbol{V}_1 , ... ,\boldsymbol{V}_r\} \equiv \{\boldsymbol{V}_a\}$ be a collection of vector fields on R^n.

(a) The set $\{\boldsymbol{V}_a\}$ is *linearly dependent over R* (linearly dependent with constant coefficients) iff there exist real constants $c_1 ,...,c_r$, *not all zero*, such that

$$c_1\boldsymbol{V}_1 (\boldsymbol{x}) + \ldots + c_r\boldsymbol{V}_r (\boldsymbol{x}) = 0 \ , \quad \forall \boldsymbol{x} \in R^n \ .$$

If the above relation is satisfied only for $c_1 = \ldots = c_r = 0$, the set $\{\boldsymbol{V}_a\}$ is *linearly independent over R*.

(b) The set $\{\boldsymbol{V}_a\}$ is *linearly dependent over* $\Lambda^0(R^n)$ iff there exist functions $f_1 (x^k)$, ..., $f_r (x^k)$, *not all identically zero over R^n*, such that

$$f_1 (x^k)\, \boldsymbol{V}_1 (\boldsymbol{x}) + \ldots + f_r (x^k)\, \boldsymbol{V}_r (\boldsymbol{x}) = 0 \ , \quad \forall \boldsymbol{x} \equiv (x^k) \in R^n \ .$$

If this relation is satisfied only for $f_1 (x^k)=\ldots= f_r (x^k) \equiv 0$, the set $\{\boldsymbol{V}_a\}$ is *linearly independent over* $\Lambda^0(R^n)$.

There can be at most n elements in a linearly independent system over $\Lambda^0(R^n)$. These elements form a *basis* $\{\boldsymbol{e}_1 , ..., \boldsymbol{e}_n\} \equiv \{\boldsymbol{e}_k\}$ for the module of all vector fields on R^n. An element of this module, i.e. an arbitrary vector field $\boldsymbol{V}$, is written as a linear combination of the $\{\boldsymbol{e}_k\}$ with coefficients $V^k \in \Lambda^0(R^n)$. Thus, at any point $\boldsymbol{x} \equiv (x^k) \in R^n$,

$$\boldsymbol{V}(\boldsymbol{x}) = V^1(x^k)\,\boldsymbol{e}_1 + ... + V^n(x^k)\,\boldsymbol{e}_n \equiv (V^1(x^k), ... , V^n(x^k)) .$$

In particular, in the basis $\{\boldsymbol{e}_k\}$,

$$\boldsymbol{e}_1 \equiv (1,0,0,...,0), \quad \boldsymbol{e}_2 \equiv (0,1,0,...,0), \quad \ldots \quad , \boldsymbol{e}_n \equiv (0,0,...,0,1) .$$

Example: Let n=3, i.e., R^n=R^3. Call $\{\boldsymbol{e}_1 , \boldsymbol{e}_2 , \boldsymbol{e}_3\} \equiv \{\boldsymbol{i} , \boldsymbol{j}, \boldsymbol{k}\}$. Let $\boldsymbol{V}$ be a vector field on R^3. Then, at any point $\boldsymbol{x} \equiv (x, y, z) \in R^3$,

$$\boldsymbol{V}(\boldsymbol{x}) = V_x (x, y, z)\, \boldsymbol{i} + V_y (x, y, z)\, \boldsymbol{j} + V_z (x, y, z)\, \boldsymbol{k} \equiv (V_x , V_y , V_z) .$$

Now, consider the six vector fields

$$\boldsymbol{V}_1 = \boldsymbol{i} , \; \boldsymbol{V}_2 = \boldsymbol{j} , \; \boldsymbol{V}_3 = \boldsymbol{k} , \; \boldsymbol{V}_4 = x\boldsymbol{j} - y\boldsymbol{i} , \; \boldsymbol{V}_5 = y\boldsymbol{k} - z\boldsymbol{j} , \; \boldsymbol{V}_6 = z\boldsymbol{i} - x\boldsymbol{k} .$$

Clearly, the $\{\boldsymbol{V}_1 , \boldsymbol{V}_2 , \boldsymbol{V}_3 \}$ are linearly independent over $\Lambda^0(R^3)$, since they constitute the basis $\{\boldsymbol{i} , \boldsymbol{j}, \boldsymbol{k}\}$. On the other hand, the $\boldsymbol{V}_4 , \boldsymbol{V}_5 , \boldsymbol{V}_6$ are separately linearly dependent on the $\{\boldsymbol{V}_1 , \boldsymbol{V}_2 , \boldsymbol{V}_3 \}$ over $\Lambda^0(R^3)$. Moreover, the set $\{\boldsymbol{V}_4 , \boldsymbol{V}_5 , \boldsymbol{V}_6 \}$ is also linearly dependent over $\Lambda^0(R^3)$, since $z\boldsymbol{V}_4 + x\boldsymbol{V}_5 + y\boldsymbol{V}_6 = 0$. Thus, the set $\{\boldsymbol{V}_1 , \ldots, \boldsymbol{V}_6\}$ is *linearly dependent over* $\Lambda^0(R^3)$. On the other hand, the system $\{\boldsymbol{V}_1 , \ldots, \boldsymbol{V}_6\}$ is *linearly independent over R*, since the equation $c_1\boldsymbol{V}_1 + \ldots + c_6\boldsymbol{V}_6 = 0$, with $c_1 ,...,c_6 \in R$ (constant coefficients), can only be satisfied for $c_1 = ...= c_6 = 0$. In general,

> *there is an infinite number of linearly independent vector fields on R^n over R , but only n linearly independent fields over $\Lambda^0(R^n)$.*

Derivation on an algebra

Let L be an operation on an algebra A (an *operator* on A):

$$L : A \to A ; \quad (\boldsymbol{x} \in A) \to \boldsymbol{y} = L\,\boldsymbol{x} \in A .$$

L is a *derivation* on A iff, $\forall \boldsymbol{x}, \boldsymbol{y} \in A$ and $a, b \in R$,

$$L\,(a\boldsymbol{x}+b\boldsymbol{y}) = aL(\boldsymbol{x}) + bL(\boldsymbol{y}) \qquad (\textit{linearity})$$
$$L\,(\boldsymbol{x}\,|\,\boldsymbol{y}) = (L\,\boldsymbol{x}\,|\,\boldsymbol{y}) + (\boldsymbol{x}\,|\,L\,\boldsymbol{y}) \qquad (\textit{Leibniz rule})$$

Example: Let A=$\Lambda^0(R^n)$=$\{ f\,(x^1,\ldots, x^n)\}$, and let L be the linear operator

$$L = \varphi^1(x^k)\, \partial/\partial x^1 + \ldots + \varphi^n (x^k)\, \partial/\partial x^n \equiv \varphi^i (x^k)\, \partial/\partial x^i ,$$

where the $\varphi^i(x^k)$ are given functions. As can be shown,

$$L\,[\, f\,(x^k)\, g(x^k)] = [L\, f\,(x^k)]\, g(x^k) + f\,(x^k)\, L\, g(x^k) .$$

Hence, L is a derivation on $\Lambda^0(R^n)$.

Lie algebra

An algebra $\mathcal{L}$ over R is a (real) *Lie algebra* with binary operation $[\cdot\,,\cdot]: \mathcal{L}\times\mathcal{L}\rightarrow\mathcal{L}$ (*Lie bracket*) iff this operation satisfies the properties:

$$[a\boldsymbol{x} + b\boldsymbol{y}\,, \boldsymbol{z}] = a\,[\boldsymbol{x}\,, \boldsymbol{z}] + b\,[\boldsymbol{y}\,, \boldsymbol{z}]$$

$$[\boldsymbol{x}\,, \boldsymbol{y}] = -[\boldsymbol{y}\,, \boldsymbol{x}] \qquad (\textit{antisymmetry})$$

$$[\boldsymbol{x}\,, [\boldsymbol{y}\,, \boldsymbol{z}]] + [\boldsymbol{y}\,, [\boldsymbol{z}\,, \boldsymbol{x}]] + [\boldsymbol{z}\,, [\boldsymbol{x}\,, \boldsymbol{y}]] = 0 \qquad (\textit{Jacobi identity})$$

(where $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in\mathcal{L}$ and $a, b \in R$). Note that, by the antisymmetry of the Lie bracket, the first and third properties are written, alternatively,

$$[\boldsymbol{x}\,, a\boldsymbol{y} + b\boldsymbol{z}] = a\,[\boldsymbol{x}\,, \boldsymbol{y}] + b\,[\boldsymbol{x}\,, \boldsymbol{z}]\,,$$

$$[[\boldsymbol{x}\,, \boldsymbol{y}]\,, \boldsymbol{z}] + [[\boldsymbol{y}\,, \boldsymbol{z}]\,, \boldsymbol{x}] + [[\boldsymbol{z}\,, \boldsymbol{x}]\,, \boldsymbol{y}] = 0\,.$$

A Lie algebra is a *non-associative* algebra, since, as follows by the above properties,

$$[\boldsymbol{x}\,, [\boldsymbol{y}\,, \boldsymbol{z}]] \neq [[\boldsymbol{x}\,, \boldsymbol{y}]\,, \boldsymbol{z}]\,.$$

Example: The algebra of $n\times n$ matrices, with $[A\,, B]=AB-BA$ (commutator).

Example: The algebra of all vectors in R^3, with $[\boldsymbol{a}\,, \boldsymbol{b}] = \boldsymbol{a}\times\boldsymbol{b}$ (vector product).

Lie algebra of derivations

Consider the algebra $A=\Lambda^0(R^n)=\{\,f\,(x^1,\ldots, x^n)\}$. Consider also the set $D(A)$ of linear operators on A, of the form

$$L = \varphi^i\,(x^k)\,\partial/\partial x^i \quad (\text{sum on } i = 1, 2, \ldots, n)\,.$$

These first-order differential operators are *derivations* on A (the Leibniz rule is satisfied). Now, given two such operators L_1, L_2, we construct the linear operator (*Lie bracket* of L_1 and L_2), as follows:

$$[L_1\,, L_2] = L_1 L_2 - L_2 L_1\;;$$

$$[L_1\,, L_2]\,f\,(x^k) = L_1\,(L_2\,f\,(x^k)) - L_2\,(L_1\,f\,(x^k))\,.$$

It can be shown that $[L_1\,, L_2]$ is a *first*-order differential operator (a derivation), hence is a member of $D(A)$. (This is *not* the case with second-order operators like L_1L_2!) Moreover, the Lie bracket of operators satisfies all the properties of the Lie bracket of a general Lie algebra (such as antisymmetry and Jacobi identity). It follows that

> *the set $D(A)$ of derivations on $\Lambda^0(R^n)$ is a Lie algebra, with binary operation defined as the Lie bracket of operators.*

Direct sum of subspaces

Let V be a vector space over a field K (where K may be R or C), of dimension $\dim V=n$. Let S_1 , S_2 be *disjoint* (i.e., $S_1 \cap S_2 =\{\mathbf{0}\}$) subspaces of V. We say that V is the *direct sum* of S_1 and S_2 if each vector of V can be *uniquely* represented as the sum of a vector of S_1 and a vector of S_2 . We write: $V= S_1 \oplus S_2$. In terms of dimensions, $\dim V=\dim S_1 +\dim S_2$. We similarly define the vector sum of three subspaces of V, each of which is disjoint from the direct sum of the other two (i.e., $S_1 \cap (S_2 \oplus S_3)=\{\mathbf{0}\}$, etc.).

Homomorphism of vector spaces

Let V, W be vector spaces over a field K. A mapping $\Phi: V \to W$ is said to be a *linear mapping* or *homomorphism* if it preserves linear operations, i.e.,

$$\Phi(\boldsymbol{x}+\boldsymbol{y}) = \Phi(\boldsymbol{x}) + \Phi(\boldsymbol{y}) , \quad \Phi(k\boldsymbol{x}) = k\, \Phi(\boldsymbol{x}) , \quad \forall\ \boldsymbol{x}, \boldsymbol{y} \in V \text{ and } k \in K .$$

A homomorphism which is also *bijective* (1-1) is called an *isomorphism*.

The set of vectors $\boldsymbol{x} \in V$ mapping under Φ into the zero of W is called the *kernel* of the homomorphism Φ:

$$\operatorname{Ker}\Phi = \{ \boldsymbol{x} \in V : \Phi(\boldsymbol{x}) = \mathbf{0}\} .$$

Note that $\Phi(\mathbf{0})=\mathbf{0}$, for *any* homomorphism (clearly, the two zeros refer to *different* vector spaces). Thus, in general, $\mathbf{0} \in \operatorname{Ker}\Phi$.

If $\operatorname{Ker}\Phi =\{\mathbf{0}\}$, then the homomorphism Φ is also an isomorphism of V onto a subspace of W. If, moreover, $\dim V=\dim W$, then the map $\Phi: V \to W$ is itself an *isomorphism*. In this case, $\operatorname{Im}\Phi =W$, where, in general, $\operatorname{Im}\Phi$ (*image of the homomorphism*) is the collection of images of all vectors of V under the map Φ.

The algebra of linear operators

Let V be a vector space over a field K. A *linear operator* $\boldsymbol{A}$ on V is a homomorphism $\boldsymbol{A} : V \to V$. Thus,

$$\boldsymbol{A}(\boldsymbol{x}+\boldsymbol{y}) = \boldsymbol{A}(\boldsymbol{x}) + \boldsymbol{A}(\boldsymbol{y}) , \quad \boldsymbol{A}(k\boldsymbol{x}) = k\boldsymbol{A}(\boldsymbol{x}) , \quad \forall\ \boldsymbol{x}, \boldsymbol{y} \in V \text{ and } k \in K .$$

The sum $\boldsymbol{A}+\boldsymbol{B}$ and the scalar multiplication $k\boldsymbol{A}$ ($k \in K$) are linear operators defined by

$$(\boldsymbol{A}+\boldsymbol{B})\,\boldsymbol{x} = \boldsymbol{A}\,\boldsymbol{x} + \boldsymbol{B}\,\boldsymbol{x} \ , \quad (k\boldsymbol{A})\,\boldsymbol{x} = k\,(\boldsymbol{A}\,\boldsymbol{x}) .$$

Under these operations, the set $Op(V)$ of all linear operators on V is a vector space. The zero element of that space is a zero operator $\boldsymbol{0}$ such that $\boldsymbol{0}\boldsymbol{x}=\mathbf{0}$, $\forall \boldsymbol{x} \in V$.

Since $\boldsymbol{A}$ and $\boldsymbol{B}$ are mappings, their composition may be defined. This is regarded as their *product* $\boldsymbol{AB}$:

$$(\boldsymbol{AB})\,\boldsymbol{x} \equiv \boldsymbol{A}(\boldsymbol{Bx}) \;\;, \quad \forall \boldsymbol{x} \in V.$$

Note that $\boldsymbol{AB}$ *is a linear operator on* V, hence belongs to $Op(V)$. In general, operator products are non-commutative: $\boldsymbol{AB} \neq \boldsymbol{BA}$. However, they are *associative* and distributive over addition:

$$(\boldsymbol{AB})\boldsymbol{C} = \boldsymbol{A}\,(\boldsymbol{BC}) \equiv \boldsymbol{ABC} \;\;, \quad \boldsymbol{A}\,(\boldsymbol{B}+\boldsymbol{C}) = \boldsymbol{AB}+\boldsymbol{AC} \;\; .$$

The *identity operator* $\boldsymbol{E}$ is the mapping of $Op(V)$ which leaves every element of V fixed: $\boldsymbol{E}\,\boldsymbol{x} = \boldsymbol{x}$. Thus, $\boldsymbol{AE}=\boldsymbol{EA}=\boldsymbol{A}$. Operators of the form $k\boldsymbol{E}$ ($k \in K$), called *scalar operators*, are commutative with all operators. In fact, any operator commutative with every operator of $Op(V)$ is a scalar operator.

It follows from the above that *the set* $Op(V)$ *of all linear operators on a given vector space* V *is an algebra*. This algebra is associative but (generally) non-commutative.

An operator $\boldsymbol{A}$ is said to be *invertible* if it represents a *bijective* (1-1) mapping, i.e., if it is an isomorphism of V onto itself. In this case, an *inverse operator* $\boldsymbol{A}^{-1}$ exists such that $\boldsymbol{AA}^{-1}= \boldsymbol{A}^{-1}\boldsymbol{A}=\boldsymbol{E}$. Practically this means that, if $\boldsymbol{A}$ maps $\boldsymbol{x} \in V$ onto $\boldsymbol{y} \in V$, then $\boldsymbol{A}^{-1}$ maps $\boldsymbol{y}$ back onto $\boldsymbol{x}$. For an invertible operator $\boldsymbol{A}$, $\mathrm{Ker}\,\boldsymbol{A}=\{\boldsymbol{0}\}$ and $\mathrm{Im}\,\boldsymbol{A}=V$.

Matrix representation of a linear operator

Let $\boldsymbol{A}$ be a linear operator on V. Let $\{\boldsymbol{e}_i\}$ (i=1,..., n) be a basis of V. Let

$$\boldsymbol{A}\,\boldsymbol{e}_k = \boldsymbol{e}_i\, A_{i\,k} \qquad \text{(sum on } i)$$

where the A_{ik} are real or complex, depending on whether V is a vector space over R or C. The $n \times n$ matrix $A=[A_{i\,k}]$ is called the *matrix of the operator* $\boldsymbol{A}$ *in the basis* $\{\boldsymbol{e}_i\}$.

Now, let $\boldsymbol{x}=x_i\,\boldsymbol{e}_i$ (sum on i) be a vector in V, and let $\boldsymbol{y}=\boldsymbol{A}\,\boldsymbol{x}$. If $\boldsymbol{y}= y_i\,\boldsymbol{e}_i$, then, by the linearity of $\boldsymbol{A}$,

$$y_i = A_{i\,k}\, x_k \qquad \text{(sun on } k) \; .$$

In matrix form,

$$[\,y\,]_{n\times 1} = [\,A\,]_{n\times n}\;[\,x\,]_{n\times 1} \;\; .$$

Next, let $\boldsymbol{A}$, $\boldsymbol{B}$ be linear operators on V. Define their product $\boldsymbol{C}=\boldsymbol{AB}$ by

$$\boldsymbol{C}\,\boldsymbol{x} = (\boldsymbol{AB})\,\boldsymbol{x} \equiv \boldsymbol{A}\,(\boldsymbol{Bx}) \;\;, \quad \boldsymbol{x} \in V \; .$$

Then, for any basis $\{\boldsymbol{e}_i\}$, $\boldsymbol{C}\,\boldsymbol{e}_k = \boldsymbol{A}\,(\boldsymbol{B}\,\boldsymbol{e}_k\,) = \boldsymbol{e}_i\, A_{i\,j}\, B_{j\,k} \equiv \boldsymbol{e}_i\, C_{i\,k} \quad \Rightarrow$

$$C_{ik} = A_{ij} B_{jk}$$

or, in matrix form,

$$C = A B \,.$$

That is,

> *the matrix of the product of two operators is the product of the matrices of these operators, in any basis of V.*

Consider now a change of basis defined by the *transition matrix* $T = [T_{ik}]$:

$$\boldsymbol{e}_k{}' = \boldsymbol{e}_i \, T_{ik} \;\; .$$

The inverse transformation is

$$\boldsymbol{e}_k = \boldsymbol{e}_i{}' \, (T^{-1})_{ik} \;\; .$$

Under this basis change, the matrix A of an operator $\boldsymbol{A}$ transforms as

$$A' = T^{-1} A\, T \quad (\textit{similarity transformation}) \,.$$

Under basis transformations, *the trace and the determinant of A remain unchanged*:

$$trA' = trA \;\; , \qquad detA' = detA \;\; .$$

An operator $\boldsymbol{A}$ is said to be *nonsingular* (*singular*) if $detA \neq 0$ ($detA = 0$). Note that this is a *basis-independent* property. *Any nonsingular operator is invertible*, i.e., there exists an inverse operator $\boldsymbol{A}^{-1} \in Op(V)$ such that $\boldsymbol{A}\,\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{E}$. Since an invertible operator represents a bijective mapping (i.e., both 1-1 and onto), it follows that $\mathrm{Ker}\boldsymbol{A} = \{\boldsymbol{0}\}$ and $\mathrm{Im}\boldsymbol{A} = V$. If $\boldsymbol{A}$ is invertible (nonsingular), then, for any basis $\{\boldsymbol{e}_i\}$ ($i = 1, \ldots, n$) of V, the vectors $\{\boldsymbol{A}\boldsymbol{e}_i\}$ are linearly independent and hence also constitute a basis.

Invariant subspaces and eigenvectors

Let V be an n-dimensional vector space over a field K, and let $\boldsymbol{A}$ be a linear operator on V. The subspace S of V is said to be *invariant under* $\boldsymbol{A}$ if, for every vector $\boldsymbol{x}$ of S, the vector $\boldsymbol{A}\boldsymbol{x}$ again belongs to S. Symbolically, $\boldsymbol{A}S \subseteq S$.

A vector $\boldsymbol{x} \neq \boldsymbol{0}$ is said to be an *eigenvector* of $\boldsymbol{A}$ if it generates a one-dimensional invariant subspace of V under $\boldsymbol{A}$. This means that an element $\lambda \in K$ exists, such that

$$\boldsymbol{A}\,\boldsymbol{x} = \lambda\,\boldsymbol{x} \;\; .$$

The element λ is called an *eigenvalue* of $\boldsymbol{A}$, to which eigenvalue the eigenvector $\boldsymbol{x}$ belongs. Note that, trivially, the null vector $\boldsymbol{0}$ is an eigenvector of $\boldsymbol{A}$, belonging to any eigenvalue λ. The set of all eigenvectors of $\boldsymbol{A}$, belonging to a given λ, is a subspace of V called the *proper subspace belonging to* λ.

It can be shown that *the eigenvalues of* $\boldsymbol{A}$ *are basis-independent quantities*. Indeed, let $A=[A_{ik}]$ be the $(n\times n)$ matrix representation of $\boldsymbol{A}$ in some basis $\{\boldsymbol{e}_i\}$ of V, and let $\boldsymbol{x}=x_i\boldsymbol{e}_i$ be an eigenvector belonging to λ. We denote by $X=[x_i]$ the column vector representing $\boldsymbol{x}$ in that basis. The eigenvalue equation for $\boldsymbol{A}$ is written, in matrix form,

$$A_{ik}\, x_k = \lambda\, x_i \quad \text{or} \quad A\, X = \lambda\, X \ .$$

This is written

$$(A-\lambda 1_n)\, X = 0 \ .$$

This equation constitutes a linear homogeneous system for $X=[x_i]$, which has a nontrivial solution iff

$$det\,(A-\lambda 1_n) = 0 \ .$$

This polynomial equation determines the eigenvalues λ_i $(i=1,...,n)$ (not necessarily all different from each-other) of $\boldsymbol{A}$. Since the determinant of the matrix representation of an operator [in particular, of the operator $(\boldsymbol{A}-\lambda\boldsymbol{E})$ for any given λ] is a basis-independent quantity, it follows that, if the above equation is satisfied for a certain λ in a certain basis (where $\boldsymbol{A}$ is represented by the matrix A), it will also be satisfied *for the same* λ in any other basis (where $\boldsymbol{A}$ is represented by another matrix, say, A'). We conclude that *the eigenvalues of an operator are a property of the operator itself and do not depend on the choice of basis of V*.

If we can find n linearly independent eigenvectors $\{\boldsymbol{x}_i\}$ of $\boldsymbol{A}$, belonging to the corresponding eigenvalues λ_i , we can use these vectors to define a basis for V. In this basis, the matrix representation of $\boldsymbol{A}$ has a particularly simple *diagonal* form:

$$A = diag\,(\lambda_1\, , \,...\, , \lambda_n\,) \ .$$

Using this expression, and the fact that the quantities trA, $detA$ and λ_i are invariant under basis transformations, we conclude that, in *any* basis of V,

$$trA = \lambda_1 + \lambda_2 + ... + \lambda_n \ , \qquad detA = \lambda_1\, \lambda_2 ... \lambda_n \ .$$

We note, in particular, that *all eigenvalues of an invertible (nonsingular) operator are nonzero*. Indeed, if even one is zero, then $detA=0$ and $\boldsymbol{A}$ is singular.

An operator $\boldsymbol{A}$ is called *nilpotent* if $\boldsymbol{A}^m=\boldsymbol{0}$ for some natural number $m>1$. The smallest such value of m is called the *degree of nilpotency*, and it cannot exceed n. *All eigenvalues of a nilpotent operator are zero*. Thus, such an operator is *singular* (non-invertible).

An operator $\boldsymbol{A}$ is called *idempotent* (or *projection operator*) if $\boldsymbol{A}^2=\boldsymbol{A}$. It follows that $\boldsymbol{A}^m=\boldsymbol{A}$, for any natural number m. *The eigenvalues of an idempotent operator can take the values 0 or 1*.

BASIC MATRIX PROPERTIES

$(A+B)^T = A^T + B^T$; $(AB)^T = B^T A^T$

$(A+B)^\dagger = A^\dagger + B^\dagger$; $(AB)^\dagger = B^\dagger A^\dagger$ where $M^\dagger \equiv (M^T)^* = (M^*)^T$

$(kA)^T = kA^T$; $(kA)^\dagger = k^* A^\dagger$ $(k \in C)$

$(AB)^{-1} = B^{-1}A^{-1}$; $(A^T)^{-1} = (A^{-1})^T$; $(A^\dagger)^{-1} = (A^{-1})^\dagger$

$[A,B]^T = [B^T, A^T]$; $[A,B]^\dagger = [B^\dagger, A^\dagger]$

$$A^{-1} = \frac{1}{\det A}\, adjA \quad (\det A \neq 0)$$

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$tr(\kappa A + \lambda B) = \kappa\, trA + \lambda\, trB$

$trA^T = trA$; $trA^\dagger = (trA)^*$

$tr(AB) = tr(BA)$, $tr(ABC) = tr(BCA) = tr(CAB)$, *etc.*

$tr[A,B] = 0$

$\det A^T = \det A$; $\det A^\dagger = (\det A)^*$

$\det(AB) = \det(BA) = \det A \cdot \det B$

$\det(A^{-1}) = 1/\det A$

$\det(cA) = c^n \det A \quad \left(c \in C,\ A \in gl(n,C)\right)$

If any row or column of A is multiplied by c, then so is $\det A$.

$[A,B] = -[B,A] \equiv AB - BA$

$[A,\ B+C] = [A,B] + [A,C]$; $[A+B,\ C] = [A,C] + [B,C]$

$[A,BC] = [A,B]C + B[A,C]$; $[AB,C] = A[B,C] + [A,C]B$

$[A,\ [B,C]] + [B,\ [C,A]] + [C,\ [A,B]] = 0$

$[[A,B],\ C] + [[B,C],\ A] + [[C,A],\ B] = 0$

Let $A = A(t) = [a_{ij}(t)]$, $B = B(t) = [b_{ij}(t)]$, be $(n\times n)$ matrices. The derivative of A (similarly, of B) is the $(n\times n)$ matrix dA/dt , with elements

$$\left(\frac{dA}{dt}\right)_{ij} = \frac{d}{dt}\, a_{ij}(t) \ .$$

The integral of A (similarly, of B) is the $(n\times n)$ matrix defined by

$$\left(\int A(t)\,dt\right)_{ij} = \int a_{ij}(t)\,dt \ .$$

$$\frac{d}{dt}(A \pm B) = \frac{dA}{dt} \pm \frac{dB}{dt} \quad ; \quad \frac{d}{dt}(AB) = \frac{dA}{dt}B + A\frac{dB}{dt}$$

$$\frac{d}{dt}[A, B] = \left[\frac{dA}{dt}, B\right] + \left[A, \frac{dB}{dt}\right]$$

$$\frac{d}{dt}(A^{-1}) = -A^{-1}\frac{dA}{dt}A^{-1} \quad \Rightarrow \quad d\,(A^{-1}) = -A^{-1}(dA)\,A^{-1}$$

$$tr\left(\frac{dA}{dt}\right) = \frac{d}{dt}(trA)$$

Let $A = A(x, y)$. Call $\partial A / \partial x \equiv \partial_x A \equiv A_x$, etc.:

$$\partial_x(A^{-1}A_y) - \partial_y(A^{-1}A_x) + [A^{-1}A_x\,,\ A^{-1}A_y] = 0$$

$$\partial_x(A_yA^{-1}) - \partial_y(A_xA^{-1}) - [A_xA^{-1},\ A_yA^{-1}] = 0$$

$$A(A^{-1}A_x)_y\,A^{-1} = (A_yA^{-1})_x \quad \Leftrightarrow \quad A^{-1}(A_yA^{-1})_x\,A = (A^{-1}A_x)_y$$

$$e^A \equiv \exp A = \sum_{n=0}^{\infty}\frac{A^n}{n!} = 1 + A + \frac{A^2}{2} + \cdots$$

$$B\,e^A B^{-1} = e^{BAB^{-1}}$$

$$\left(e^A\right)^* = e^{A^*} \; ; \quad \left(e^A\right)^T = e^{A^T} \; ; \quad \left(e^A\right)^\dagger = e^{A^\dagger} \; ; \quad \left(e^A\right)^{-1} = e^{-A}$$

$e^A e^B = e^B e^A = e^{A+B}$ when $[A, B] = 0$

In general, $e^A e^B = e^C$ where

$$C = A + B + \frac{1}{2}[A, B] + \frac{1}{12}\big([A, [A, B]] + [B, [B, A]]\big) + \cdots$$

By definition, $\log B = A \Leftrightarrow B = e^A$.

$$\det\left(e^A\right) = e^{trA} \quad \Leftrightarrow \quad \det B = e^{tr\,(\log B)} \quad \Leftrightarrow \quad tr\,(\log B) = \log(\det B)$$

$\det(1 + \delta A) \simeq 1 + tr\,\delta A$, for infinitesimal δA

$$tr\,(A^{-1}A_x) = tr\,(A_xA^{-1}) = tr\,(\log A)_x = [tr\,(\log A)]_x = [\log(\det A)]_x$$

DETERMINANTS

Consider the 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} .$$

The *determinant* of *A* is defined by

$$\det A \equiv \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc \quad . \qquad (1)$$

Next, consider the 3×3 matrix

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & k \end{bmatrix} .$$

To evaluate its determinant, we work as follows: First, we draw a 3×3 "chessboard" consisting of + (plus) and – (minus) signs, as shown below. *Careful:* At the *top left* we always put a *plus* sign!

$$\begin{vmatrix} + & - & + \\ - & + & - \\ + & - & + \end{vmatrix} .$$

We may now develop the determinant of *A with respect to any row or any column*; the result will always be the same. Let us assume, e.g., that we choose to develop with respect to the *first row*. Its first element is *a*. At the position where this element is located (top left) the "chessboard" has a + sign; we thus leave the sign of *a* unchanged. Imagine now that we *cross off both the row and the column to which this element belongs* (first row, first column in this case). What is left over is a lower-order, 2×2 matrix with determinant

$$\begin{vmatrix} e & f \\ h & k \end{vmatrix} .$$

We multiply this determinant by *a* and we save the result.

The second element in the first row is *b*. At its location, the chessboard has a – sign; we thus write –*b*. We "cross off" the row and the column to which *b* belongs (first row, second column) and we get the 2×2 determinant

$$\begin{vmatrix} d & f \\ g & k \end{vmatrix} \ .$$

We multiply this by –*b* and we save this result, too.

The third element in the first row is *c*. At its location the chessboard has a + sign, thus we leave the sign of *c* unchanged. Crossing off the first row and the third column, where *c* is located, we find the determinant

$$\begin{vmatrix} d & e \\ g & h \end{vmatrix} \ .$$

We multiply this by *c* and again we save this result in the "memory".

Summing the contents of the memory, we finally find the determinant of *A*:

$$\det A \equiv \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = a\begin{vmatrix} e & f \\ h & k \end{vmatrix} - b\begin{vmatrix} d & f \\ g & k \end{vmatrix} + c\begin{vmatrix} d & e \\ g & h \end{vmatrix} \ . \qquad (2)$$

Of course, to complete the job we must evaluate the minor determinants according to Eq. (1), which is an easy task.

Exercise: Evaluate again the determinant of *A*, this time by developing with respect to the *second column*, and show that

$$\det A \equiv \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & k \end{vmatrix} = -b\begin{vmatrix} d & f \\ g & k \end{vmatrix} + e\begin{vmatrix} a & c \\ g & k \end{vmatrix} - h\begin{vmatrix} a & c \\ d & f \end{vmatrix} \ .$$

Verify that your result is the same as before.

Exercise: With the aid of the chessboard

$$\begin{vmatrix} + & - \\ - & + \end{vmatrix}$$

(the + sign always on the *top left*!) and by following an analogous procedure, verify formula (1) for a 2×2 determinant. (By definition, the determinant of a 1×1 matrix [*a*] is equal to the single element of the matrix.)

For a 4×4 matrix, the chessboard is of the form (with a + sign always on the top left)

$$\begin{vmatrix} + & - & + & - \\ - & + & - & + \\ + & - & + & - \\ - & + & - & + \end{vmatrix} .$$

The development of a 4×4 determinant leads to 3×3 determinants that are developed as shown previously. As is obvious, the problem becomes harder as the dimension of the determinant increases!

Exercise: Show that

$$\begin{vmatrix} 1 & -1 & 1 \\ 2 & 0 & -2 \\ -1 & 1 & -1 \end{vmatrix} = 0 ,$$

by developing with respect to a row and, again, with respect to a column. Choose the row and column that will make your calculations easier. (Obviously, as a general rule, it is in our best interest to choose a row or a column with *as many zeros as possible*!)

Properties of determinants

Let A be an $n \times n$ matrix and let $\det A$ be the determinant of A. The following statements are true:

1. If all elements of a row or a column of A are zero, then $\det A = 0$.

2. If every element of a row or a column of A is multiplied by λ, then $\det A$ is multiplied by λ as well.

3. If *all* elements of A are multiplied by λ, then $\det A$ is multiplied by λ^n (where n is the dimension of A). That is,

$$\det(\lambda A) = \lambda^n \det A .$$

4. If any two rows or any two columns of A are interchanged, the value of $\det A$ is multiplied by (–1).

5. If two rows or two columns of A are identical, then $\det A = 0$. The same is true, more generally, if two rows or two columns are multiples of each other.

6. The value of det$A$ remains the same if the rows and columns of A are interchanged. That is,

$$\det(A^T) = \det A \ ,$$

where A^T is the *transpose* of A: $(A^T)_{ij} = A_{ji}$.

7. If A and B are $n\times n$ matrices,

$$\det(AB) = \det(BA) = \det A \,.\, \det B \ .$$

Also,

$$\det(A^k) = (\det A)^k \ , \quad k=1,2,3,\dots \ .$$

8. If A^{-1} is the *inverse* of A (see below),

$$\det(A^{-1}) = 1/\det A \ .$$

9. The determinant of a *diagonal* (or, more generally, a *triangular*) matrix A is equal to the product of the elements of the diagonal of A.

10. The value of det$A$ is unchanged if to any row or any column of A we add an arbitrary multiple of any other row or column, respectively.

Evaluation of a matrix inverse

Consider a 3×3 matrix A:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \equiv [a_{ij}] \qquad (i, j = 1,2,3) \ .$$

Let a_{ij} be an arbitrary element of A (the one that belongs to the i-th row and the j-th column). By "crossing off" the row and the column to which a_{ij} belongs, we obtain a 2×2 matrix. We call D_{ij} the determinant of this latter matrix.

We now construct a 3×3 matrix M, as follows: We replace every element a_{ij} of the given matrix A by the corresponding quantity

$$(-1)^{i+j} D_{ij} \ .$$

That is, in place of a_{ij} we put the minor determinant D_{ij} *multiplied by the sign that exists on the chessboard at the position of* a_{ij} . We thus get

$$M = \begin{bmatrix} D_{11} & -D_{12} & D_{13} \\ -D_{21} & D_{22} & -D_{23} \\ D_{31} & -D_{32} & D_{33} \end{bmatrix} .$$

Finally, we take the *transpose* M^T of M, which is called the *adjoint* of the matrix A:

$$\text{adj}\, A = M^T = \begin{bmatrix} D_{11} & -D_{21} & D_{31} \\ -D_{12} & D_{22} & -D_{32} \\ D_{13} & -D_{23} & D_{33} \end{bmatrix} .$$

The *inverse* A^{-1} of A, satisfying $AA^{-1} = A^{-1}A = I$ (where I is the 3×3 unit matrix) is given by

$$A^{-1} = \frac{1}{\det A}\ \text{adj}\, A \qquad (3)$$

Obviously, a necessary condition in order that the inverse of A may exist (i.e., in order that the matrix A be *invertible*) is that $\det A \neq 0$. The process described above, leading to relation (3), is generally valid for *any* $n\times n$ matrix (n=2,3,4,...).

Exercise: For the 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} ,$$

show that

$$A^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} .$$

Verify that

$$AA^{-1} = A^{-1}A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} .$$

Exercise: By using (3), show that

$$\begin{bmatrix} 0 & 1 & -3 \\ -1 & -1 & 3 \\ 0 & 1 & -1 \end{bmatrix}^{-1} = \begin{bmatrix} -1 & -1 & 0 \\ -1/2 & 0 & 3/2 \\ -1/2 & 0 & 1/2 \end{bmatrix} .$$

Verify that your result satisfies the relation $AA^{-1} = A^{-1}A = I$.

Solution of linear systems

The method we will describe applies to any *linear system of equations*; i.e., system of n linear equations with n unknowns (n=2,3,4,...). For simplicity, we consider a system of two equations:

$$\begin{aligned} a_{11} x_1 + a_{12} x_2 &= b_1 \\ a_{21} x_1 + a_{22} x_2 &= b_2 \end{aligned} \tag{4}$$

In matrix form, this is written

$$A\mathbf{x} = \mathbf{b} \quad \Leftrightarrow \quad \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \tag{5}$$

where A is the matrix of the coefficients of the unknowns, $\mathbf{x}$ is the column vector of the unknowns and $\mathbf{b}$ is the column vector of the constants. In the case where $\mathbf{b}=0 \Leftrightarrow b_1=b_2=0$, the given system is said to be *homogeneous linear*.

We note the following:

1. If $\det A \neq 0$, the matrix A is *invertible* and the system has a *unique solution* that is obtained as follows:

$$A\mathbf{x} = \mathbf{b} \;\Rightarrow\; A^{-1}(A\mathbf{x}) = A^{-1}\mathbf{b} \;\Rightarrow\; (A^{-1}A)\,\mathbf{x} = A^{-1}\mathbf{b} \;\Rightarrow$$

$$\mathbf{x} = A^{-1}\mathbf{b} \tag{6}$$

In the case where $\mathbf{b}=0$ (homogeneous system), the only solution of the system is the trivial one: $\mathbf{x}=0 \Leftrightarrow x_1=x_2=0$.

2. If $\det A=0$ (the matrix A is *non-invertible*), the system either has no solution (is *inconsistent*) or has an *infinite number* of solutions (see below).

The difficulty in solving (6) lies in the necessity of determining the inverse matrix. Let us now see an alternative expression for the solution of the system, based on *Cramer's method* (or *method of determinants*). As before, we call A the matrix of the coefficients of the unknowns in system (4):

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} .$$

Furthermore, we call A_1 the matrix obtained from A by replacement of its first column (i.e., the column of the coefficients a_{11} and a_{21} of x_1) with the column of the constant terms b_1 and b_2 . Similarly, we call A_2 the matrix obtained from A by replacing its second column (the one with the coefficients of x_2) with the column of the constants. Analytically,

$$A_1 = \begin{bmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{bmatrix} \quad , \quad A_2 = \begin{bmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{bmatrix} .$$

Then, the solution of system (4) – if it exists – is written

$$x_1 = \frac{\det A_1}{\det A} \quad , \quad x_2 = \frac{\det A_2}{\det A} \tag{7}$$

The determinants of the matrices A_1 and A_2 are called *Cramer's determinants*.

Exercise: Write the analytical expression of the general solution (7), for any given a_{ij} and b_i .

Exercise: Consider the system

$$a\,x + b\,y = c$$
$$e\,x + f\,y = g$$

(where we have put $x_1=x$, $x_2=y$) . Show that its solution is

$$x = \frac{c\,f - bg}{a\,f - be} \quad , \quad y = \frac{ag - ce}{a\,f - be} \quad .$$

Assume now that we "rewrite" the system by inverting the order of the two equations:

$$e\,x + f\,y = g$$
$$a\,x + b\,y = c$$

Must we expect a different solution? How is your answer related to the properties of determinants?

More generally, for a linear system of n equations with n unknowns,

$$\begin{aligned} &a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ &a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ &\vdots \\ &a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{aligned} \tag{8}$$

the solution is written

$$x_i = \frac{\det A_i}{\det A} \quad , \quad i = 1, 2, \cdots, n \tag{9}$$

where A is the $n\times n$ matrix of the coefficients a_{jk} of the unknowns, while A_i is the matrix obtained from A by replacing the column of the coefficients of x_i with the column of the constants b_k .

We note the following:

1. If $\det A \neq 0$ (i.e., if the matrix A is invertible) a unique solution (9) of the system (8) exists.

2. If $\det A=0$ (the matrix A is *not* invertible) and if *even one* of the Cramer determinants $\det A_k$ in (9) is non-vanishing, the system (8) *has no solution* (is *inconsistent*), as follows from (9).

3. If $\det A=0$ and if *all* Cramer determinants $\det A_k$ (k=1,2,...,n) are zero, the system (8) has an *infinite number* of solutions.

Particularly significant for applications is the case of a *homogeneous* system, in which all constant terms b_k (k=1,2,...,n) are zero:

$$
\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = 0 \\
&a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = 0 \\
&\vdots \\
&a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = 0
\end{aligned}
\tag{10}
$$

In this case *all* Cramer determinants $\det A_k$ (k=1,2,...,n) are zero (explain this!). The following possibilities thus exist:

1. If the determinant of the matrix A of the coefficients of the unknowns is non-zero ($\det A\neq 0$), the only possible solution of the system (10) is the *trivial solution* $x_1 = x_2 = \ldots = x_n = 0$, as follows from (9).

2. If $\det A=0$, the system (10) admits an *infinite number* of nontrivial solutions.

Exercise: Show the following: (*a*) A homogeneous linear system always has a solution, i.e., is never inconsistent. (*b*) For such a system to possess a *nontrivial* solution (different, that is, from the zero solution) the determinant of the matrix of coefficients of the unknowns must be zero.

Example: Consider the homogeneous system

$$
\begin{aligned}
2x - y &= 0 \\
-6x + 3y &= 0
\end{aligned}
$$

(where we have put $x_1=x$, $x_2=y$). The determinant of the coefficients of the unknowns is

$$\begin{vmatrix} 2 & -1 \\ -6 & 3 \end{vmatrix} = 6-6=0 \ .$$

This occurs because the second line is a multiple (by –3) of the first. And this, in turn, reflects the fact that the equations in the system *are not independent* of each other (the second one is just a multiple of the first, thus does not provide any useful new information). The only thing we can say is that $y=2x$, with *arbitrary* x. This means that the system has an *infinite number* of solutions, one for each chosen value of x.

Application to the vector product

Consider the vectors

$$\vec{A} = A_x\,\hat{u}_x + A_y\,\hat{u}_y + A_z\,\hat{u}_z \equiv (A_x, A_y, A_z) \ ,$$
$$\vec{B} = B_x\,\hat{u}_x + B_y\,\hat{u}_y + B_z\,\hat{u}_z \equiv (B_x, B_y, B_z) \ ,$$

where $\hat{u}_x$, $\hat{u}_y$, $\hat{u}_z$ are the *unit vectors* on the axes x, y, z, respectively, of a standard Cartesian system. As we know from vector analysis, the *vector product* (or "cross product") of $\vec{A}$ and $\vec{B}$ can be written in determinant form, as follows:

$$\vec{A}\times\vec{B} = \begin{vmatrix} \hat{u}_x & \hat{u}_y & \hat{u}_z \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix} \ .$$

Moreover, the necessary condition in order that $\vec{A}$ and $\vec{B}$ be *parallel* to each other is $\vec{A}\times\vec{B}=0$.

Example: Find the values of α and β for which the vectors $\vec{A}\equiv(1,\alpha,3)$ and $\vec{B}\equiv(-2,-4,\beta)$ are parallel to each other.

Solution: We must have $\vec{A}\times\vec{B}=0 \ \Rightarrow$

$$\begin{vmatrix} \hat{u}_x & \hat{u}_y & \hat{u}_z \\ 1 & \alpha & 3 \\ -2 & -4 & \beta \end{vmatrix} = 0 \ \Rightarrow \ \hat{u}_x\,(\alpha\beta+12) - \hat{u}_y\,(\beta+6) + \hat{u}_z\,(-4+2\alpha) = 0$$

(where the determinant has been developed with respect to the first row, i.e., the row of the unit vectors). Given that the unit vectors constitute a linearly independent set, the only way the above equality may be satisfied is by setting all three coefficients of the corresponding unit vectors equal to zero. We thus obtain a system of *three* equations with *two* unknowns:

$$2\alpha - 4 = 0\,, \qquad \beta + 6 = 0\,, \qquad \alpha\beta + 12 = 0\,.$$

The first two equations yield $\alpha=2,\ \beta=-6$. The third equation simply verifies this result. That is, the third equation is *compatible* with the other two but furnishes no additional information, since this last equation *is not independent* of the preceding ones but follows directly from them. Note that, with the values of α and β found above, the third row of the determinant that represents $\vec{A}\times\vec{B}$ becomes a multiple (by –2) of the second row, so that the determinant automatically vanishes.

Exercise: Show that no values of α and β exist for which the vectors $\vec{A}\equiv(1,\alpha,3)$ and $\vec{B}\equiv(-2,\beta,6)$ are parallel to each other.

Exercise: Show that there is an *infinite* number of values of α and β for which the vectors $\vec{A}\equiv(1,\alpha,3)$ and $\vec{B}\equiv(-2,\beta,-6)$ are parallel to each other. What relation must exist between α and β?

THE EXPONENTIAL FUNCTION

Problem: Let a be a positive real number. We know how to define $a^{m/n}$ with m, n integers. But, how do we define a^x for a general, real x that may be an *irrational* number, i.e., cannot be written as a quotient of integers m, n ?

Well, if it is difficult to define a function directly, we may try defining the inverse function (assuming it exists). To this end, we consider the function

$$\ln x = \int_1^x \frac{1}{t}\, dt\ , \quad x > 0 \qquad (1)$$

Then,

$$(\ln x)' = 1/x$$

where the prime denotes differentiation with respect to x. Note in particular that $\ln 1=0$. It can also be shown [1] that, for $a, b \in R^+$, $\ln(ab)=\ln a+\ln b$, $\ln(a/b)=\ln a-\ln b$. Thus, $\ln x$ is a logarithmic function in the usual sense.

The function $\ln x$ is increasing for $x>0$ (indeed, its derivative $1/x$ is positive for $x>0$). Since $\ln x$ is monotone, this function is invertible. Call $\exp x$ the inverse of $\ln x$. That is,

$$y = \exp x \quad \Leftrightarrow \quad x = \ln y\ .$$

This means that

$$\exp(\ln y) = y \quad \text{and} \quad \ln(\exp x) = x\ .$$

It can be shown [1] that $\exp x$ is an exponential function in the usual sense; i.e., it has the form $\exp x = e^x$ for some $e>0$, to be determined. We write

$$y = e^x \quad \Leftrightarrow \quad x = \ln y \quad (x \in R\ ,\ y \in R^+)$$

so that

$$e^{\ln y} = y \quad \text{and} \quad \ln(e^x) = x\ .$$

Note in particular that, for $x=0$ we have $e^0=1$ and $\ln 1=0$, as required. Also, for $x=1$ we have that $\ln(e^1) = 1$ and, by the definition (1) of the logarithmic function,

$$\ln e = \int_1^e \frac{1}{t}\, dt = 1\ .$$

We will now show that the function e^x ($x \in R$) can be expressed as the limit of a certain infinite sequence:

$$e^x = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \quad (x \in R) \qquad (2)$$

Then, for any $a \in R^+$ we will have that $a = e^{\ln a} \Rightarrow$

$$a^x = e^{x \ln a} = \lim_{n\to\infty}\left(1+\frac{x\ln a}{n}\right)^n .$$

Proposition 1: Given a function $u = f(x)$ that assumes positive values for all x in its domain of definition, the derivative of $\ln[f(x)]$ is given by

$$\frac{d}{dx}\ln f(x) = \frac{f'(x)}{f(x)} \tag{3}$$

Proof: $\dfrac{d}{dx}\ln f(x) = \dfrac{d(\ln u)}{du}\dfrac{du}{dx} = \dfrac{1}{u}\dfrac{du}{dx} = \dfrac{f'(x)}{f(x)}$.

Proposition 2: The derivative of e^x is given by $(e^x)' = e^x$.

Proof: $\ln(e^x) = x \Rightarrow [\ln(e^x)]' = 1 \Rightarrow (e^x)'/\, e^x = 1 \Rightarrow (e^x)' = e^x$, where we have used relation (3) for the derivative of $\ln(e^x)$.

Corollary: $[\exp f(x)]' = f'(x)\exp f(x)$.

Now, consider the function $g(x) = \lim_{n\to\infty}\left(1+\dfrac{x}{n}\right)^n$ $(x \in R)$. We have:

$$g'(x) = \lim_{n\to\infty}\left[n\left(1+\frac{x}{n}\right)^{n-1}\frac{1}{n}\right] = \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^{n-1} = \lim_{n\to\infty}\left[\left(1+\frac{x}{n}\right)^n\left(1+\frac{x}{n}\right)^{-1}\right]$$

$$= \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n \cdot \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^{-1} = g(x)\cdot 1 = g(x).$$

Moreover, $g(0)=1$. Hence the function $y=g(x)$ satisfies the differential equation $y'=y$ with initial condition $y=1$ for $x=0$. On the other hand, the function $y=e^x$ satisfies the same differential equation with the same initial condition. Since the solution of this differential equation with given initial condition is unique, we conclude that the functions $g(x)$ and e^x must be identical. Therefore relation (2) must be true.

We note that, for $x=1$, Eq. (2) gives

$$e = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n \quad (\simeq 2.72) \tag{4}$$

This is the formula by which the number e is usually defined.

In the same spirit we may show that another possible representation of the exponential function e^x is in the form of a power (Maclaurin) series:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots \quad (x \in R) \tag{5}$$

Indeed, notice that the x-derivative of this series is the series itself, as well as that the value of the series is equal to 1 for x=0. Although expressions (2) and (5) do not look alike, they represent the *same* function, $\exp x$! (*Note:* Two functions of x are considered identical if they have the same domain D of definition and assume equal values for all $x \in D$.)

We defined a^x (a>0, $x \in R$) in a rather indirect way by first defining the function e^x as the inverse of the function $\ln x$ and then by writing $a^x = e^{x \ln a}$. There is, however, a more direct definition of a^x. Let $x_1, x_2, \ldots, x_n, \ldots$ be *any* infinite sequence of *rational* numbers x_n such that $\lim_{n \to \infty} x_n = x \in R$. [Question: Can a sequence of rational numbers have an *irrational* limit? Yes! See, e.g., the expression (4) for e, where the latter number *is* irrational (see, e.g., [2]).] We now define a^x as follows:

$$a^x = \lim_{n \to \infty} a^{x_n} \quad (a > 0,\ x \in R).$$

Since x_n is a rational number for all n, raising a to a rational number should not be a problem! Note that the value of a^x does not depend on the specific choice of the sequence x_n, as long as the limit of this sequence is x.

References

[1] D. D. Berkey, *Calculus*, 2nd Edition (Saunders College, 1988), Chap. 8.

[2] https://mindyourdecisions.com/blog/2015/06/18/lets-prove-e-2-718-is-irrational-3-methods/