KARL SVOZIL

MATHEMATICAL METHODS OF THEORETICAL PHYSICS

EDITION FUNZL

Contents

2   *Multilinear Algebra and Tensors   75*

Introduction

THIS IS A FOURTH ATTEMPT to provide some written material of a course in mathemathical methods of theoretical physics. I have presented this course to an undergraduate audience at the Vienna University of Technology. Only God knows (see Ref. [1] part one, question 14, article 13; and also Ref. [2], p. 243) if I have succeeded to teach them the subject! I kindly ask the perplexed to please be patient, do not panic under any circumstances, and do not allow themselves to be too upset with mistakes, omissions & other problems of this text. At the end of the day, everything will be fine, and in the long run we will be dead anyway.

THE PROBLEM with all such presentations is to learn the formalism in sufficient depth, yet not to get buried by the formalism. As every individual has his or her own mode of comprehension there is no canonical answer to this challenge.

I AM RELEASING THIS text to the public domain because it is my conviction and experience that content can no longer be held back, and access to it be restricted, as its creators see fit. On the contrary, in the *attention economy* – subject to the scarcity as well as the compound accumulation of attention – we experience a push toward so much content that we can hardly bear this information flood, so we have to be selective and restrictive rather than aquisitive. I hope that there are some readers out there who actually enjoy and profit from the text, in whatever form and way they find appropriate.

SUCH UNIVERSITY TEXTS AS THIS ONE – and even recorded video transcripts of lectures – present a transitory, almost outdated form of teaching. Future generations of students will most likely enjoy *massive open online courses* (MOOCs) that might integrate interactive elements and will allow a more individualized – and at the same time automated – form of learning. What is most important from the viewpoint of university administrations is that (i) MOOCs are cost-effective (that is, cheaper than standard tuition) and (ii) the know-how of university teachers and researchers gets transferred to the university administration and management; thereby the dependency of the university management on

"It is not enough to have no concept, one must also be capable of expressing it." From the German original in *Karl Kraus,* Die Fackel **697**, *60 (1925)*: *"Es genügt nicht, keinen Gedanken zu haben: man muss ihn auch ausdrücken können."*

[1] Thomas Aquinas. *Summa Theologica. Translated by Fathers of the English Dominican Province.* Christian Classics Ethereal Library, Grand Rapids, MI, 1981. URL http://www.ccel.org/ccel/aquinas/summa.html

[2] Ernst Specker. Die Logik nicht gleichzeitig entscheidbarer Aussagen. *Dialectica*, 14(2-3):239–246, 1960. DOI: 10.1111/j.1746-8361.1960.tb00422.x. URL http://doi.org/10.1111/j.1746-8361.1960.tb00422.x

teaching staff is considerably alleviated. In the latter way, MOOCs are the implementation of assembly line methods (first introduced by Henry Ford for the production of affordable cars) in the university setting. Together with "scientometric" methods which have their origin in both Bolshevism as well as in Taylorism [3], automated teaching is transforming schools and universites, and in particular, the old Central European universities, as much as the *Ford Motor Company* (NYSE:F) has transformed the car industry, and the Soviets have transformed Czarist Russia. To this end, for better or worse, university teachers become accountants [4], and *"science becomes bureaucratized; indeed, a higher police function. The retrieval is taught to the professors.* [5]*"*

TO NEWCOMERS in the area of theoretical physics (and beyond) I strongly recommend to consider and acquire two related proficiencies:

- to learn to speak and publish in LaTeX and BibTeX; in particular, in the implementation of TeX Live. LaTeX's various dialects and formats, such as REVTeX, provide a kind of template for structured scientific texts, thereby assisting you writing and publishing consistently and with methodologic rigour;

- to subsribe to and browse through preprints published at the website `arXiv.org`, which provides open access to more than three quarters of a million scientific texts; most of them written in and compiled by LaTeX. Over time, this database has emerged as a *de facto* standard from the initiative of an individual researcher working at the *Los Alamos National Laboratory* (the site at which also the first nuclear bomb has been developed and assembled). Presently it happens to be administered by *Cornell University.* I suspect (this is a personal subjective opinion) that (the successors of) `arXiv.org` will eventually bypass if not supersede most scientific journals of today.

So this very text is written in LaTeX and accessible freely *via* `arXiv.org` under eprint number *arXiv:1203.4558.*

MY OWN ENCOUNTER with many researchers of different fields and different degrees of formalization has convinced me that there is no single, unique "optimal" way of formally comprehending a subject [6]. With regards to formal rigour, there appears to be a rather questionable chain of contempt – all too often theoretical physicists look upon the experimentalists suspiciously, mathematical physicists look upon the theoreticians skeptically, and mathematicians look upon the mathematical physicists dubiously. I have even experienced the distrust formal logicians expressed about their collegues in mathematics! For an anectodal evidence, take the claim of a prominent member of the mathematical physics community, who once dryly remarked in front of a fully packed audience, "what other people call 'proof' I call 'conjecture'!"

SO PLEASE BE AWARE that not all I present here will be acceptable to everybody; for various reasons. Some people will claim that I am too confusing and utterly formalistic, others will claim my arguments are in

[3] Frederick Winslow Taylor. *The Principles of Scientific Management.* Harper Bros., New York, NY, 1911. URL https://archive.org/details/principlesofscie00taylrich

[4] Karl Svozil. Versklavung durch Verbuchhalterung. *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare*, 60(1):101–111, 2013. URL http://eprints.rclis.org/19560/

[5] Ernst Jünger. *Heliopolis. Rückblick auf eine Stadt.* Heliopolis-Verlag Ewald Katzmann, Tübingen, 1949

German original *"die Wissenschaft wird bürokratisiert, ja Funktion der höheren Polizei. Den Professoren wird das Apportieren beigebracht."*
If you excuse a maybe utterly displaced comparison, this might be tantamount only to studying the Austrian family code ("Ehegesetz") from §49 onward, available through http://www.ris.bka.gv.at/Bundesrecht/ before getting married.

http://arxiv.org/abs/1203.4558

[6] Philip W. Anderson. More is different. Broken symmetry and the nature of the hierarchical structure of science. *Science*, 177(4047):393–396, August 1972. DOI: 10.1126/science.177.4047.393. URL http://doi.org/10.1126/science.177.4047.393

desparate need of rigour. Many formally fascinated readers will demand to go deeper into the meaning of the subjects; others may want some easy-to-identify pragmatic, syntactic rules of deriving results. I apologise to both groups from the outset. This is the best I can do; from certain different perspectives, others, maybe even some tutors or students, might perform much better.

I AM CALLING for more tolerance and a greater unity in physics; as well as for a greater esteem on "both sides of the same effort;" I am also opting for more pragmatism; one that acknowledges the mutual benefits and oneness of theoretical and empirical physical world perceptions. Schrödinger [7] cites Democritus with arguing against a too great separation of the intellect ($\delta\iota\alpha\nu o\iota\alpha$, dianoia) and the senses ($\alpha\iota\sigma\theta\eta\sigma\epsilon\iota\varsigma$, aitheseis). In fragment D 125 from Galen [8], p. 408, footnote 125 , the intellect claims "ostensibly there is color, ostensibly sweetness, ostensibly bitterness, actually only atoms and the void;" to which the senses retort: "Poor intellect, do you hope to defeat us while from us you borrow your evidence? Your victory is your defeat."

In his 1987 *Abschiedsvorlesung* professor Ernst Specker at the *Eidgenössische Hochschule Zürich* remarked that the many books authored by David Hilbert carry his name first, and the name(s) of his co-author(s) second, although the subsequent author(s) had actually written these books; the only exception of this rule being Courant and Hilbert's 1924 book *Methoden der mathematischen Physik*, comprising around 1000 densly packed pages, which allegedly none of these authors had actually written. It appears to be some sort of collective effort of scholars from the University of Göttingen.

So, in sharp distinction from these activities, I most humbly present my own version of what is important for standard courses of contemporary physics. Thereby, I am quite aware that, not dissimilar with some attempts of that sort undertaken so far, I might fail miserably. Because even if I manage to induce some interest, affaction, passion and understanding in the audience – as Danny Greenberger put it, inevitably four hundred years from now, all our present physical theories of today will appear transient [9], if not laughable. And thus in the long run, my efforts will be forgotten (although, I do hope, not totally futile); and some other brave, courageous guy will continue attempting to (re)present the most important mathematical methods in theoretical physics.

ALL THINGS CONSIDERED, it is mind-boggling why formalized thinking and numbers utilize our comprehension of nature. Even today eminent researchers muse about the *"unreasonable effectiveness of mathematics in the natural sciences"* [10].

Zeno of Elea and Parmenides, for instance, wondered how there can be motion if our universe is either infinitely divisible or discrete. Because, in the dense case (between any two points there is another point), the slightest finite move would require an infinity of actions. Likewise in the discrete case, how can there be motion if everything is not moving at all times [11]?

[7] Erwin Schrödinger. *Nature and the Greeks*. Cambridge University Press, Cambridge, 1954, 2014. ISBN 9781107431836. URL http://www.cambridge.org/9781107431836

[8] Hermann Diels. *Die Fragmente der Vorsokratiker, griechisch und deutsch*. Weidmannsche Buchhandlung, Berlin, 1906. URL http://www.archive.org/details/diefragmentederv01dieluoft German: Nachdem D. [[Demokritos]] sein Mißtrauen gegen die Sinneswahrnehmungen in dem Satze ausgesprochen: 'Scheinbar (d. i. konventionell) ist Farbe, scheinbar Süßigkeit, scheinbar Bitterkeit: wirklich nur Atome und Leeres" läßt er die Sinne gegen den Verstand reden: 'Du armer Verstand, von uns nimmst du deine Beweisstücke und willst uns damit besiegen? Dein Sieg ist dein Fall!'

[9] Imre Lakatos. *Philosophical Papers. 1. The Methodology of Scientific Research Programmes*. Cambridge University Press, Cambridge, 1978. ISBN 9781316038765, 9780521280310, 0521280311

[10] Eugene P. Wigner. The unreasonable effectiveness of mathematics in the natural sciences. Richard Courant Lecture delivered at New York University, May 11, 1959. *Communications on Pure and Applied Mathematics*, 13:1–14, 1960. DOI: 10.1002/cpa.3160130102. URL http://doi.org/10.1002/cpa.3160130102

[11] H. D. P. Lee. *Zeno of Elea*. Cambridge University Press, Cambridge, 1936; Paul Benacerraf. Tasks and supertasks, and the modern Eleatics. *Journal of Philosophy*, LIX(24):765–784, 1962. URL http://www.jstor.org/stable/2023500; A. Grünbaum. *Modern Science and Zeno's paradoxes*. Allen and Unwin, London, second edition, 1968; and Richard Mark

The Pythagoreans are often cited to have believed that the universe is natural numbers or simple fractions thereof, and thus physics is just a part of mathematics; or that there is no difference between these realms. They took their conception of numbers and world-as-numbers so seriously that the existence of irrational numbers which cannot be written as some ratio of integers shocked them; so much so that they allegedly drowned the poor guy who had discovered this fact. That appears to be a saddening case of a state of mind in which a subjective metaphysical belief in and wishful thinking about one's own constructions of the world overwhelms critical thinking; and what should be wisely taken as an epistemic finding is taken to be ontologic truth.

The relationship between physics and formalism has been debated by Bridgman [12], Feynman [13], and Landauer [14], among many others. It has many twists, anecdotes and opinions. Take, for instance, Heaviside's not uncontroversial stance [15] on it:

> *I suppose all workers in mathematical physics have noticed how the mathematics seems made for the physics, the latter suggesting the former, and that practical ways of working arise naturally. … But then the rigorous logic of the matter is not plain! Well, what of that? Shall I refuse my dinner because I do not fully understand the process of digestion? No, not if I am satisfied with the result. Now a physicist may in like manner employ unrigorous processes with satisfaction and usefulness if he, by the application of tests, satisfies himself of the accuracy of his results. At the same time he may be fully aware of his want of infallibility, and that his investigations are largely of an experimental character, and may be repellent to unsympathetically constituted mathematicians accustomed to a different kind of work. [§225]*

Dietrich Küchemann, the ingenious German-British aerodynamicist and one of the main contributors to the wing design of the *Concord* supersonic civil aercraft, tells us [16]

> *[Again,] the most drastic simplifying assumptions must be made before we can even think about the flow of gases and arrive at equations which are amenable to treatment. Our whole science lives on highly-idealised concepts and ingenious abstractions and approximations. We should remember this in all modesty at all times, especially when somebody claims to have obtained "the right answer" or "the exact solution". At the same time, we must acknowledge and admire the intuitive art of those scientists to whom we owe the many useful concepts and approximations with which we work [page 23].*

The question, for instance, is imminent whether we should take the formalism very seriously and literally, using it as a guide to new territories, which might even appear absurd, inconsistent and mind-boggling; just like Carroll's *Alice's Adventures in Wonderland.* Should we expect that all the wild things formally imaginable have a physical realization?

It might be prudent to adopt a contemplative strategy of *evenly-suspended attention* outlined by Freud [17], who admonishes analysts to be aware of the dangers caused by *"temptations to project, what [the analyst] in dull self-perception recognizes as the peculiarities of his own personality, as generally valid theory into science."* Nature is thereby treated as a client-patient, and whatever findings come up are accepted as is without any immediate emphasis or judgment. This also alleviates the dangers

[12] Percy W. Bridgman. A physicist's second reaction to Mengenlehre. *Scripta Mathematica*, 2:101–117, 224–234, 1934

[13] Richard Phillips Feynman. *The Feynman lectures on computation.* Addison-Wesley Publishing Company, Reading, MA, 1996. edited by A.J.G. Hey and R. W. Allen

[14] Rolf Landauer. Information is physical. *Physics Today*, 44(5):23–29, May 1991. DOI: 10.1063/1.881299. URL http://doi.org/10.1063/1.881299

[15] Oliver Heaviside. *Electromagnetic theory.* "The Electrician" Printing and Publishing Corporation, London, 1894-1912. URL http://archive.org/details/electromagnetict02heavrich

[16] Dietrich Küchemann. *The Aerodynamic Design of Aircraft.* Pergamon Press, Oxford, 1978

[17] Sigmund Freud. Ratschläge für den Arzt bei der psychoanalytischen Behandlung. In Anna Freud, E. Bibring, W. Hoffer, E. Kris, and O. Isakower, editors, *Gesammelte Werke. Chronologisch geordnet. Achter Band. Werke aus den Jahren 1909–1913*, pages 376–387. Fischer, Frankfurt am Main, 1912, 1999. URL http://gutenberg.spiegel.de/buch/kleine-schriften-ii-7122/15

of becoming embittered with the reactions of "the peers," a problem sometimes encountered when "surfing on the edge" of contemporary knowledge; such as, for example, Everett's case [18].

Jaynes has warned of the *"Mind Projection Fallacy"* [19], pointing out that *"we are all under an ego-driven temptation to project our private thoughts out onto the real world, by supposing that the creations of one's own imagination are real properties of Nature, or that one's own ignorance signifies some kind of indecision on the part of Nature."*

And yet, despite all aforementioned *provisos*, science finally succeeded to do what the alchemists sought for so long: we are capable of producing gold from mercury [20].

[18] Hugh Everett III. The Everett interpretation of quantum mechanics: Collected works 1955-1980 with commentary. Princeton University Press, Princeton, NJ, 2012. ISBN 9780691145075. URL http://press.princeton.edu/titles/9770.html

[19] Edwin Thompson Jaynes. Clearing up mysteries - the original goal. In John Skilling, editor, *Maximum-Entropy and Bayesian Methods: Proceedings of the 8th Maximum Entropy Workshop, held on August 1-5, 1988, in St. John's College, Cambridge, England*, pages 1–28. Kluwer, Dordrecht, 1989. URL http://bayes.wustl.edu/etj/articles/cmystery.pdf; and Edwin Thompson Jaynes. Probability in quantum theory. In Wojciech Hubert Zurek, editor, *Complexity, Entropy, and the Physics of Information: Proceedings of the 1988 Workshop on Complexity, Entropy, and the Physics of Information, held May - June, 1989, in Santa Fe, New Mexico*, pages 381–404. Addison-Wesley, Reading, MA, 1990. ISBN 9780201515091. URL http://bayes.wustl.edu/etj/articles/prob.in.qm.pdf

[20] R. Sherr, K. T. Bainbridge, and H. H. Anderson. Transmutation of mercury by fast neutrons. *Physical Review*, 60(7):473–479, Oct 1941. DOI: 10.1103/PhysRev.60.473. URL http://doi.org/10.1103/PhysRev.60.473

Part I
Linear vector spaces

1

Finite-dimensional vector spaces and linear algebra

VECTOR SPACES are prevalent in physics; they are essential for an understanding of mechanics, relativity theory, quantum mechanics, and statistical physics.

1.1 Conventions and basic definitions

This presentation is greatly inspired by Halmos' compact yet comprehensive treatment "Finite-Dimensional Vector Spaces" [1]. I greatly encourage the reader to have a look into that book. Of course, there exist zillions of other very nice presentations, among them Greub's "Linear algebra," and Strang's "Introduction to Linear Algebra," among many others, even freely downloadable ones [2] competing for your attention.

Unless stated differently, only finite-dimensional vector spaces will be considered.

In what follows the overline sign stands for complex conjugation; that is, if $a = \Re a + i \Im a$ is a complex number, then $\overline{a} = \Re a - i \Im a$.

A supercript "T" means transposition.

The physically oriented notation in Mermin's book on quantum information theory [3] is adopted. Vectors are either typed in bold face, or in Dirac's "bra-ket" notation[4]. Both notations will be used simultaneously and equivalently; not to confuse or obfuscate, but to make the reader familiar with the bra-ket notation used in quantum physics.

Thereby, the vector $\mathbf{x}$ is identified with the "ket vector" $|\mathbf{x}\rangle$. Ket vectors will be represented by column vectors, that is, by vertically arranged tuples of scalars, or, equivalently, as $n \times 1$ matrices; that is,

$$\mathbf{x} \equiv |\mathbf{x}\rangle \equiv \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{1.1}$$

The vector $\mathbf{x}^*$ from the dual space (see later, Section 1.8 on page 15) is identified with the "bra vector" $\langle \mathbf{x}|$. Bra vectors will be represented

by row vectors, that is, by horizontally arranged tuples of scalars, or, equivalently, as $1 \times n$ matrices; that is,

$$\mathbf{x}^* \equiv \langle \mathbf{x} | \equiv \left(x_1, x_2, \ldots, x_n \right). \tag{1.2}$$

Dot (scalar or inner) products between two vectors $\mathbf{x}$ and $\mathbf{y}$ in Euclidean space are then denoted by "$\langle$bra$|$(c)$|$ket$\rangle$" form; that is, by $\langle \mathbf{x} | \mathbf{y} \rangle$.

For an $n \times m$ matrix $\mathbf{A}$ we shall use the following *index notation*: suppose the (column) index i indicating their column number in a matrix-like object $\mathbf{A}$ "runs *horizontally*," and the (row) index j indicating their row number in a matrix-like object $\mathbf{A}$ "runs *vertically*," so that, with $1 \le i \le n$ and $1 \le j \le m$,

$$\mathbf{A} \equiv \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \equiv a_{ij}. \tag{1.3}$$

Stated differently, a_{ij} is the element of the table representing $\mathbf{A}$ which is in the ith row and in the jth column.

A *matrix multiplication* (written with or without dot) $\mathbf{A} \cdot \mathbf{B} = \mathbf{AB}$ of an $n \times m$ matrix $\mathbf{A} \equiv a_{ij}$ with an $m \times l$ matrix $\mathbf{B} \equiv b_{pq}$ can then be written as an $n \times l$ matrix $\mathbf{A} \cdot \mathbf{B} \equiv a_{ij} b_{jk}$, $1 \le i \le n$, $1 \le j \le m$, $1 \le k \le l$. Here the *Einstein summation convention* $a_{ij} b_{jk} = \sum_j a_{ij} b_{jk}$ has been used, which requires that, when an index variable appears twice in a single term, one has to sum over all of the possible index values. Stated differently, if $\mathbf{A}$ is an $n \times m$ matrix and $\mathbf{B}$ is an $m \times l$ matrix, their matrix product $\mathbf{AB}$ is an $n \times l$ matrix, in which the m entries across the rows of $\mathbf{A}$ are multiplied with the m entries down the columns of $\mathbf{B}$.

As stated earlier ket and bra vectors (from the original or the dual vector space; exact definitions will be given later) will be coded – with respect to a basis or coordinate system (see below) – as an n-tuple of numbers; which are arranged either a $n \times 1$ matrices (column vectors), or a $1 \times n$ matrices (row vectors), respectively. We can then write certain terms very compactly (alas often misleadingly). Suppose, for instance, that $|\mathbf{x}\rangle \equiv \mathbf{x} \equiv \left(x_1, x_2, \ldots, x_n \right)^T$ and $|\mathbf{y}\rangle \equiv \mathbf{y} \equiv \left(y_1, y_2, \ldots, y_n \right)^T$ are two (column) vectors (with respect to a given basis). Then, $x_i y_j a_{ij}$ can (somewhat superficially) be represented as a matrix multiplication $\mathbf{x}^T \mathbf{A} \mathbf{y}$ of a row vector with a matrix and a column vector yielding a scalar; which in turn can be interpreted as a 1×1 matrix. Note that, as "T" indicates transposition $\mathbf{y}^T \equiv \left[\left(y_1, y_2, \ldots, y_n \right)^T \right]^T = \left(y_1, y_2, \ldots, y_n \right)$ represents a row vector, whose components or coordinates with respect to a particular (here undisclosed) basis are the scalars y_i.

Note that double transposition yields the identity.

1.1.1   Fields of real and complex numbers

In physics, scalars occur either as real or complex numbers. Thus we shall restrict our attention to these cases.

A *field* $\langle \mathbb{F}, +, \cdot, -, ^{-1}, 0, 1 \rangle$ is a set together with two operations, usually called *addition* and *multiplication*, denoted by "$+$" and "$\cdot$" (often "$a \cdot b$" is

identified with the expression "ab" without the center dot) respectively, such that the following conditions (or, stated differently, axioms) hold:

(i)   closure of $\mathbb{F}$ with respect to addition and multiplication: for all $a, b \in \mathbb{F}$, both $a + b$ and ab are in $\mathbb{F}$;

(ii)   associativity of addition and multiplication: for all a, b, and c in $\mathbb{F}$, the following equalities hold: $a + (b + c) = (a + b) + c$, and $a(bc) = (ab)c$;

(iii)   commutativity of addition and multiplication: for all a and b in $\mathbb{F}$, the following equalities hold: $a + b = b + a$ and $ab = ba$;

(iv)   additive and multiplicative identity: there exists an element of $\mathbb{F}$, called the additive identity element and denoted by 0, such that for all a in $\mathbb{F}$, $a + 0 = a$. Likewise, there is an element, called the multiplicative identity element and denoted by 1, such that for all a in $\mathbb{F}$, $1 \cdot a = a$. (To exclude the trivial ring, the additive identity and the multiplicative identity are required to be distinct.)

(v)   additive and multiplicative inverses: for every a in $\mathbb{F}$, there exists an element $-a$ in $\mathbb{F}$, such that $a + (-a) = 0$. Similarly, for any a in $\mathbb{F}$ other than 0, there exists an element a^{-1} in $\mathbb{F}$, such that $a \cdot a^{-1} = 1$. (The elements $+(-a)$ and a^{-1} are also denoted $-a$ and $\frac{1}{a}$, respectively.) Stated differently: subtraction and division operations exist.

(vi)   Distributivity of multiplication over addition: For all a, b and c in $\mathbb{F}$, the following equality holds: $a(b + c) = (ab) + (ac)$.

1.1.2   Vectors and vector space

Vector spaces are merely structures allowing the sum (addition) of objects called "vectors," and multiplication of these objects by scalars; thereby remaining in this structure. That is, for instance, the "coherent superposition" $\mathbf{a} + \mathbf{b} \equiv |\mathbf{a} + \mathbf{b}\rangle$ of two vectors $\mathbf{a} \equiv |\mathbf{a}\rangle$ and $\mathbf{b} \equiv |\mathbf{b}\rangle$ can be guaranteed to be a vector. At this stage, little can be said about the length or relative direction or orientation of these "vectors." Algebraically, "vectors" are elements of vector spaces. Geometrically a vector may be interpreted as "a quantity which is usefully represented by an arrow" [5].

A *linear vector space* $\langle \mathfrak{V}, +, \cdot, -, 0, 1 \rangle$ is a set $\mathfrak{V}$ of elements called *vectors*, here denoted by bold face symbols such as $\mathbf{a}, \mathbf{x}, \mathbf{v}, \mathbf{w}, \ldots$, or, equivalently, denoted by $|\mathbf{a}\rangle, |\mathbf{x}\rangle, |\mathbf{v}\rangle, |\mathbf{w}\rangle, \ldots$, satisfying certain conditions (or, stated differently, axioms); among them, with respect to addition of vectors:

(i)   commutativity, that is, $|\mathbf{x}\rangle + |\mathbf{y}\rangle = |\mathbf{y}\rangle + |\mathbf{x}\rangle$;

(ii)   associativity, that is, $(|\mathbf{x}\rangle + |\mathbf{y}\rangle) + |\mathbf{z}\rangle = |\mathbf{x}\rangle + (|\mathbf{y}\rangle + |\mathbf{z}\rangle)$;

(iii)   the uniqueness of the origin or null vector 0; as well as

(iv)   the uniqueness of the negative vector;

with respect to multiplication of vectors with scalars:

(v)   the existence of a unit factor 1; and

For proofs and additional information see §2 in

Paul Richard Halmos. *Finite-dimensional Vector Spaces*. Undergraduate Texts in Mathematics. Springer, New York, 1958. ISBN 978-1-4612-6387-6,978-0-387-90093-3. DOI: 10.1007/978-1-4612-6387-6. URL http://doi.org/10.1007/978-1-4612-6387-6

[5] Gabriel Weinreich. *Geometrical Vectors (Chicago Lectures in Physics)*. The University of Chicago Press, Chicago, IL, 1998
In order to define *length*, we have to engage an additional structure, namely the *norm* $\|\mathbf{a}\|$ of a vector $\mathbf{a}$. And in order to define relative *direction* and *orientation*, and, in particular, *orthogonality* and *collinearity* we have to define the *scalar product* $\langle \mathbf{a} | \mathbf{b} \rangle$ of two vectors $\mathbf{a}$ and $\mathbf{b}$.

(vi) distributivity with respect to scalar and vector additions; that is,

$$(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x},$$
$$\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y},$$

(1.4)

with $\mathbf{x}, \mathbf{y} \in \mathfrak{V}$ and scalars $\alpha, \beta \in \mathbb{F}$, respectively.

Examples of vector spaces are:

(i)   The set $\mathbb{C}$ of complex numbers: $\mathbb{C}$ can be interpreted as a complex vector space by interpreting as vector addition and scalar multiplication as the usual addition and multiplication of complex numbers, and with 0 as the null vector;

(ii) The set $\mathbb{C}^n$, $n \in \mathbb{N}$ of n-tuples of complex numbers: Let $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$. $\mathbb{C}^n$ can be interpreted as a complex vector space by interpreting the ordinary addition $\mathbf{x} + \mathbf{y} = (x_1 + y_1, \ldots, x_n + y_n)$ and the multiplication $\alpha\mathbf{x} = (\alpha x_1, \ldots, \alpha x_n)$ by a complex number α as vector addition and scalar multiplication, respectively; the null tuple $0 = (0, \ldots, 0)$ is the neutral element of vector addition;

(iii) The set $\mathfrak{P}$ of all polynomials with complex coefficients in a variable t: $\mathfrak{P}$ can be interpreted as a complex vector space by interpreting the ordinary addition of polynomials and the multiplication of a polynomial by a complex number as vector addition and scalar multiplication, respectively; the null polynomial is the neutral element of vector addition.

1.2   Linear independence

A set $\mathfrak{S} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\} \subset \mathfrak{V}$ of vectors $\mathbf{x}_i$ in a linear vector space is *linearly independent* if $\mathbf{x}_i \neq 0 \,\forall 1 \leq i \leq k$, and additionally, if either $k = 1$, or if no vector in $\mathfrak{S}$ can be written as a linear combination of other vectors in this set $\mathfrak{S}$; that is, there are no scalars α_j satisfying $\mathbf{x}_i = \sum_{1 \leq j \leq k, \, j \neq i} \alpha_j \mathbf{x}_j$.

Equivalently, if $\sum_{i=1}^{k} \alpha_i \mathbf{x}_i = 0$ implies $\alpha_i = 0$ for each i, then the set $\mathfrak{S} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k\}$ is linearly independent.

Note that a the vectors of a basis are linear independent and "maximal" insofar as any inclusion of an additional vector results in a linearly dependent set; that ist, this additional vector can be expressed in terms of a linear combination of the existing basis vectors; see also Section 1.4 on page 9.

1.3   Subspace

A nonempty subset $\mathfrak{M}$ of a vector space is a *subspace* or, used synonymously, a *linear manifold*, if, along with every pair of vectors $\mathbf{x}$ and $\mathbf{y}$ contained in $\mathfrak{M}$, every linear combination $\alpha\mathbf{x} + \beta\mathbf{y}$ is also contained in $\mathfrak{M}$.

If $\mathfrak{U}$ and $\mathfrak{V}$ are two subspaces of a vector space, then $\mathfrak{U} + \mathfrak{V}$ is the subspace spanned by $\mathfrak{U}$ and $\mathfrak{V}$; that is, it contains all vectors $\mathbf{z} = \mathbf{x} + \mathbf{y}$, with $\mathbf{x} \in \mathfrak{U}$ and $\mathbf{y} \in \mathfrak{V}$.

$\mathfrak{M}$ is the *linear span*

$$\mathfrak{M} = \text{span}(\mathfrak{U}, \mathfrak{V}) = \text{span}(\mathbf{x}, \mathbf{y}) = \{\alpha\mathbf{x} + \beta\mathbf{y} \mid \alpha, \beta \in \mathbb{F}, \mathbf{x} \in \mathfrak{U}, \mathbf{y} \in \mathfrak{V}\}. \quad (1.5)$$

A generalization to more than two vectors and more than two subspaces is straightforward.

For every vector space $\mathfrak{V}$, the vector space containing only the null vector, and the vector space $\mathfrak{V}$ itself are subspaces of $\mathfrak{V}$.

1.3.1   Scalar or inner product

A *scalar* or *inner* product presents some form of measure of "distance" or "apartness" of two vectors in a linear vector space. It should not be confused with the bilinear functionals (introduced on page 15) that connect a vector space with its dual vector space, although for real Euclidean vector spaces these may coincide, and although the scalar product is also bilinear in its arguments. It should also not be confused with the tensor product introduced in Section 1.10 on page 23.

An inner product space is a vector space $\mathfrak{V}$, together with an inner product; that is, with a map $\langle\cdot|\cdot\rangle : \mathfrak{V} \times \mathfrak{V} \longrightarrow \mathbb{F}$ (usually $\mathbb{F} = \mathbb{C}$ or $\mathbb{F} = \mathbb{R}$) that satisfies the following three conditions (or, stated differently, axioms) for all vectors and all scalars:

(i)   Conjugate (Hermitian) symmetry: $\langle\mathbf{x}|\mathbf{y}\rangle = \overline{\langle\mathbf{y}|\mathbf{x}\rangle}$;

(ii)   linearity in the first argument:

$$\langle\alpha\mathbf{x} + \beta\mathbf{y}|\mathbf{z}\rangle = \alpha\langle\mathbf{x}|\mathbf{z}\rangle + \beta\langle\mathbf{y}|\mathbf{z}\rangle;$$

(iii)   positive-definiteness: $\langle\mathbf{x}|\mathbf{x}\rangle \geq 0$; with equality if and only if $\mathbf{x} = 0$.

Note that from the first two properties, it follows that the inner product is *antilinear,* or synonymously, *conjugate-linear,* in its second argument (note that $\overline{(uv)} = (\overline{u})(\overline{v})$ for all $u, v \in \mathbb{C}$):

$$\langle\mathbf{z}|\alpha\mathbf{x} + \beta\mathbf{y}\rangle = \overline{\langle\alpha\mathbf{x} + \beta\mathbf{y}|\mathbf{z}\rangle} = \overline{\alpha}\,\overline{\langle\mathbf{x}|\mathbf{z}\rangle} + \overline{\beta}\,\overline{\langle\mathbf{y}|\mathbf{z}\rangle} = \overline{\alpha}\langle\mathbf{z}|\mathbf{x}\rangle + \overline{\beta}\langle\mathbf{z}|\mathbf{y}\rangle. \quad (1.6)$$

One example of an inner product is is the *dot product*

$$\langle\mathbf{x}|\mathbf{y}\rangle = \sum_{i=1}^{n} x_i \overline{y_i} \quad (1.7)$$

of two vectors $\mathbf{x} = (x_1, \ldots, x_n)$ and $\mathbf{y} = (y_1, \ldots, y_n)$ in $\mathbb{C}^n$, which, for real Euclidean space, reduces to the well-known dot product $\langle\mathbf{x}|\mathbf{y}\rangle = x_1 y_1 + \cdots + x_n y_n = \|\mathbf{x}\|\|\mathbf{y}\| \cos\angle(\mathbf{x}, \mathbf{y})$.

It is mentioned without proof that the most general form of an inner product in $\mathbb{C}^n$ is $\langle\mathbf{x}|\mathbf{y}\rangle = \mathbf{y}\mathbf{A}\mathbf{x}^{\dagger}$, where the symbol "$\dagger$" stands for the *conjugate transpose* (also denoted as *Hermitian conjugate* or *Hermitian adjoint*), and $\mathbf{A}$ is a positive definite Hermitian matrix (all of its eigenvalues are positive).

The *norm* of a vector $\mathbf{x}$ is defined by

$$\|\mathbf{x}\| = \sqrt{\langle\mathbf{x}|\mathbf{x}\rangle}. \quad (1.8)$$

For real, Euclidean vector spaces, this function is symmetric; that is $\langle\mathbf{x}|\mathbf{y}\rangle = \langle\mathbf{y}|\mathbf{x}\rangle$.

Conversely, the *polarization identity* expresses the inner product of two vectors in terms of the norm of their differences; that is,

$$\langle \mathbf{x}|\mathbf{y}\rangle = \frac{1}{4}\left[\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 + i\left(\|\mathbf{x}+i\mathbf{y}\|^2 - \|\mathbf{x}-i\mathbf{y}\|^2\right)\right]. \qquad (1.9)$$

In complex vector space, a direct but tedious calculation – with linearity in the first argument and conjugate-linearity in the second argument of the inner product – yields

$$\frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2 + i\|\mathbf{x}+i\mathbf{y}\|^2 - i\|\mathbf{x}-i\mathbf{y}\|^2\right)$$

$$= \frac{1}{4}\left(\langle\mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y}\rangle - \langle\mathbf{x}-\mathbf{y}|\mathbf{x}-\mathbf{y}\rangle + i\langle\mathbf{x}+i\mathbf{y}|\mathbf{x}+i\mathbf{y}\rangle - i\langle\mathbf{x}-i\mathbf{y}|\mathbf{x}-i\mathbf{y}\rangle\right)$$

$$= \frac{1}{4}\left[(\langle\mathbf{x}|\mathbf{x}\rangle + \langle\mathbf{x}|\mathbf{y}\rangle + \langle\mathbf{y}|\mathbf{x}\rangle + \langle\mathbf{y}|\mathbf{y}\rangle) - (\langle\mathbf{x}|\mathbf{x}\rangle - \langle\mathbf{x}|\mathbf{y}\rangle - \langle\mathbf{y}|\mathbf{x}\rangle + \langle\mathbf{y}|\mathbf{y}\rangle)\right.$$

$$\left.+ i(\langle\mathbf{x}|\mathbf{x}\rangle + \langle\mathbf{x}|i\mathbf{y}\rangle + \langle i\mathbf{y}|\mathbf{x}\rangle + \langle i\mathbf{y}|i\mathbf{y}\rangle) - i(\langle\mathbf{x}|\mathbf{x}\rangle - \langle\mathbf{x}|i\mathbf{y}\rangle - \langle i\mathbf{y}|\mathbf{x}\rangle + \langle i\mathbf{y}|i\mathbf{y}\rangle)\right]$$

$$= \frac{1}{4}\left[\langle\mathbf{x}|\mathbf{x}\rangle + \langle\mathbf{x}|\mathbf{y}\rangle + \langle\mathbf{y}|\mathbf{x}\rangle + \langle\mathbf{y}|\mathbf{y}\rangle - \langle\mathbf{x}|\mathbf{x}\rangle + \langle\mathbf{x}|\mathbf{y}\rangle + \langle\mathbf{y}|\mathbf{x}\rangle - \langle\mathbf{y}|\mathbf{y}\rangle\right.$$

$$\left.+ i(\langle\mathbf{x}|\mathbf{x}\rangle + \langle\mathbf{x}|i\mathbf{y}\rangle + \langle i\mathbf{y}|\mathbf{x}\rangle + \langle i\mathbf{y}|i\mathbf{y}\rangle - \langle\mathbf{x}|\mathbf{x}\rangle + \langle\mathbf{x}|i\mathbf{y}\rangle + \langle i\mathbf{y}|\mathbf{x}\rangle - \langle i\mathbf{y}|i\mathbf{y}\rangle)\right]$$

$$= \frac{1}{4}\left[2(\langle\mathbf{x}|\mathbf{y}\rangle + \langle\mathbf{y}|\mathbf{x}\rangle) + 2i(\langle\mathbf{x}|i\mathbf{y}\rangle + \langle i\mathbf{y}|\mathbf{x}\rangle)\right]$$

$$= \frac{1}{2}\left[(\langle\mathbf{x}|\mathbf{y}\rangle + \langle\mathbf{y}|\mathbf{x}\rangle) + i(-i\langle\mathbf{x}|\mathbf{y}\rangle + i\langle\mathbf{y}|\mathbf{x}\rangle)\right]$$

$$= \frac{1}{2}\left[(\langle\mathbf{x}|\mathbf{y}\rangle + \langle\mathbf{y}|\mathbf{x}\rangle) + (\langle\mathbf{x}|\mathbf{y}\rangle - \langle\mathbf{y}|\mathbf{x}\rangle)\right] = \langle\mathbf{x}|\mathbf{y}\rangle. \qquad (1.10)$$

For any real vector space the imaginary terms in (1.9) are absent, and (1.9) reduces to

$$\langle\mathbf{x}|\mathbf{y}\rangle = \frac{1}{4}\left(\langle\mathbf{x}+\mathbf{y}|\mathbf{x}+\mathbf{y}\rangle - \langle\mathbf{x}-\mathbf{y}|\mathbf{x}-\mathbf{y}\rangle\right) = \frac{1}{4}\left(\|\mathbf{x}+\mathbf{y}\|^2 - \|\mathbf{x}-\mathbf{y}\|^2\right). \qquad (1.11)$$

Two nonzero vectors $\mathbf{x},\mathbf{y}\in\mathfrak{V}$, $\mathbf{x},\mathbf{y}\neq 0$ are *orthogonal*, denoted by "$\mathbf{x}\perp\mathbf{y}$" if their scalar product vanishes; that is, if

$$\langle\mathbf{x}|\mathbf{y}\rangle = 0. \qquad (1.12)$$

Let $\mathfrak{E}$ be any set of vectors in an inner product space $\mathfrak{V}$. The symbol

$$\mathfrak{E}^{\perp} = \left\{\mathbf{x}\mid\langle\mathbf{x}|\mathbf{y}\rangle = 0, \mathbf{x}\in\mathfrak{V}, \forall\mathbf{y}\in\mathfrak{E}\right\} \qquad (1.13)$$

denotes the set of all vectors in $\mathfrak{V}$ that are orthogonal to every vector in $\mathfrak{E}$.

Note that, regardless of whether or not $\mathfrak{E}$ is a subspace, $\mathfrak{E}^{\perp}$ is a subspace. Furthermore, $\mathfrak{E}$ is contained in $(\mathfrak{E}^{\perp})^{\perp} = \mathfrak{E}^{\perp\perp}$. In case $\mathfrak{E}$ is a subspace, we call $\mathfrak{E}^{\perp}$ the *orthogonal complement* of $\mathfrak{E}$.

The following *projection theorem* is mentioned without proof. If $\mathfrak{M}$ is any subspace of a finite-dimensional inner product space $\mathfrak{V}$, then $\mathfrak{V}$ is the direct sum of $\mathfrak{M}$ and $\mathfrak{M}^{\perp}$; that is, $\mathfrak{M}^{\perp\perp} = \mathfrak{M}$.

For the sake of an example, suppose $\mathfrak{V} = \mathbb{R}^2$, and take $\mathfrak{E}$ to be the set of all vectors spanned by the vector $(1,0)$; then $\mathfrak{E}^{\perp}$ is the set of all vectors spanned by $(0,1)$.

1.3.2   Hilbert space

A (quantum mechanical) *Hilbert space* is a linear vector space $\mathfrak{V}$ over the field $\mathbb{C}$ of complex numbers (sometimes only $\mathbb{R}$ is used) equipped with vector addition, scalar multiplication, and some inner (scalar) product. Furthermore, *closure* is an additional requirement, but nobody has made operational sense of that so far: If $\mathbf{x}_n \in \mathfrak{V}$, $n = 1, 2, \ldots$, and if $\lim_{n,m \to \infty} (\mathbf{x}_n - \mathbf{x}_m, \mathbf{x}_n - \mathbf{x}_m) = 0$, then there exists an $\mathbf{x} \in \mathfrak{V}$ with $\lim_{n \to \infty} (\mathbf{x}_n - \mathbf{x}, \mathbf{x}_n - \mathbf{x}) = 0$.

Infinite dimensional vector spaces and continuous spectra are nontrivial extensions of the finite dimensional Hilbert space treatment. As a heuristic rule – which is not always correct – it might be stated that the sums become integrals, and the Kronecker delta function δ_{ij} defined by

$$\delta_{ij} = \begin{cases} 0 & \text{for } i \neq j, \\ 1 & \text{for } i = j. \end{cases} \tag{1.14}$$

becomes the Dirac delta function $\delta(x - y)$, which is a generalized function in the continuous variables x, y. In the Dirac bra-ket notation, the resolution of the identity operator, sometimes also referred to as *completeness*, is given by $\mathbb{I} = \int_{-\infty}^{+\infty} |x\rangle\langle x| \, dx$. For a careful treatment, see, for instance, the books by Reed and Simon [6], or wait for Chapter 7, page 139.

1.4   Basis

We shall use bases of vector spaces to formally represent vectors (elements) therein.

A (linear) *basis* [or a *coordinate system*, or a *frame (of reference)*] is a set $\mathfrak{B}$ of linearly independent vectors such that every vector in $\mathfrak{V}$ is a linear combination of the vectors in the basis; hence $\mathfrak{B}$ spans $\mathfrak{V}$.

What particular basis should one choose? *A priori* no basis is privileged over the other. Yet, in view of certain (mutual) properties of elements of some bases (such as orthogonality or orthonormality) we shall prefer (s)ome over others.

Note that a vector is some directed entity with a particular length, oriented in some (vector) "space." It is "laid out there" in front of our eyes, as it is: some directed entity. *A priori*, this space, in its most primitive form, is not equipped with a basis, or synonymuously, frame of reference, or reference frame. Insofar it is not yet coordinatized. In order to formalize the notion of a vector, we have to code this vector by "coordinates" or "components" which are the coeffitients with respect to a (de)composition into basis elements. Therefore, just as for numbers (e.g., by different numeral bases, or by prime decomposition), there exist many "competing" ways to code a vector.

Some of these ways appear to be rather straightforward, such as, in particular, the *Cartesian basis,* also synonymuosly called the *standard basis.* It is, however, not in any way *a priori* "evident" or "necessary" what should be specified to be "the Cartesian basis." Actually, specification of a "Cartesian basis" seems to be mainly motivated by physical inertial motion – and thus identified with some inertial frame of ref-

[6] Michael Reed and Barry Simon. *Methods of Mathematical Physics I: Functional Analysis.* Academic Press, New York, 1972; and Michael Reed and Barry Simon. *Methods of Mathematical Physics II: Fourier Analysis, Self-Adjointness.* Academic Press, New York, 1975

For proofs and additional information see §7 in

Paul Richard Halmos. *Finite-dimensional Vector Spaces.* Undergraduate Texts in Mathematics. Springer, New York, 1958. ISBN 978-1-4612-6387-6,978-0-387-90093-3. DOI: 10.1007/978-1-4612-6387-6. URL http://doi.org/10.1007/978-1-4612-6387-6

erence – "without any friction and forces," resulting in a "straight line motion at constant speed." (This sentence is cyclic, because heuristically any such absence of "friction and force" can only be operationalized by testing if the motion is a "straight line motion at constant speed.") If we grant that in this way straight lines can be defined, then Cartesian bases in Euclidean vector spaces can be characterized by orthogonal (orthogonality is defined *via* vanishing scalar products between nonzero vectors) straight lines spanning the entire space. In this way, we arrive, say for a planar situation, at the coordinates characteried by some basis $\{(0,1),(1,0)\}$, where, for instance, the basis vector "$(1,0)$" literally and physically means "a unit arrow pointing in some particular, specified direction."

Alas, if we would prefer, say, cyclic motion in the plane, we might want to call a frame based on the polar coordinates r and θ "Cartesian," resulting in some "Cartesian basis" $\{(0,1),(1,0)\}$; but this "Cartesian basis" would be very different from the Cartesian basis mentioned earlier, as "$(1,0)$" would refer to some specific unit radius, and "$(0,1)$" would refer to some specific unit angle (with respect to a specific zero angle). In terms of the "straight" coordinates (with respect to "the usual Cartesian basis") x, y, the polar coordinates are $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}(y/x)$. We obtain the original "straight" coordinates (with respect to "the usual Cartesian basis") back if we take $x = r\cos\theta$ and $y = r\sin\theta$.

Other bases than the "Cartesian" one may be less suggestive at first; alas it may be "economical" or pragmatical to use them; mostly to cope with, and adapt to, the *symmetry* of a physical configuration: if the physical situation at hand is, for instance, rotationally invariant, we might want to use rotationally invariant bases – such as, for instance, polar coordinares in two dimensions, or spherical coordinates in three dimensions – to represent a vector, or, more generally, to code any given representation of a physical entity (e.g., tensors, operators) by such bases.

1.5 Dimension

The *dimension* of $\mathfrak{V}$ is the number of elements in $\mathfrak{B}$.

All bases $\mathfrak{B}$ of $\mathfrak{V}$ contain the same number of elements.

A vector space is finite dimensional if its bases are finite; that is, its bases contain a finite number of elements.

In quantum physics, the dimension of a quantized system is associated with the *number of mutually exclusive measurement outcomes*. For a spin state measurement of an electron along a particular direction, as well as for a measurement of the linear polarization of a photon in a particular direction, the dimension is two, since both measurements may yield two distinct outcomes which we can interpret as vectors in two-dimensional Hilbert space, which, in Dirac's bra-ket notation [7], can be written as $|\uparrow\rangle$ and $|\downarrow\rangle$, or $|+\rangle$ and $|-\rangle$, or $|H\rangle$ and $|V\rangle$, or $|0\rangle$ and $|1\rangle$, or $|\text{☺}\rangle$ and $|\text{☹}\rangle$, respectively.

[7] Paul Adrien Maurice Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, Oxford, fourth edition, 1930, 1958. ISBN 9780198520115

1.6   Vector coordinates or components

The coordinates or components of a vector with respect to some basis represent the coding of that vector in that particular basis. It is important to realize that, as bases change, so do coordinates. Indeed, the changes in coordinates have to "compensate" for the bases change, because the same coordinates in a different basis would render an altogether different vector. Thus it is often said that, in order to represent one and the same vector, if the base vectors *vary*, the corresponding components or coordinates have to *contra-vary*. Figure 1.1 presents some geometrical demonstration of these thoughts, for your contemplation.
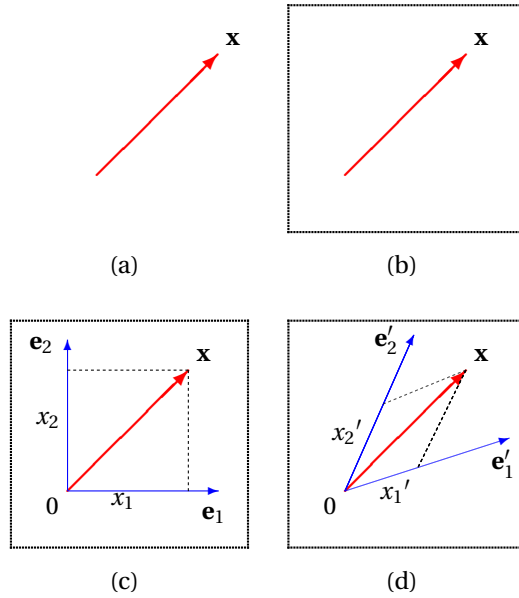
(a)   (b)

(c)   (d)

Figure 1.1: Coordinazation of vectors: (a) some primitive vector; (b) some primitive vectors, laid out in some space, denoted by dotted lines (c) vector coordinates x_1 and x_2 of the vector $\mathbf{x} = (x_1, x_2) = x_1 \mathbf{e}_1 + x_2 \mathbf{e}_2$ in a standard basis; (d) vector coordinates x_1' and x_2' of the vector $\mathbf{x} = (x_1', x_2') = x_1' \mathbf{e}_1' + x_2' \mathbf{e}_2'$ in some nonorthogonal basis.

   Elementary high school tutorials often condition students into believing that the components of the vector "is" the vector, rather then emphasizing that these components *represent* or *encode* the vector with respect to some (mostly implicitly assumed) basis. A similar situation occurs in many introductions to quantum theory, where the span (i.e., the onedimensional linear subspace spanned by that vector) $\{\mathbf{y} \mid \mathbf{y} = \alpha\mathbf{x}, \alpha \in \mathbb{C}\}$, or, equivalently, for orthogonal projections, the *projection* (i.e., the projection operator; see also page 27) $\mathbf{E_x} \equiv \mathbf{x} \otimes \mathbf{x}^T \equiv |\mathbf{x}\rangle\langle\mathbf{x}|$ corresponding to a unit (of length 1) vector $\mathbf{x}$ often is identified with that vector. In many instances, this is a great help and, if administered properly, is consistent and fine (at least for all practical purposes).

   The Cartesian standard basis in n-dimensional complex space $\mathbb{C}^n$ is the set of (usually "straight") vectors $x_i, i = 1, \ldots, n$, of "unit length" – the unit is conventional and thus needs to be fixed as operationally precisely as possible, such as in the *International System of Units (SI)* – represented by n-tuples, defined by the condition that the i'th coordinate of the j'th basis vector $\mathbf{e}_j$ is given by δ_{ij}, where δ_{ij} is the Kronecker delta function

$$\delta_{ij} = \begin{cases} 0 & \text{for } i \neq j, \\ 1 & \text{for } i = j. \end{cases} \tag{1.15}$$

Thus we can represent the basis vectors by

$$|\mathbf{e}_1\rangle \equiv \mathbf{e}_1 \equiv \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad |\mathbf{e}_2\rangle \equiv \mathbf{e}_2 \equiv \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad \dots \quad |\mathbf{e}_n\rangle \equiv \mathbf{e}_n \equiv \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \qquad (1.16)$$

In terms of these standard base vectors, every vector $\mathbf{x}$ can be written as a linear combination – in quantum physics, this is called *coherent superposition*

$$|\mathbf{x}\rangle \equiv \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{e}_i \equiv \sum_{i=1}^{n} x_i |\mathbf{e}_i\rangle \equiv \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \qquad (1.17)$$

With the notation defined by

$$X = \left(x_1, x_2, \dots, x_n \right)^T, \text{ and}$$
$$\mathbf{U} = \left(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \right) \equiv \left(|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle, \dots, |\mathbf{e}_n\rangle \right), \qquad (1.18)$$

such that $u_{ij} = e_{i,j}$ is the jth component of the ith vector, Eq. (1.17) can be written in "Euclidean dot product notation," that is, "column times row" and "row times column" (the dot is usually omitted)

For reasons demonstrated later in Eq. (1.178) $\mathbf{U}$ is a unitary matrix, that is, $\mathbf{U}^{-1} = \mathbf{U}^\dagger = \overline{\mathbf{U}}^T$, where the overline stands for complex conjugation $\overline{u}_{ij}$ of the entries u_{ij} of $\mathbf{U}$, and the superscript T indicates transposition; that is, $\mathbf{U}^T$ has entries u_{ji}.

$$|\mathbf{x}\rangle \equiv \mathbf{x} = \left(\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n \right) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \equiv \left(|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle, \dots, |\mathbf{e}_n\rangle \right) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \equiv$$
$$\equiv \begin{pmatrix} \mathbf{e}_{1,1} & \mathbf{e}_{2,1} & \cdots & \mathbf{e}_{n,1} \\ \mathbf{e}_{1,2} & \mathbf{e}_{2,2} & \cdots & \mathbf{e}_{n,2} \\ \cdots & \cdots & \ddots & \cdots \\ \mathbf{e}_{1,n} & \mathbf{e}_{2,n} & \cdots & \mathbf{e}_{n,n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \equiv \mathbf{U} X. \qquad (1.19)$$

Of course, with the Cartesian standard basis (1.16), $\mathbf{U} = \mathbb{I}_n$, but (1.19) remains valid for general bases.

In (1.19) the identification of the tuple $X = \left(x_1, x_2, \dots, x_n \right)^T$ containing the vector components x_i with the vector $|\mathbf{x}\rangle \equiv \mathbf{x}$ really means "coded *with respect*, or *relative*, to the basis $\mathfrak{B} = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$." Thus in what follows, we shall often identify the column vector $\left(x_1, x_2, \dots, x_n \right)^T$ containing the coordinates of the vector with the vector $\mathbf{x} \equiv |\mathbf{x}\rangle$, but we always need to keep in mind that the tuples of coordinates are defined only with respect to a particular basis $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$; otherwise these numbers lack any meaning whatsoever.

Indeed, with respect to some arbitrary basis $\mathfrak{B} = \{\mathbf{f}_1, \dots, \mathbf{f}_n\}$ of some n-dimensional vector space $\mathfrak{V}$ with the base vectors $\mathbf{f}_i$, $1 \le i \le n$, every vector $\mathbf{x}$ in $\mathfrak{V}$ can be written as a unique linear combination

$$|\mathbf{x}\rangle \equiv \mathbf{x} = \sum_{i=1}^{n} x_i \mathbf{f}_i \equiv \sum_{i=1}^{n} x_i |\mathbf{f}_i\rangle \equiv \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \qquad (1.20)$$

of the product of the coordinates x_i with respect to the basis $\mathfrak{B}$.

The uniqueness of the coordinates is proven indirectly by *reductio ad absurdum:* Suppose there is another decomposition $\mathbf{x} = \sum_{i=1}^{n} y_i \mathbf{f}_i = (y_1, y_2, \ldots, y_n)$; then by subtraction, $0 = \sum_{i=1}^{n} (x_i - y_i)\mathbf{f}_i = (0, 0, \ldots, 0)$. Since the basis vectors $\mathbf{f}_i$ are linearly independent, this can only be valid if all coefficients in the summation vanish; thus $x_i - y_i = 0$ for all $1 \le i \le n$; hence finally $x_i = y_i$ for all $1 \le i \le n$. This is in contradiction with our assumption that the coordinates x_i and y_i (or at least some of them) are different. Hence the only consistent alternative is the assumption that, with respect to a given basis, the coordinates are uniquely determined.

A set $\mathfrak{B} = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ of vectors of the inner product space $\mathfrak{V}$ is *orthonormal* if, for all $\mathbf{a}_i \in \mathfrak{B}$ and $\mathbf{a}_j \in \mathfrak{B}$, it follows that

$$\langle \mathbf{a}_i \mid \mathbf{a}_j \rangle = \delta_{ij}. \tag{1.21}$$

Any such set is called *complete* if it is not a subset of any larger orthonormal set of vectors of $\mathfrak{V}$. Any complete set is a basis. If, instead of Eq. (1.21), $\langle \mathbf{a}_i \mid \mathbf{a}_j \rangle = \alpha_i \delta_{ij}$ with nontero factors α_i, the set is called *orthogonal.*

1.7   *Finding orthogonal bases from nonorthogonal ones*

A *Gram-Schmidt process* [8] is a systematic method for orthonormalising a set of vectors in a space equipped with a *scalar product,* or by a synonym preferred in mathematics, *inner product.* The Gram-Schmidt process takes a finite, linearly independent set of base vectors and generates an orthonormal basis that spans the same (sub)space as the original set.

The general method is to start out with the original basis, say, $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \ldots, \mathbf{x}_n\}$, and generate a new orthogonal basis $\{\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \ldots, \mathbf{y}_n\}$ by

$$\mathbf{y}_1 = \mathbf{x}_1,$$
$$\mathbf{y}_2 = \mathbf{x}_2 - P_{\mathbf{y}_1}(\mathbf{x}_2),$$
$$\mathbf{y}_3 = \mathbf{x}_3 - P_{\mathbf{y}_1}(\mathbf{x}_3) - P_{\mathbf{y}_2}(\mathbf{x}_3),$$
$$\vdots$$
$$\mathbf{y}_n = \mathbf{x}_n - \sum_{i=1}^{n-1} P_{\mathbf{y}_i}(\mathbf{x}_n), \tag{1.22}$$

where

$$P_{\mathbf{y}}(\mathbf{x}) = \frac{\langle \mathbf{x}|\mathbf{y} \rangle}{\langle \mathbf{y}|\mathbf{y} \rangle}\mathbf{y}, \text{ and } P_{\mathbf{y}}^{\perp}(\mathbf{x}) = \mathbf{x} - \frac{\langle \mathbf{x}|\mathbf{y} \rangle}{\langle \mathbf{y}|\mathbf{y} \rangle}\mathbf{y} \tag{1.23}$$

are the orthogonal projections of $\mathbf{x}$ onto $\mathbf{y}$ and $\mathbf{y}^{\perp}$, respectively (the latter is mentioned for the sake of completeness and is not required here). Note that these orthogonal projections are idempotent and mutually orthogonal; that is,

$$P_{\mathbf{y}}^2(\mathbf{x}) = P_{\mathbf{y}}(P_{\mathbf{y}}(\mathbf{x})) = \frac{\langle \mathbf{x}|\mathbf{y} \rangle}{\langle \mathbf{y}|\mathbf{y} \rangle}\frac{\langle \mathbf{y}|\mathbf{y} \rangle}{\langle \mathbf{y}|\mathbf{y} \rangle}\mathbf{y} = P_{\mathbf{y}}(\mathbf{x}),$$

$$(P_{\mathbf{y}}^{\perp})^2(\mathbf{x}) = P_{\mathbf{y}}^{\perp}(P_{\mathbf{y}}^{\perp}(\mathbf{x})) = \mathbf{x} - \frac{\langle \mathbf{x}|\mathbf{y} \rangle}{\langle \mathbf{y}|\mathbf{y} \rangle}\mathbf{y} - \left(\frac{\langle \mathbf{x}|\mathbf{y} \rangle}{\langle \mathbf{y}|\mathbf{y} \rangle} - \frac{\langle \mathbf{x}|\mathbf{y} \rangle \langle \mathbf{y}|\mathbf{y} \rangle}{\langle \mathbf{y}|\mathbf{y} \rangle^2} \right)\mathbf{y}, = P_{\mathbf{y}}^{\perp}(\mathbf{x}), \tag{1.24}$$

$$P_{\mathbf{y}}(P_{\mathbf{y}}^{\perp}(\mathbf{x})) = P_{\mathbf{y}}^{\perp}(P_{\mathbf{y}}(\mathbf{x})) = \frac{\langle \mathbf{x}|\mathbf{y} \rangle}{\langle \mathbf{y}|\mathbf{y} \rangle}\mathbf{y} - \frac{\langle \mathbf{x}|\mathbf{y} \rangle \langle \mathbf{y}|\mathbf{y} \rangle}{\langle \mathbf{y}|\mathbf{y} \rangle^2}\mathbf{y} = 0.$$

[8] Steven J. Leon, Åke Björck, and Walter Gander. Gram-Schmidt orthogonalization: 100 years and more. *Numerical Linear Algebra with Applications,* 20(3):492–532, 2013. ISSN 1070-5325. DOI: 10.1002/nla.1839. URL http://doi.org/10.1002/nla.1839
The scalar or inner product $\langle \mathbf{x}|\mathbf{y} \rangle$ of two vectors $\mathbf{x}$ and $\mathbf{y}$ is defined on page 7. In Euclidean space such as $\mathbb{R}^n$, one often identifies the "dot product" $\mathbf{x} \cdot \mathbf{y} = x_1 y_1 + \cdots + x_n y_n$ of two vectors $\mathbf{x}$ and $\mathbf{y}$ with their scalar or inner product.

For a more general discussion of projections, see also page 27.

Subsequently, in order to obtain an orthonormal basis, one can divide every basis vector by its length.

The idea of the proof is as follows (see also Greub [9], section 7.9). In order to generate an orthogonal basis from a nonorthogonal one, the first vector of the old basis is identified with the first vector of the new basis; that is $\mathbf{y}_1 = \mathbf{x}_1$. Then, as depicted in Fig. 1.2, the second vector of the new basis is obtained by taking the second vector of the old basis and subtracting its projection on the first vector of the new basis.

More precisely, take the Ansatz

$$\mathbf{y}_2 = \mathbf{x}_2 + \lambda \mathbf{y}_1, \tag{1.25}$$

thereby determining the arbitrary scalar λ such that $\mathbf{y}_1$ and $\mathbf{y}_2$ are orthogonal; that is, $\langle \mathbf{y}_1 | \mathbf{y}_2 \rangle = 0$. This yields

$$\langle \mathbf{y}_2 | \mathbf{y}_1 \rangle = \langle \mathbf{x}_2 | \mathbf{y}_1 \rangle + \lambda \langle \mathbf{y}_1 | \mathbf{y}_1 \rangle = 0, \tag{1.26}$$

and thus, since $\mathbf{y}_1 \neq 0$,

$$\lambda = -\frac{\langle \mathbf{x}_2 | \mathbf{y}_1 \rangle}{\langle \mathbf{y}_1 | \mathbf{y}_1 \rangle}. \tag{1.27}$$

To obtain the third vector $\mathbf{y}_3$ of the new basis, take the Ansatz

$$\mathbf{y}_3 = \mathbf{x}_3 + \mu \mathbf{y}_1 + \nu \mathbf{y}_2, \tag{1.28}$$

and require that it is orthogonal to the two previous orthogonal basis vectors $\mathbf{y}_1$ and $\mathbf{y}_2$; that is $\langle \mathbf{y}_1 | \mathbf{y}_3 \rangle = \langle \mathbf{y}_2 | \mathbf{y}_3 \rangle = 0$. We already know that $\langle \mathbf{y}_1 | \mathbf{y}_2 \rangle = 0$. Consider the scalar products of $\mathbf{y}_1$ and $\mathbf{y}_2$ with the *Ansatz* for $\mathbf{y}_3$ in Eq. (1.28); that is,

$$\begin{aligned} \langle \mathbf{y}_3 | \mathbf{y}_1 \rangle &= \langle \mathbf{y}_3 | \mathbf{x}_1 \rangle + \mu \langle \mathbf{y}_1 | \mathbf{y}_1 \rangle + \nu \langle \mathbf{y}_2 | \mathbf{y}_1 \rangle \\ 0 &= \langle \mathbf{y}_3 | \mathbf{x}_1 \rangle + \mu \langle \mathbf{y}_1 | \mathbf{y}_1 \rangle + \nu \cdot 0, \end{aligned} \tag{1.29}$$

and

$$\begin{aligned} \langle \mathbf{y}_3 | \mathbf{y}_2 \rangle &= \langle \mathbf{y}_3 | \mathbf{x}_2 \rangle + \mu \langle \mathbf{y}_1 | \mathbf{y}_2 \rangle + \nu \langle \mathbf{y}_2 | \mathbf{y}_2 \rangle \\ 0 &= \langle \mathbf{y}_3 | \mathbf{x}_2 \rangle + \mu \cdot 0 + \nu \langle \mathbf{y}_2 | \mathbf{y}_2 \rangle. \end{aligned} \tag{1.30}$$

As a result,

$$\mu = -\frac{\langle \mathbf{x}_3 | \mathbf{y}_1 \rangle}{\langle \mathbf{y}_1 | \mathbf{y}_1 \rangle}, \quad \nu = -\frac{\langle \mathbf{x}_3 | \mathbf{y}_2 \rangle}{\langle \mathbf{y}_2 | \mathbf{y}_2 \rangle}. \tag{1.31}$$

A generalization of this construction for all the other new base vectors $\mathbf{y}_3, \ldots, \mathbf{y}_n$, and thus a proof by complete induction, proceeds by a generalized construction.

Consider, as an example, the standard Euclidean scalar product denoted by "$\cdot$" and the basis $\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$. Then two orthogonal bases are obtained obtained by taking

(i) either the basis vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$, together with $\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \dfrac{\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}}{\begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$,
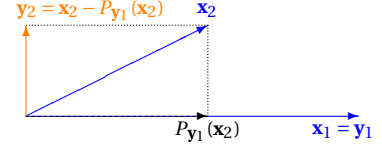
Figure 1.2: Gram-Schmidt construction for two nonorthogonal vectors $\mathbf{x}_1$ and $\mathbf{x}_2$, yielding two orthogonal vectors $\mathbf{y}_1$ and $\mathbf{y}_2$.

(ii)   or the basis vector $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, together with $\begin{pmatrix} 0 \\ 1 \end{pmatrix} - \dfrac{\begin{pmatrix} 0 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}}{\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$.

1.8   Dual space

Every vector space $\mathfrak{V}$ has a corresponding *dual vector space* (or just *dual space*) $\mathfrak{V}^*$ consisting of all linear functionals on $\mathfrak{V}$.

A *linear functional* on a vector space $\mathfrak{V}$ is a scalar-valued linear function $\mathbf{y}$ defined for every vector $\mathbf{x} \in \mathfrak{V}$, with the linear property that

$$\mathbf{y}(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 \mathbf{y}(\mathbf{x}_1) + \alpha_2 \mathbf{y}(\mathbf{x}_2). \tag{1.32}$$

For example, let $\mathbf{x} = (x_1, \ldots, x_n)$, and take $\mathbf{y}(\mathbf{x}) = x_1$.

For another example, let again $\mathbf{x} = (x_1, \ldots, x_n)$, and let $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ be scalars; and take $\mathbf{y}(\mathbf{x}) = \alpha_1 x_1 + \cdots + \alpha_n x_n$.

The following supermarket example has been communicated to me by Hans Havlicek [10]: suppose you visit a supermarket, with a variety of products therein. Suppose further that you select some items and collect them in a cart or trolley. Suppose further that, in order to complete your purchase, you finally go to the cash desk, where the sum total of your purchase is computed from the price-per-product information stored in the memory of the cash register.

In this example, the vector space can be identified with all conceivable configurations of products in a cart or trolley. Its dimension is determined by the number of different, mutually distinct products in the supermarket. Its "base vectors" can be identified with the mutually distinct products in the supermarket. The respective functional is the computation of the price of any such purchase. It is based on a particular price information. Every such price information contains one price per item for all mutually distinct products. The dual space consists of all conceivable price informations.

We adopt a square bracket notation "$[\cdot, \cdot]$" for the functional

$$\mathbf{y}(\mathbf{x}) = [\mathbf{x}, \mathbf{y}]. \tag{1.33}$$

Note that the usual arithmetic operations of addition and multiplication, that is,

$$(a\mathbf{y} + b\mathbf{z})(\mathbf{x}) = a\mathbf{y}(\mathbf{x}) + b\mathbf{z}(\mathbf{x}), \tag{1.34}$$

together with the "zero functional" (mapping every argument to zero) induce a kind of linear vector space, the "vectors" being identified with the linear functionals. This vector space will be called *dual space* $\mathfrak{V}^*$.

As a result, this "bracket" functional is *bilinear* in its two arguments; that is,

$$[\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2, \mathbf{y}] = \alpha_1 [\mathbf{x}_1, \mathbf{y}] + \alpha_2 [\mathbf{x}_2, \mathbf{y}], \tag{1.35}$$

and

$$[\mathbf{x}, \alpha_1 \mathbf{y}_1 + \alpha_2 \mathbf{y}_2] = \alpha_1 [\mathbf{x}, \mathbf{y}_1] + \alpha_2 [\mathbf{x}, \mathbf{y}_2]. \tag{1.36}$$

[10] Hans Havlicek. private communication, 2016

The square bracket can be identified with the scalar dot product $[\mathbf{x}, \mathbf{y}] = \langle \mathbf{x} \mid \mathbf{y} \rangle$ only for Euclidean vector spaces $\mathbb{R}^n$, since for complex spaces this would no longer be positive definite. That is, for Euclidean vector spaces $\mathbb{R}^n$ the inner or scalar product is bilinear.

Because of linearity, we can completely characterize an arbitrary linear functional $\mathbf{y} \in \mathfrak{V}^*$ by its values of the vectors of some basis of $\mathfrak{V}$: If we know the functional value on the basis vectors in $\mathfrak{B}$, we know the functional on all elements of the vector space $\mathfrak{V}$. If $\mathfrak{V}$ is an n-dimensional vector space, and if $\mathfrak{B} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ is a basis of $\mathfrak{V}$, and if $\{\alpha_1, \ldots, \alpha_n\}$ is any set of n scalars, then there is a unique linear functional $\mathbf{y}$ on $\mathfrak{V}$ such that $[\mathbf{f}_i, \mathbf{y}] = \alpha_i$ for all $0 \le i \le n$.

A constructive proof of this theorem can be given as follows: Because every $\mathbf{x} \in \mathfrak{V}$ can be written as a linear combination $\mathbf{x} = x_1 \mathbf{f}_1 + \cdots + x_n \mathbf{f}_n$ of the basis vectors of $\mathfrak{B} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ in one and only one (unique) way, we obtain for any arbitrary linear functional $\mathbf{y} \in \mathfrak{V}^*$ a unique decomposition in terms of the basis vectors of $\mathfrak{B} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$; that is,

$$[\mathbf{x}, \mathbf{y}] = x_1 [\mathbf{f}_1, \mathbf{y}] + \cdots + x_n [\mathbf{f}_n, \mathbf{y}]. \tag{1.37}$$

By identifying $[\mathbf{f}_i, \mathbf{y}] = \alpha_i$ we obtain

$$[\mathbf{x}, \mathbf{y}] = x_1 \alpha_1 + \cdots + x_n \alpha_n. \tag{1.38}$$

Conversely, if we *define* $\mathbf{y}$ by $[\mathbf{x}, \mathbf{y}] = \alpha_1 x_1 + \cdots + \alpha_n x_n$, then $\mathbf{y}$ can be interpreted as a linear functional in $\mathfrak{V}^*$ with $[\mathbf{f}_i, \mathbf{y}] = \alpha_i$.

If we introduce a *dual basis* by requiring that $[\mathbf{f}_i, \mathbf{f}_j^*] = \delta_{ij}$ (cf. Eq. 1.39 below), then the coefficients $[\mathbf{f}_i, \mathbf{y}] = \alpha_i$, $1 \le i \le n$, can be interpreted as the *coordinates* of the linear functional $\mathbf{y}$ with respect to the dual basis $\mathfrak{B}^*$, such that $\mathbf{y} = (\alpha_1, \alpha_2, \ldots, \alpha_n)^T$.

Likewise, as will be shown in (1.46), $x_i = [\mathbf{x}, \mathbf{f}_i^*]$; that is, the vector coordinates can be represented by the functionals of the elements of the dual basis.

Let us explicitly construct an example of a linear functional $\varphi(\mathbf{x}) \equiv [\mathbf{x}, \varphi]$ that is defined on all vectors $\mathbf{x} = \alpha \mathbf{e}_1 + \beta \mathbf{e}_2$ of a two-dimensional vector space with the basis $\{\mathbf{e}_1, \mathbf{e}_2\}$ by enumerating its "performance on the basis vectors" $\mathbf{e}_1 = (1, 0)$ and $\mathbf{e}_2 = (0, 1)$; more explicitly, say, for an example's sake, $\varphi(\mathbf{e}_1) \equiv [\mathbf{e}_1, \varphi] = 2$ and $\varphi(\mathbf{e}_2) \equiv [\mathbf{e}_2, \varphi] = 3$. Therefore, for example, $\varphi((5, 7)) \equiv [(5, 7), \varphi] = 5[\mathbf{e}_1, \varphi] + 7[\mathbf{e}_2, \varphi] = 10 + 21 = 31$.

1.8.1   Dual basis

We now can define a *dual basis*, or, used synonymuously, a *reciprocal* or *contravariant* basis. If $\mathfrak{V}$ is an n-dimensional vector space, and if $\mathfrak{B} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ is a basis of $\mathfrak{V}$, then there is a unique *dual basis* $\mathfrak{B}^* = \{\mathbf{f}_1^*, \ldots, \mathbf{f}_n^*\}$ in the dual vector space $\mathfrak{V}^*$ defined by

$$\mathbf{f}_j^*(\mathbf{f}_i) = [\mathbf{f}_i, \mathbf{f}_j^*] = \delta_{ij}, \tag{1.39}$$

where δ_{ij} is the Kronecker delta function. The dual space $\mathfrak{V}^*$ spanned by the dual basis $\mathfrak{B}^*$ is n-dimensional.

In a different notation involving subscripts (lower indices) for (basis) vectors of the base vector space, and superscripts (upper indices) $\mathbf{f}^j = \mathbf{f}_j^*$, for (basis) vectors of the dual vector space, Eq. (1.39) can be written as

$$\mathbf{f}^j(\mathbf{f}_i) = [\mathbf{f}_i, \mathbf{f}^j] = \delta_{ij}. \tag{1.40}$$

Suppose g is a *metric*, facilitating the translation from vectors of the base vectors into vectors of the dual space and *vice versa* (cf. Section 2.7.1 on page 84 for a definition and more details), in particular, $\mathbf{f}_i = g_{il}\mathbf{f}^l$ as well as $\mathbf{f}_j^* = \mathbf{f}^j = g^{jk}\mathbf{f}_k$. Then Eqs. (1.39) and (1.39) can be rewritten as

$$[g_{il}\mathbf{f}^l, \mathbf{f}^j] = [\mathbf{f}_i, g^{jk}\mathbf{f}_k] = \delta_{ij}. \tag{1.41}$$

Note that the vectors $\mathbf{f}_i^* = \mathbf{f}^i$ of the dual basis can be used to "retreive" the components of arbitrary vectors $\mathbf{x} = \sum_j X_j\mathbf{f}_j$ through

$$\mathbf{f}_i^*(\mathbf{x}) = \mathbf{f}_i^*\left(\sum_j X_j\mathbf{f}_j\right) = \sum_j X_j\mathbf{f}_i^*\left(\mathbf{f}_j\right) = \sum_j X_j\delta_{ij} = X_i. \tag{1.42}$$

Likewise, the basis vectors $\mathbf{f}_i$ can be used to obtain the coordinates of any dual vector.

In terms of the inner products of the base vector space and its dual vector space the representation of the metric may be defined by $g_{ij} = g(\mathbf{f}_i, \mathbf{f}_j) = \langle \mathbf{f}_i \mid \mathbf{f}_j \rangle$, as well as $g^{ij} = g(\mathbf{f}^i, \mathbf{f}^j) = \langle \mathbf{f}^i \mid \mathbf{f}^j \rangle$, respectively. Note however, that the coordinates g_{ij} of the metric g need not necessarily be positive definite. For example, special relativity uses the "pseudo-Euclidean" metric $g = \text{diag}(+1, +1, +1, -1)$ (or just $g = \text{diag}(+, +, +, -)$), where "diag" stands for the *diagonal matrix* with the arguments in the diagonal.

In a real Euclidean vector space $\mathbb{R}^n$ with the dot product as the scalar product, the dual basis of an orthogonal basis is also orthogonal, and contains vectors with the same directions, although with *reciprocal length* (thereby exlaining the wording "reciprocal basis"). Moreover, for an orthonormal basis, the basis vectors are uniquely identifiable by $\mathbf{e}_i \longrightarrow \mathbf{e}_i^* = \mathbf{e}_i^T$. This identification can only be made for orthonomal bases; it is *not* true for nonorthonormal bases.

A "reverse construction" of the elements $\mathbf{f}_j^*$ of the dual basis $\mathfrak{B}^*$ – thereby using the definition "$[\mathbf{f}_i, \mathbf{y}] = \alpha_i$ for all $1 \le i \le n$" for any element $\mathbf{y}$ in $\mathfrak{V}^*$ introduced earlier – can be given as follows: for every $1 \le j \le n$, we can *define* a vector $\mathbf{f}_j^*$ in the dual basis $\mathfrak{B}^*$ by the *requirement* $[\mathbf{f}_i, \mathbf{f}_j^*] = \delta_{ij}$. That is, in words: the dual basis element, when applied to the elements of the original n-dimensional basis, yields one if and only if it corresponds to the respective equally indexed basis element; for all the other $n-1$ basis elements it yields zero.

What remains to be proven is the conjecture that $\mathfrak{B}^* = \{\mathbf{f}_1^*, \ldots, \mathbf{f}_n^*\}$ is a basis of $\mathfrak{V}^*$; that is, that the vectors in $\mathfrak{B}^*$ are linear independent, and that they span $\mathfrak{V}^*$.

First observe that $\mathfrak{B}^*$ is a set of linear independent vectors, for if $\alpha_1\mathbf{f}_1^* + \cdots + \alpha_n\mathbf{f}_n^* = 0$, then also

$$[\mathbf{x}, \alpha_1\mathbf{f}_1^* + \cdots + \alpha_n\mathbf{f}_n^*] = \alpha_1[\mathbf{x}, \mathbf{f}_1^*] + \cdots + \alpha_n[\mathbf{x}, \mathbf{f}_n^*] = 0 \tag{1.43}$$

for arbitrary $\mathbf{x} \in \mathfrak{V}$. In particular, by identifying $\mathbf{x}$ with $\mathbf{f}_i \in \mathfrak{B}$, for $1 \le i \le n$,

$$\alpha_1[\mathbf{f}_i, \mathbf{f}_1^*] + \cdots + \alpha_n[\mathbf{f}_i, \mathbf{f}_n^*] = \alpha_j[\mathbf{f}_i, \mathbf{f}_j^*] = \alpha_j\delta_{ij} = \alpha_i = 0. \tag{1.44}$$

Second, every $\mathbf{y} \in \mathfrak{V}^*$ is a linear combination of elements in $\mathfrak{B}^* = \{\mathbf{f}_1^*, \ldots, \mathbf{f}_n^*\}$, because by starting from $[\mathbf{f}_i, \mathbf{y}] = \alpha_i$, with $\mathbf{x} = x_1\mathbf{f}_1 + \cdots + x_n\mathbf{f}_n$

The metric tensor g_{ij} represents a *bilinear functional* $g(\mathbf{x}, \mathbf{y}) = x^i y^j g_{ij}$ that is *symmetric*; that is, $g(\mathbf{x}, \mathbf{y}) = g(\mathbf{x}, \mathbf{y})$ and *nondegenerate*; that is, for any nonzero vector $\mathbf{x} \in \mathfrak{V}$, $\mathbf{x} \ne 0$, there is some vector $\mathbf{y} \in \mathfrak{V}$, so that $g(\mathbf{x}, \mathbf{y}) \ne 0$. g also satisfies the triangle inequality $||\mathbf{x} - \mathbf{z}|| \le ||\mathbf{x} - \mathbf{y}|| + ||\mathbf{y} - \mathbf{z}||$.

we obtain

$$[\mathbf{x}, \mathbf{y}] = x_1 [\mathbf{f}_1, \mathbf{y}] + \cdots + x_n [\mathbf{f}_n, \mathbf{y}] = x_1 \alpha_1 + \cdots + x_n \alpha_n. \qquad (1.45)$$

Note that , for arbitrary $\mathbf{x} \in \mathfrak{V}$,

$$[\mathbf{x}, \mathbf{f}_i^*] = x_1 [\mathbf{f}_1, \mathbf{f}_i^*] + \cdots + x_n [\mathbf{f}_n, \mathbf{f}_i^*] = x_j [\mathbf{f}_j, \mathbf{f}_i^*] = x_j \delta_{ji} = x_i, \qquad (1.46)$$

and by substituting $[\mathbf{x}, \mathbf{f}_i]$ for x_i in Eq. (1.45) we obtain

$$\begin{aligned}
[\mathbf{x}, \mathbf{y}] &= x_1 \alpha_1 + \cdots + x_n \alpha_n \\
&= [\mathbf{x}, \mathbf{f}_1] \alpha_1 + \cdots + [\mathbf{x}, \mathbf{f}_n] \alpha_n \\
&= [\mathbf{x}, \alpha_1 \mathbf{f}_1 + \cdots + \alpha_n \mathbf{f}_n],
\end{aligned} \qquad (1.47)$$

and therefore $\mathbf{y} = \alpha_1 \mathbf{f}_1 + \cdots + \alpha_n \mathbf{f}_n = \alpha_i \mathbf{f}_i$.

How can one determine the dual basis from a given, not necessarily orthogonal, basis? For the rest of this section, suppose that the metric is identical to the Euclidean metric $\mathrm{diag}(+, +, \cdots, +)$ representable as the usual "dot product." The tuples of *column vectors* of the basis $\mathfrak{B} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ can be arranged into a $n \times n$ matrix

$$\mathbf{B} \equiv \Big(|\mathbf{f}_1\rangle, |\mathbf{f}_2\rangle, \cdots, |\mathbf{f}_n\rangle \Big) \equiv \Big(\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_n \Big) = \begin{pmatrix} \mathbf{f}_{1,1} & \cdots & \mathbf{f}_{n,1} \\ \mathbf{f}_{1,2} & \cdots & \mathbf{f}_{n,2} \\ \vdots & \vdots & \vdots \\ \mathbf{f}_{1,n} & \cdots & \mathbf{f}_{n,n} \end{pmatrix}. \qquad (1.48)$$

Then take the *inverse matrix* $\mathbf{B}^{-1}$, and interpret the *row vectors* $\mathbf{f}_i^*$ of

$$\mathbf{B}^* = \mathbf{B}^{-1} \equiv \begin{pmatrix} \langle \mathbf{f}_1 | \\ \langle \mathbf{f}_2 | \\ \vdots \\ \langle \mathbf{f}_n | \end{pmatrix} \equiv \begin{pmatrix} \mathbf{f}_1^* \\ \mathbf{f}_2^* \\ \vdots \\ \mathbf{f}_n^* \end{pmatrix} = \begin{pmatrix} \mathbf{f}_{1,1}^* & \cdots & \mathbf{f}_{1,n}^* \\ \mathbf{f}_{2,1}^* & \cdots & \mathbf{f}_{2,n}^* \\ \vdots & \vdots & \vdots \\ \mathbf{f}_{n,1}^* & \cdots & \mathbf{f}_{n,n}^* \end{pmatrix} \qquad (1.49)$$

as the tuples of elements of the dual basis of $\mathfrak{B}^*$.

For orthogonal but not orthonormal bases, the term *reciprocal* basis can be easily explained from the fact that the norm (or length) of each vector in the *reciprocal basis* is just the *inverse* of the length of the original vector.

For a direct proof, consider $\mathbf{B} \cdot \mathbf{B}^{-1} = \mathbb{I}_n$.

(i) For example, if

$$\mathfrak{B} \equiv \{|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle, \ldots, |\mathbf{e}_n\rangle\} \equiv \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} \equiv \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\} \qquad (1.50)$$

is the standard basis in n-dimensional vector space containing unit vectors of norm (or length) one, then

$$\mathfrak{B}^* \equiv \{\langle \mathbf{e}_1 |, \langle \mathbf{e}_2 |, \ldots, \langle \mathbf{e}_n |\} \qquad (1.51)$$
$$\equiv \{\mathbf{e}_1^*, \mathbf{e}_2^*, \ldots, \mathbf{e}_n^*\} \equiv \{(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)\}$$

has elements with identical components, but those tuples are the transposed ones.

(ii) If

$$\mathfrak{X} \equiv \{\alpha_1|\mathbf{e}_1\rangle, \alpha_2|\mathbf{e}_2\rangle, \ldots, \alpha_n|\mathbf{e}_n\rangle\} \equiv \{\alpha_1\mathbf{e}_1, \alpha_2\mathbf{e}_2, \ldots, \alpha_n\mathbf{e}_n\}$$

$$\equiv \left\{ \begin{pmatrix} \alpha_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ \alpha_2 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ 0 \\ \vdots \\ \alpha_n \end{pmatrix} \right\}, \quad (1.52)$$

with nonzero $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$, is a "dilated" basis in n-dimensional vector space containing vectors of norm (or length) α_i, then

$$\mathfrak{X}^* \equiv \left\{ \frac{1}{\alpha_1}\langle\mathbf{e}_1|, \frac{1}{\alpha_2}\langle\mathbf{e}_2|, \ldots, \frac{1}{\alpha_n}\langle\mathbf{e}_n| \right\}$$

$$\equiv \left\{ \frac{1}{\alpha_1}\mathbf{e}_1^*, \frac{1}{\alpha_2}\mathbf{e}_2^*, \ldots, \frac{1}{\alpha_n}\mathbf{e}_n^* \right\} \quad (1.53)$$

$$\equiv \left\{ \left(\frac{1}{\alpha_1}, 0, \ldots, 0\right), \left(0, \frac{1}{\alpha_2}, \ldots, 0\right), \ldots, \left(0, 0, \ldots, \frac{1}{\alpha_n}\right) \right\}$$

has elements with identical components of inverse length $\frac{1}{\alpha_i}$, and again those tuples are the transposed tuples.

(iii) Consider the nonorthogonal basis $\mathfrak{B} = \left\{ \begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \end{pmatrix} \right\}$. The associated column matrix is

$$\mathbf{B} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}. \quad (1.54)$$

The inverse matrix is

$$\mathbf{B}^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}, \quad (1.55)$$

and the associated dual basis is obtained from the rows of $\mathbf{B}^{-1}$ by

$$\mathfrak{B}^* = \left\{ \left(-2, 1\right), \left(\tfrac{3}{2}, -\tfrac{1}{2}\right) \right\} = \frac{1}{2}\left\{ \left(-4, 2\right), \left(3, -1\right) \right\}. \quad (1.56)$$

1.8.2 Dual coordinates

With respect to a given basis, the components of a vector are often written as tuples of ordered ("x_i is written before x_{i+1}" – not "$x_i < x_{i+1}$") scalars as *column vectors*

$$|\mathbf{x}\rangle \equiv \mathbf{x} \equiv \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad (1.57)$$

whereas the components of vectors in dual spaces are often written in terms of tuples of ordered scalars as *row vectors*

$$\langle\mathbf{x}| \equiv \mathbf{x}^* \equiv \left(x_1^*, x_2^*, \ldots, x_n^* \right). \quad (1.58)$$

The coordinates of vectors $|\mathbf{x}\rangle \equiv \mathbf{x}$ of the base vector space $\mathfrak{V}$ – and by definition (or rather, declaration) the vectors $|\mathbf{x}\rangle \equiv \mathbf{x}$ themselves – are called *contravariant*: because in order to compensate for scale changes

of the reference axes (the basis vectors) $|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle, \ldots, |\mathbf{e}_n\rangle \equiv \mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ these coordinates have to contra-vary (inversely vary) with respect to any such change.

In contradistinction the coordinates of dual vectors, that is, vectors of the dual vector space $\mathfrak{V}^*$, $\langle\mathbf{x}| \equiv \mathbf{x}^*$ – and by definition (or rather, declaration) the vectors $\langle\mathbf{x}| \equiv \mathbf{x}^*$ themselves – are called *covariant*.

Alternatively to the notation used so far (and throughout this chapter) covariant coordinates can be denoted by subscripts (lower indices), and contravariant coordinates can be denoted by superscripts (upper indices); that is (see also Havlicek [11], Section 11.4),

[11] Hans Havlicek. *Lineare Algebra für Technische Mathematiker*. Heldermann Verlag, Lemgo, second edition, 2008

$$\mathbf{x} \equiv |\mathbf{x}\rangle \equiv \left[\begin{pmatrix} x^1 \\ x^2 \\ \vdots \\ x^n \end{pmatrix} \right], \text{ and} \tag{1.59}$$

$$\mathbf{x}^* \equiv \langle\mathbf{x}| \equiv \left[(x_1^*, x_2^*, \ldots, x_n^*) \right]_i \equiv [(x_1, x_2, \ldots, x_n)].$$

This notation will be used in the chapter 2 on tensors. Note again that the covariant and contravariant components x_k and x^k are not absolute, but always defined *with respect to* a particular (dual) basis.

Note that, for orthormal bases it is possible to interchange contravariant and covariant coordinates by taking the conjugate transpose; that is,

$$(\langle\mathbf{x}|)^\dagger = |\mathbf{x}\rangle, \text{ and } (|\mathbf{x}\rangle)^\dagger = \langle\mathbf{x}|. \tag{1.60}$$

Note also that the *Einstein summation convention* requires that, when an index variable appears twice in a single term, one has to sum over all of the possible index values. This saves us from drawing the sum sign "$\sum_i$" for the index i; for instance $x_i y_i = \sum_i x_i y_i$.

In the particular context of covariant and contravariant components – made necessary by nonorthogonal bases whose associated dual bases are *not* identical – the summation always is between some superscript (upper index) and some subscript (lower index); e.g., $x_i y^i$.

Note again that for orthonormal basis, $x^i = x_i$.

1.8.3   *Representation of a functional by inner product*

The following representation theorem, often called *Riesz representation theorem* (sometimes also called the *Fréchet-Riesz theorem*), is about the connection between any functional in a vector space and its inner product: To any linear functional $\mathbf{z}$ on a finite-dimensional inner product space $\mathfrak{V}$ there corresponds a unique vector $\mathbf{y} \in \mathfrak{V}$, such that

For proofs and additional information see §67 in

Paul Richard Halmos. *Finite-dimensional Vector Spaces*. Undergraduate Texts in Mathematics. Springer, New York, 1958. ISBN 978-1-4612-6387-6,978-0-387-90093-3. DOI: 10.1007/978-1-4612-6387-6. URL http://doi.org/10.1007/978-1-4612-6387-6

$$\mathbf{z}(\mathbf{x}) = [\mathbf{x}, \mathbf{z}] = \langle\mathbf{x} \mid \mathbf{y}\rangle \tag{1.61}$$

for all $\mathbf{x} \in \mathfrak{V}$.

For a proof, let us first consider the case of $\mathbf{z} = 0$, for which we can just *ad hoc* identify $\mathbf{y} = 0$.

For any nonzero $\mathbf{z} \neq 0$ we can locate the subspace $\mathfrak{M}$ of $\mathfrak{V}$ consisting of all vectors $\mathbf{x}$ for which $\mathbf{z}$ vanishes; that is, $\mathbf{z}(\mathbf{x}) = [\mathbf{x}, \mathbf{z}] = 0$. Let $\mathfrak{N} = \mathfrak{M}^\perp$ be the orthogonal complement of $\mathfrak{M}$ with respect to $\mathfrak{V}$; that is, $\mathfrak{N}$ consists of

all vectors orthogonal to all vectors in $\mathfrak{M}$. Surely $\mathfrak{N}$ consists of a nonzero unit vector $\mathbf{y}_0$. Based on $\mathbf{y}_0$ we can define the vector $\mathbf{y} = \overline{\mathbf{z}(\mathbf{y}_0)}\mathbf{y}_0$. $\mathbf{y}$ satisfies Eq. (1.61) at least (i) for $\mathbf{x} = \mathbf{y}_0$, (ii) as well as for all $\mathbf{x} \in \mathfrak{M}$ through the orthogonality of $\mathbf{y}_0$ with all vectors in $\mathfrak{M}$.

For an arbitrary vector $\mathbf{x} \in \mathfrak{V}$, define $\mathbf{x}_0 = \mathbf{x} - \frac{\mathbf{z}(\mathbf{x})}{\mathbf{z}(\mathbf{y}_0)}\mathbf{y}_0$, or $\mathbf{z}(\mathbf{x})\mathbf{y}_0 = \mathbf{z}(\mathbf{y}_0)\,[\mathbf{x} - \mathbf{x}_0]$, such that $\mathbf{z}(\mathbf{x}_0) = 0$; that is, $\mathbf{x}_0 \in \mathfrak{M}$. Therefore, $\mathbf{x} = \mathbf{x}_0 + \frac{\mathbf{z}(\mathbf{x})}{\mathbf{z}(\mathbf{y}_0)}\mathbf{y}_0$ is a linear combination of two vectors $\mathbf{x}_0 \in \mathfrak{M}$ and $\mathbf{y}_0 \in \mathfrak{N}$, for each of which Eq. (1.61) is valid. Because $\mathbf{x}$ is a linear combination of $\mathbf{x}_0$ and $\mathbf{y}_0$, Eq. (1.61) holds for $\mathbf{x}$ as well.

The proof of uniqueness is by supposing that there exist two (presumably different) $\mathbf{y}_1$ and $\mathbf{y}_2$ such that $(\mathbf{x}, \mathbf{y}_1) = (\mathbf{x}, \mathbf{y}_2)$. Due to linearity of the scalar product, $(\mathbf{x}, \mathbf{y}_1 - \mathbf{y}_2) = 0$ for all $\mathbf{x} \in \mathfrak{V}$. In particular, if we identify $\mathbf{x} = \mathbf{y}_1 - \mathbf{y}_2$, then $(\mathbf{y}_1 - \mathbf{y}_2, \mathbf{y}_1 - \mathbf{y}_2) = 0$ and thus $\mathbf{y}_1 = \mathbf{y}_2$.

This proof is constructive in the sense that it yields $\mathbf{y}$, given $\mathbf{z}$.

Note that in real or complex vector space $\mathbb{R}^n$ or $\mathbb{C}^n$, and with the dot product, $\mathbf{y}^\dagger = \mathbf{z}$.

Note also that every inner product $\langle \mathbf{x} \mid \mathbf{y} \rangle = \phi_y(x)$ defines a linear functional $\phi_y(x)$ for all $\mathbf{x} \in \mathfrak{V}$.

In quantum mechanics, this representation of a functional by the inner product suggests the (unique) existence of the bra vector $\langle \psi | \in \mathfrak{V}^*$ associated with every ket vector $|\psi\rangle \in \mathfrak{V}$.

It also suggests a "natural" duality between propositions and states – that is, between (i) dichotomic (yes/no, or 1/0) observables represented by projections $\mathbf{E}_x = |\mathbf{x}\rangle\langle\mathbf{x}|$ and their associated linear subspaces spanned by unit vectors $|\mathbf{x}\rangle$ on the one hand, and (ii) pure states, which are also represented by projections $\boldsymbol{\rho}_\psi = |\psi\rangle\langle\psi|$ and their associated subspaces spanned by unit vectors $|\psi\rangle$ on the other hand – *via* the scalar product "$\langle\cdot|\cdot\rangle$." In particular [12],

[12] Jan Hamhalter. *Quantum Measure Theory. Fundamental Theories of Physics, Vol. 134.* Kluwer Academic Publishers, Dordrecht, Boston, London, 2003. ISBN 1-4020-1714-6

$$\psi(\mathbf{x}) \equiv [\mathbf{x}, \psi] = \langle \mathbf{x} \mid \psi \rangle \tag{1.62}$$

represents the *probability amplitude*. By the *Born rule* for pure states, the absolute square $|\langle \mathbf{x} \mid \psi \rangle|^2$ of this probability amplitude is identified with the probability of the occurrence of the proposition $\mathbf{E}_x$, given the state $|\psi\rangle$.

More general, due to linearity and the spectral theorem (cf. Section 1.28.1 on page 58), the statistical expectation for a Hermitian (normal) operator $\mathbf{A} = \sum_{i=0}^{k} \lambda_i \mathbf{E}_i$ and a quantized system prepared in pure state (cf.

Sec. 1.12) $\boldsymbol{\rho}_\psi = |\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle$ is given by the *Born rule*

$$
\begin{aligned}
\langle \mathbf{A} \rangle_\psi = \mathrm{Tr}(\boldsymbol{\rho}_\psi \mathbf{A}) &= \mathrm{Tr}\left[\boldsymbol{\rho}_\psi \left(\sum_{i=0}^k \lambda_i \mathbf{E}_i\right)\right] = \mathrm{Tr}\left(\sum_{i=0}^k \lambda_i \boldsymbol{\rho}_\psi \mathbf{E}_i\right) \\
&= \mathrm{Tr}\left(\sum_{i=0}^k \lambda_i (|\psi\rangle\langle\psi|)(|\mathbf{x}_i\rangle\langle\mathbf{x}_i|)\right) = \mathrm{Tr}\left(\sum_{i=0}^k \lambda_i |\psi\rangle\langle\psi|\mathbf{x}_i\rangle\langle\mathbf{x}_i|\right) \\
&= \sum_{j=0}^k \langle\mathbf{x}_j|\left(\sum_{i=0}^k \lambda_i |\psi\rangle\langle\psi|\mathbf{x}_i\rangle\langle\mathbf{x}_i|\right)|\mathbf{x}_j\rangle \\
&= \sum_{j=0}^k \sum_{i=0}^k \lambda_i \langle\mathbf{x}_j|\psi\rangle\langle\psi|\mathbf{x}_i\rangle \underbrace{\langle\mathbf{x}_i|\mathbf{x}_j\rangle}_{\delta_{ij}} \\
&= \sum_{i=0}^k \lambda_i \langle\mathbf{x}_i|\psi\rangle\langle\psi|\mathbf{x}_i\rangle = \sum_{i=0}^k \lambda_i |\langle\mathbf{x}_i|\psi\rangle|^2,
\end{aligned}
\tag{1.63}
$$

where Tr stands for the trace (cf. Section 1.18 on page 40), and we have used the spectral decomposition $\mathbf{A} = \sum_{i=0}^k \lambda_i \mathbf{E}_i$ (cf. Section 1.28.1 on page 58).

1.8.4 Double dual space

In the following, we strictly limit the discussion to finite dimensional vector spaces.

Because to every vector space $\mathfrak{V}$ there exists a dual vector space $\mathfrak{V}^*$ "spanned" by all linear functionals on $\mathfrak{V}$, there exists also a dual vector space $(\mathfrak{V}^*)^* = \mathfrak{V}^{**}$ to the dual vector space $\mathfrak{V}^*$ "spanned" by all linear functionals on $\mathfrak{V}^*$. We state without proof that $\mathfrak{V}^{**}$ is closely related to, and can be *canonically identified* with $\mathfrak{V}$ *via* the *canonical bijection*

$$
\begin{aligned}
\mathfrak{V} \to \mathfrak{V}^{**} : \mathbf{x} \mapsto \langle\cdot|\mathbf{x}\rangle, \text{ with} \\
\langle\cdot|\mathbf{x}\rangle : \mathfrak{V}^* \to \mathbb{R} \text{ or } \mathbb{C} : \mathbf{a}^* \mapsto \langle\mathbf{a}^*|\mathbf{x}\rangle;
\end{aligned}
\tag{1.64}
$$

indeed, more generally,

$$
\begin{aligned}
\mathfrak{V} &\equiv \mathfrak{V}^{**}, \\
\mathfrak{V}^* &\equiv \mathfrak{V}^{***}, \\
\mathfrak{V}^{**} &\equiv \mathfrak{V}^{****} \equiv \mathfrak{V}, \\
\mathfrak{V}^{***} &\equiv \mathfrak{V}^{*****} \equiv \mathfrak{V}^*, \\
&\vdots
\end{aligned}
\tag{1.65}
$$

1.9 Direct sum

Let $\mathfrak{U}$ and $\mathfrak{V}$ be vector spaces (over the same field, say $\mathbb{C}$). Their *direct sum* is a vector space $\mathfrak{W} = \mathfrak{U} \oplus \mathfrak{V}$ consisting of all ordered pairs $(\mathbf{x}, \mathbf{y})$, with $\mathbf{x} \in \mathfrak{U}$ in $\mathbf{y} \in \mathfrak{V}$, and with the linear operations defined by

$$
(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2, \alpha\mathbf{y}_1 + \beta\mathbf{y}_2) = \alpha(\mathbf{x}_1, \mathbf{y}_1) + \beta(\mathbf{x}_2, \mathbf{y}_2).
\tag{1.66}
$$

Note that, just like vector addition, addition is defined coordinate-wise.

We state without proof that the dimension of the direct sum is the

For proofs and additional information see §18 in

Paul Richard Halmos. *Finite-dimensional Vector Spaces*. Undergraduate Texts in Mathematics. Springer, New York, 1958. ISBN 978-1-4612-6387-6,978-0-387-90093-3. DOI: 10.1007/978-1-4612-6387-6. URL http://doi.org/10.1007/978-1-4612-6387-6

For proofs and additional information see §19 in

Paul Richard Halmos. *Finite-dimensional Vector Spaces*. Undergraduate Texts in Mathematics. Springer, New York, 1958. ISBN 978-1-4612-6387-6,978-0-387-90093-3. DOI: 10.1007/978-1-4612-6387-6. URL http://doi.org/10.1007/978-1-4612-6387-6

sum of the dimensions of its summands.

We also state without proof that, if $\mathfrak{U}$ and $\mathfrak{V}$ are subspaces of a vector space $\mathfrak{W}$, then the following three conditions are equivalent:

(i) $\mathfrak{W} = \mathfrak{U} \oplus \mathfrak{V}$;

(ii) $\mathfrak{U} \bigcap \mathfrak{V} = 0$ and $\mathfrak{U} + \mathfrak{V} = \mathfrak{W}$, that is, $\mathfrak{W}$ is spanned by $\mathfrak{U}$ and $\mathfrak{V}$ (i.e., $\mathfrak{U}$ and $\mathfrak{V}$ are complements of each other);

(iii) every vector $\mathbf{z} \in \mathfrak{W}$ can be written as $\mathbf{z} = \mathbf{x} + \mathbf{y}$, with $\mathbf{x} \in \mathfrak{U}$ and $\mathbf{y} \in \mathfrak{V}$, in one and only one way.

Very often the direct sum will be used to "compose" a vector space by the direct sum of its subspaces. Note that there is no "natural" way of composition. A different way of putting two vector spaces together is by the *tensor product*.

1.10   Tensor product

1.10.1   Sloppy definition

For the moment, suffice it to say that the *tensor product* $\mathfrak{V} \otimes \mathfrak{U}$ of two linear vector spaces $\mathfrak{V}$ and $\mathfrak{U}$ should be such that, to every $\mathbf{x} \in \mathfrak{V}$ and every $\mathbf{y} \in \mathfrak{U}$ there corresponds a tensor product $\mathbf{z} = \mathbf{x} \otimes \mathbf{y} \in \mathfrak{V} \otimes \mathfrak{U}$ which is bilinear; that is, linear in both factors.

A generalization to more factors appears to present no further conceptual difficulties.

1.10.2   Definition

A more rigorous definition is as follows: The *tensor product* $\mathfrak{V} \otimes \mathfrak{U}$ of two vector spaces $\mathfrak{V}$ and $\mathfrak{U}$ (over the same field, say $\mathbb{C}$) is the dual vector space of all bilinear forms on $\mathfrak{V} \oplus \mathfrak{U}$.

For each pair of vectors $\mathbf{x} \in \mathfrak{V}$ and $\mathbf{y} \in \mathfrak{U}$ the tensor product $\mathbf{z} = \mathbf{x} \otimes \mathbf{y}$ is the element of $\mathfrak{V} \otimes \mathfrak{U}$ such that $\mathbf{z}(\mathbf{w}) = \mathbf{w}(\mathbf{x}, \mathbf{y})$ for every bilinear form $\mathbf{w}$ on $\mathfrak{V} \oplus \mathfrak{U}$.

Alternatively we could define the *tensor product* as the coherent superpositions of products $\mathbf{e}_i \otimes \mathbf{f}_j$ of all basis vectors $\mathbf{e}_i \in \mathfrak{V}$, with $1 \leq i \leq n$, and $\mathbf{f}_j \in \mathfrak{U}$, with $1 \leq i \leq m$ as follows. First we note without proof that if $\mathfrak{A} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\mathfrak{B} = \{\mathbf{f}_1, \ldots, \mathbf{f}_m\}$ are bases of n- and m- dimensional vector spaces $\mathfrak{V}$ and $\mathfrak{U}$, respectively, then the set of vectors $\mathbf{e}_i \otimes \mathbf{f}_j$ with $i = 1, \ldots n$ and $j = 1, \ldots m$ is a basis of the tensor product $\mathfrak{V} \otimes \mathfrak{U}$. Then an arbitrary tensor product can be written as the coherent superposition of all its basis vectors $\mathbf{e}_i \otimes \mathbf{f}_j$ with $\mathbf{e}_i \in \mathfrak{V}$, with $1 \leq i \leq n$, and $\mathbf{f}_j \in \mathfrak{U}$, with $1 \leq i \leq m$; that is,

$$\mathbf{z} = \sum_{i,j} c_{ij} \, \mathbf{e}_i \otimes \mathbf{f}_j. \qquad (1.67)$$

We state without proof that the dimension of $\mathfrak{V} \otimes \mathfrak{U}$ of an n-dimensional vector space $\mathfrak{V}$ and an m-dimensional vector space $\mathfrak{U}$ is multiplicative, that is, the dimension of $\mathfrak{V} \otimes \mathfrak{U}$ is nm. Informally, this is evident from the number of basis pairs $\mathbf{e}_i \otimes \mathbf{f}_j$.

1.10.3   Representation

A tensor (dyadic) product $\mathbf{z} = \mathbf{x} \otimes \mathbf{y}$ of two vectors $\mathbf{x}$ and $\mathbf{y}$ has three equivalent notations or representations:

(i)   as the scalar coordinates $x_i y_j$ with respect to the basis in which the vectors $\mathbf{x}$ and $\mathbf{y}$ have been defined and coded;

(ii)   as a quasi-matrix $z_{ij} = x_i y_j$, whose components z_{ij} are defined with respect to the basis in which the vectors $\mathbf{x}$ and $\mathbf{y}$ have been defined and coded;

(iii)   as a quasi-vector or "flattened matrix" defined by the Kronecker product $\mathbf{z} = (x_1\mathbf{y}, x_2\mathbf{y}, \ldots, x_n\mathbf{y})^T = (x_1 y_1, x_1 y_2, \ldots, x_n y_n)^T$. Again, the scalar coordinates $x_i y_j$ are defined with respect to the basis in which the vectors $\mathbf{x}$ and $\mathbf{y}$ have been defined and coded.

In all three cases, the pairs $x_i y_j$ are properly represented by distinct mathematical entities.

Take, for example, $\mathbf{x} = (2,3)^T$ and $\mathbf{y} = (5,7,11)^T$. Then $\mathbf{z} = \mathbf{x} \otimes \mathbf{y}$ can be represented by (i) the four scalars $x_1 y_1 = 10$, $x_1 y_2 = 14$, $x_1 y_3 = 22$, $x_2 y_1 = 15$, $x_2 y_2 = 21$, $x_2 y_3 = 33$, or by (ii) a 2×3 matrix $\begin{pmatrix} 10 & 14 & 22 \\ 15 & 21 & 33 \end{pmatrix}$, or by (iii) a 4-touple $\left(10, 14, 22, 15, 21, 33\right)^T$.

Note, however, that this kind of quasi-matrix or quasi-vector representation of vector products can be misleading insofar as it (wrongly) suggests that all vectors in the tensor product space are accessible (representable) as quasi-vectors – they are, however, accessible by *coherent superpositions* (1.67) of such quasi-vectors. For instance, take the arbitrary form of a (quasi-)vector in $\mathbb{C}^4$, which can be parameterized by

$$\left(\alpha_1, \alpha_2, \alpha_3, \alpha_4\right)^T, \text{ with } \alpha_1, \alpha_3, \alpha_3, \alpha_4 \in \mathbb{C}, \tag{1.68}$$

and compare (1.68) with the general form of a tensor product of two quasi-vectors in $\mathbb{C}^2$

$$\left(a_1, a_2\right)^T \otimes \left(b_1, b_2\right)^T \equiv \left(a_1 b_1, a_1 b_2, a_2 b_1, a_2 b_2\right)^T, \text{ with } a_1, a_2, b_1, b_2 \in \mathbb{C}. \tag{1.69}$$

A comparison of the coordinates in (1.68) and (1.69) yields

$$\alpha_1 = a_1 b_1, \quad \alpha_2 = a_1 b_2, \quad \alpha_3 = a_2 b_1, \quad \alpha_4 = a_2 b_2. \tag{1.70}$$

By taking the quotient of the two first and the two last equations, and by equating these quotients, one obtains

$$\frac{\alpha_1}{\alpha_2} = \frac{b_1}{b_2} = \frac{\alpha_3}{\alpha_4}, \text{ and thus } \alpha_1 \alpha_4 = \alpha_2 \alpha_3, \tag{1.71}$$

which amounts to a condition for the four coordinates $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ in order for this four-dimensional vector to be decomposable into a tensor product of two two-dimensional quasi-vectors. In quantum mechanics, pure states which are not decomposable into a product of single particle states are called *entangled*.

In quantum mechanics this amounts to the fact that not all pure two-particle states can be written in terms of (tensor) products of single-particle states; see also Section 1.5 of

   David N. Mermin. *Quantum Computer Science*. Cambridge University Press, Cambridge, 2007. ISBN 9780521876582. URL http://people.ccmr.cornell.edu/~mermin/qcomp/CS483.html

A typical example of an entangled state is the *Bell state*, $|\Psi^-\rangle$ or, more generally, states in the Bell basis

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|1\rangle - |1\rangle|0\rangle\right), \quad |\Psi^+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|1\rangle + |1\rangle|0\rangle\right),$$
$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|0\rangle - |1\rangle|1\rangle\right), \quad |\Phi^+\rangle = \frac{1}{\sqrt{2}}\left(|0\rangle|0\rangle + |1\rangle|1\rangle\right), \tag{1.72}$$

or just

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle - |10\rangle\right), \quad |\Psi^+\rangle = \frac{1}{\sqrt{2}}\left(|01\rangle + |10\rangle\right),$$
$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle - |11\rangle\right), \quad |\Phi^+\rangle = \frac{1}{\sqrt{2}}\left(|00\rangle + |11\rangle\right). \tag{1.73}$$

For instance, in the case of $|\Psi^-\rangle$ a comparison of coefficient yields

$$\alpha_1 = a_1 b_1 = 0, \quad \alpha_2 = a_1 b_2 = \frac{1}{\sqrt{2}},$$
$$\alpha_3 = a_2 b_1 - \frac{1}{\sqrt{2}}, \quad \alpha_4 = a_2 b_2 = 0; \tag{1.74}$$

and thus the entanglement, since

$$\alpha_1 \alpha_4 = 0 \neq \alpha_2 \alpha_3 = \frac{1}{2}. \tag{1.75}$$

This shows that $|\Psi^-\rangle$ cannot be considered as a two particle product state. Indeed, the state can only be characterized by considering the *relative properties* of the two particles – in the case of $|\Psi^-\rangle$ they are associated with the statements [13]: "the quantum numbers (in this case "0" and "1") of the two particles are always different."

1.11   Linear transformation

1.11.1   Definition

A *linear transformation,* or, used synonymuosly, a *linear operator,* **A** on a vector space $\mathfrak{V}$ is a correspondence that assigns every vector $\mathbf{x} \in \mathfrak{V}$ a vector $\mathbf{A}\mathbf{x} \in \mathfrak{V}$, in a linear way; such that

$$\mathbf{A}(\alpha\mathbf{x} + \beta\mathbf{y}) = \alpha\mathbf{A}(\mathbf{x}) + \beta\mathbf{A}(\mathbf{y}) = \alpha\mathbf{A}\mathbf{x} + \beta\mathbf{A}\mathbf{y}, \tag{1.76}$$

identically for all vectors $\mathbf{x}, \mathbf{y} \in \mathfrak{V}$ and all scalars α, β.

1.11.2   Operations

The *sum* $\mathbf{S} = \mathbf{A} + \mathbf{B}$ of two linear transformations $\mathbf{A}$ and $\mathbf{B}$ is defined by $\mathbf{S}\mathbf{x} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{x}$ for every $\mathbf{x} \in \mathfrak{V}$.

The *product* $\mathbf{P} = \mathbf{A}\mathbf{B}$ of two linear transformations $\mathbf{A}$ and $\mathbf{B}$ is defined by $\mathbf{P}\mathbf{x} = \mathbf{A}(\mathbf{B}\mathbf{x})$ for every $\mathbf{x} \in \mathfrak{V}$.

The notation $\mathbf{A}^n \mathbf{A}^m = \mathbf{A}^{n+m}$ and $(\mathbf{A}^n)^m = \mathbf{A}^{nm}$, with $\mathbf{A}^1 = \mathbf{A}$ and $\mathbf{A}^0 = \mathbf{1}$ turns out to be useful.

With the exception of commutativity, all formal algebraic properties of numerical addition and multiplication, are valid for transformations; that is $\mathbf{A}\mathbf{0} = \mathbf{0}\mathbf{A} = \mathbf{0}$, $\mathbf{A}\mathbf{1} = \mathbf{1}\mathbf{A} = \mathbf{A}$, $\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{A}\mathbf{B} + \mathbf{A}\mathbf{C}$, $(\mathbf{A} + \mathbf{B})\mathbf{C} = \mathbf{A}\mathbf{C} + \mathbf{B}\mathbf{C}$,

and $\mathbf{A}(\mathbf{B}\mathbf{C}) = (\mathbf{A}\mathbf{B})\mathbf{C}$. In *matrix notation*, $\mathbf{1} = \mathbb{1}$, and the entries of $\mathbf{0}$ are 0 everywhere.

The *inverse operator* $\mathbf{A}^{-1}$ of $\mathbf{A}$ is defined by $\mathbf{A}\mathbf{A}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbb{1}$.

The *commutator* of two matrices $\mathbf{A}$ and $\mathbf{B}$ is defined by

$$[\mathbf{A}, \mathbf{B}] = \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A}. \qquad (1.77)$$

The commutator should not be confused with the bilinear funtional introduced for dual spaces.

The *polynomial* can be directly adopted from ordinary arithmetic; that is, any finite polynomial p of degree n of an operator (transformation) $\mathbf{A}$ can be written as

$$p(\mathbf{A}) = \alpha_0 \mathbf{1} + \alpha_1 \mathbf{A}^1 + \alpha_2 \mathbf{A}^2 + \cdots + \alpha_n \mathbf{A}^n = \sum_{i=0}^{n} \alpha_i \mathbf{A}^i. \qquad (1.78)$$

The Baker-Hausdorff formula

$$e^{i\mathbf{A}} \mathbf{B} e^{-i\mathbf{A}} = B + i[\mathbf{A}, \mathbf{B}] + \frac{i^2}{2!}[\mathbf{A}, [\mathbf{A}, \mathbf{B}]] + \cdots \qquad (1.79)$$

for two arbitrary noncommutative linear operators $\mathbf{A}$ and $\mathbf{B}$ is mentioned without proof (cf. Messiah, *Quantum Mechanics, Vol. 1* [14]).

If $[\mathbf{A}, \mathbf{B}]$ commutes with $\mathbf{A}$ and $\mathbf{B}$, then

[14] A. Messiah. *Quantum Mechanics,* volume I. North-Holland, Amsterdam, 1962

$$e^{\mathbf{A}} e^{\mathbf{B}} = e^{\mathbf{A}+\mathbf{B}+\frac{1}{2}[\mathbf{A},\mathbf{B}]}. \qquad (1.80)$$

If $\mathbf{A}$ commutes with $\mathbf{B}$, then

$$e^{\mathbf{A}} e^{\mathbf{B}} = e^{\mathbf{A}+\mathbf{B}}. \qquad (1.81)$$

1.11.3   Linear transformations as matrices

Let $\mathfrak{V}$ be an n-dimensional vector space; let $\mathfrak{B} = \{|\mathbf{f}_1\rangle, |\mathbf{f}_2\rangle, \ldots, |\mathbf{f}_n\rangle\}$ be any basis of $\mathfrak{V}$, and let $\mathbf{A}$ be a linear transformation on $\mathfrak{V}$.

Because every vector is a linear combination of the basis vectors $|\mathbf{f}_i\rangle$, every linear transformation can be defined by "its performance on the basis vectors;" that is, by the particular mapping of all n basis vectors into the transformed vectors, which in turn can be represented as linear combination of the n basis vectors.

Therefore it is possible to define some $n \times n$ matrix with n^2 coefficients or coordinates α_{ij} such that

$$\mathbf{A}|\mathbf{f}_j\rangle = \sum_i \alpha_{ij}|\mathbf{f}_i\rangle \qquad (1.82)$$

for all $j = 1, \ldots, n$. Again, note that this definition of a *transformation matrix* is "tied to" a basis.

The "reverse order" of indices in (1.82) has been chosen in order for the vector coordinates to transform in the "right order:" with (1.17) on page 12,

$$\mathbf{A}|\mathbf{x}\rangle = \mathbf{A}\sum_i x_i|\mathbf{f}_i\rangle = \sum_i \mathbf{A}x_i|\mathbf{f}_i\rangle = \sum_i x_i \mathbf{A}|\mathbf{f}_i\rangle = \sum_{i,j} x_i \alpha_{ji}|\mathbf{f}_j\rangle =$$

$$\sum_{i,j} \alpha_{ji} x_i|\mathbf{f}_j\rangle = (i \leftrightarrow j) = \sum_{i,j} \alpha_{ij} x_j|\mathbf{f}_i\rangle, \qquad (1.83)$$

and, due to the linear independence of the basis vectors and by comparison of the coefficients of $|\mathbf{f}_i\rangle$, with respect to the basis $\mathfrak{B} = \{|\mathbf{f}_1\rangle, |\mathbf{f}_2\rangle, \ldots, |\mathbf{f}_n\rangle\}$,

$$\mathbf{A}x_i = \sum_j \alpha_{ij} x_j. \tag{1.84}$$

For *orthonormal bases* there is an even closer connection – representable as scalar product – between a matrix defined by an n-by-n square array and the representation in terms of the elements of the bases; that is, by inserting two resolutions of the identity (see Section 1.15 on page 36) before and after the linear transformation $\mathbf{A}$,

$$\mathbf{A} = \mathbb{I}_n \mathbf{A} \mathbb{I}_n = \sum_{i,j=1}^n |\mathbf{f}_i\rangle\langle\mathbf{f}_i|A|\mathbf{f}_j\rangle\langle\mathbf{f}_j| = \sum_{i,j=1}^n \alpha_{ij}|\mathbf{f}_i\rangle\langle\mathbf{f}_j|, \tag{1.85}$$

whereby

$$\langle\mathbf{f}_i|A|\mathbf{f}_j\rangle = \langle\mathbf{f}_i|A\mathbf{f}_j\rangle$$

$$= \langle\mathbf{f}_i|\sum_l \alpha_{lj}\mathbf{f}_l\rangle = \sum_l \alpha_{lj}\langle\mathbf{f}_i|\mathbf{f}_l\rangle = \sum_l \alpha_{lj}\delta_{il} = \alpha_{ij}$$

$$\equiv \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{pmatrix}. \tag{1.86}$$

In terms of this matrix notation, it is quite easy to present an example for which the commutator $[\mathbf{A}, \mathbf{B}]$ does not vanish; that is $\mathbf{A}$ and $\mathbf{B}$ do not commute.

Take, for the sake of an example, the *Pauli spin matrices* which are proportional to the angular momentum operators along the x, y, z-axis [15]:

[15] Leonard I. Schiff. *Quantum Mechanics*. McGraw-Hill, New York, 1955

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{1.87}$$

$$\sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Together with the identity, that is, with $\mathbb{I}_2 = \mathrm{diag}(1,1)$, they form a complete basis of all (4×4) matrices. Now take, for instance, the commutator

$$[\sigma_1, \sigma_3] = \sigma_1\sigma_3 - \sigma_3\sigma_1$$

$$= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \tag{1.88}$$

$$= 2\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

1.12   Projection or projection operator

The more I learned about quantum mechanics the more I realized the importance of projection operators for its conceptualization [16]:

For proofs and additional information see §41 in

Paul Richard Halmos. *Finite-dimensional Vector Spaces*. Undergraduate Texts in Mathematics. Springer, New York, 1958. ISBN 978-1-4612-6387-6,978-0-387-90093-3. DOI: 10.1007/978-1-4612-6387-6. URL http://doi.org/10.1007/978-1-4612-6387-6

[16] John von Neumann. *Mathematische Grundlagen der Quantenmechanik.*

(i) Pure quantum states are represented by a very particular kind of projections; namely, those that are of the *trace class one,* meaning their trace (cf. Section 1.18 below) is one, as well as being *positive* (cf. Section 1.21 below). Positivity implies that the projection is *self-adjoint* (cf. Section 1.20 below), which is equivalent to the projection being *orthogonal* (cf. Section 1.23 below).

Mixed quantum states are compositions – actually, nontrivial convex combinations – of (pure) quantum states; again they are of the trace class one, self-adjoint, and positive; yet unlike pure states they are no projectors (that is, they are not idempotent); and the trace of their square is not one (indeed, it is less than one).

(ii) Mixed states, should they ontologically exist, can be composed from projections by summing over projectors.

(iii) Projectors serve as the most elementary obsevables – they correspond to yes-no propositions.

(iv) In Section 1.28.1 we will learn that every observable can be decomposed into weighted (spectral) sums of projections.

(v) Furthermore, from dimension three onwards, Gleason's theorem (cf. Section 1.33.1) allows quantum probability theory to be based upon maximal (in terms of co-measurability) "quasi-classical" blocks of projectors.

(vi) Such maximal blocks of projectors can be bundled together to show (cf. Section 1.33.2) that the corresponding algebraic structure has no two-valued measure (interpretable as truth assignment), and therefore cannot be "embedded" into a "larger" classical (Boolean) algebra.

1.12.1 Definition

If $\mathfrak{V}$ is the direct sum of some subspaces $\mathfrak{M}$ and $\mathfrak{N}$ so that every $\mathbf{z} \in \mathfrak{V}$ can be uniquely written in the form $\mathbf{z} = \mathbf{x} + \mathbf{y}$, with $\mathbf{x} \in \mathfrak{M}$ and with $\mathbf{y} \in \mathfrak{N}$, then the *projection*, or, used synonymuosly, *projection* on $\mathfrak{M}$ along $\mathfrak{N}$ is the transformation $\mathbf{E}$ defined by $\mathbf{Ez} = \mathbf{x}$. Conversely, $\mathbf{Fz} = \mathbf{y}$ is the projection on $\mathfrak{N}$ along $\mathfrak{M}$.

A (nonzero) linear transformation $\mathbf{E}$ is a projection if and only if it is idempotent; that is, $\mathbf{EE} = \mathbf{E} \neq 0$.

For a proof note that, if $\mathbf{E}$ is the projection on $\mathfrak{M}$ along $\mathfrak{N}$, and if $\mathbf{z} = \mathbf{x} + \mathbf{y}$, with $\mathbf{x} \in \mathfrak{M}$ and with $\mathbf{y} \in \mathfrak{N}$, the decomposition of $\mathbf{x}$ yields $\mathbf{x} + 0$, so that $\mathbf{E}^2\mathbf{z} = \mathbf{EEz} = \mathbf{Ex} = \mathbf{x} = \mathbf{Ez}$. The converse – idempotence "$\mathbf{EE} = \mathbf{E}$" implies that $\mathbf{E}$ is a projection – is more difficult to prove. For this proof we refer to the literature; e.g., Halmos [17].

We also mention without proof that a linear transformation $\mathbf{E}$ is a projection if and only if $\mathbf{1} - \mathbf{E}$ is a projection. Note that $(\mathbf{1} - \mathbf{E})^2 = \mathbf{1} - \mathbf{E} - \mathbf{E} + \mathbf{E}^2 = \mathbf{1} - \mathbf{E}$; furthermore, $\mathbf{E}(\mathbf{1} - \mathbf{E}) = (\mathbf{1} - \mathbf{E})\mathbf{E} = \mathbf{E} - \mathbf{E}^2 = 0$.

Furthermore, if $\mathbf{E}$ is the projection on $\mathfrak{M}$ along $\mathfrak{N}$, then $\mathbf{1} - \mathbf{E}$ is the projection on $\mathfrak{N}$ along $\mathfrak{M}$.

1.12.2   Construction of projections from unit vectors

How can we construct projections from unit vectors, or systems of orthogonal projections from some vector in some orthonormal basis with the standard dot product?

Let $\mathbf{x}$ be the coordinates of a unit vector; that is $\|\mathbf{x}\| = 1$. Transposition is indicated by the superscript "T" in real vector space. In complex vector space the transposition has to be substituted for the *conjugate transpose* (also denoted as *Hermitian conjugate* or *Hermitian adjoint*), "†," standing for transposition and complex conjugation of the coordinates. More explicitly,

$$\left(x_1,\ldots,x_n\right)^{\dagger} = \begin{pmatrix} \overline{x_1} \\ \vdots \\ \overline{x_n} \end{pmatrix}, \text{ and } \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}^{\dagger} = (\overline{x_1},\ldots,\overline{x_n}). \tag{1.89}$$

Note that, just as for real vector spaces, $\left(\mathbf{x}^T\right)^T = \mathbf{x}$, or, in the bra-ket notation, $\left(|\mathbf{x}\rangle^T\right)^T = |\mathbf{x}\rangle$, so is $\left(\mathbf{x}^{\dagger}\right)^{\dagger} = \mathbf{x}$, or $\left(|\mathbf{x}\rangle^{\dagger}\right)^{\dagger} = |\mathbf{x}\rangle$ for complex vector spaces.

As already mentioned on page 20, Eq. (1.60), for orthonormal bases of complex Hilbert space we can express the dual vector in terms of the original vector by taking the conjugate transpose, and *vice versa*; that is,

$$\langle \mathbf{x}| = (|\mathbf{x}\rangle)^{\dagger}, \text{ and } |\mathbf{x}\rangle = (\langle \mathbf{x}|)^{\dagger}. \tag{1.90}$$

In real vector space, the *dyadic product*, or *tensor product*, or *outer product*

$$\mathbf{E_x} = \mathbf{x} \otimes \mathbf{x}^T = |\mathbf{x}\rangle\langle\mathbf{x}| \equiv \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \left(x_1, x_2, \ldots, x_n\right)$$

$$\tag{1.91}$$

$$= \begin{pmatrix} x_1\left(x_1, x_2, \ldots, x_n\right) \\ x_2\left(x_1, x_2, \ldots, x_n\right) \\ \vdots \\ x_n\left(x_1, x_2, \ldots, x_n\right) \end{pmatrix} = \begin{pmatrix} x_1 x_1 & x_1 x_2 & \cdots & x_1 x_n \\ x_2 x_1 & x_2 x_2 & \cdots & x_2 x_n \\ \vdots & \vdots & \vdots & \vdots \\ x_n x_1 & x_n x_2 & \cdots & x_n x_n \end{pmatrix}$$

is the projection associated with $\mathbf{x}$.

If the vector $\mathbf{x}$ is not normalized, then the associated projection is

$$\mathbf{E_x} \equiv \frac{\mathbf{x} \otimes \mathbf{x}^T}{\langle \mathbf{x} \mid \mathbf{x}\rangle} \equiv \frac{|\mathbf{x}\rangle\langle\mathbf{x}|}{\langle \mathbf{x} \mid \mathbf{x}\rangle} \tag{1.92}$$

This construction is related to $P_{\mathbf{x}}$ on page 13 by $P_{\mathbf{x}}(\mathbf{y}) = \mathbf{E_x}\mathbf{y}$.

For a proof, consider only normalized vectors $\mathbf{x}$, and let $\mathbf{E_x} = \mathbf{x} \otimes \mathbf{x}^T$, then

$$\mathbf{E_x}\mathbf{E_x} = (|\mathbf{x}\rangle\langle\mathbf{x}|)(|\mathbf{x}\rangle\langle\mathbf{x}|) = |\mathbf{x}\rangle \underbrace{\langle\mathbf{x}|\mathbf{x}\rangle}_{=1}\langle\mathbf{x}| = \mathbf{E_x}.$$

More explicitly, by writing out the coordinate tuples, the equivalent proof is

$$\mathbf{E_x}\mathbf{E_x} \equiv (\mathbf{x} \otimes \mathbf{x}^T) \cdot (\mathbf{x} \otimes \mathbf{x}^T)$$

$$\equiv \left[\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} (x_1, x_2, \ldots, x_n) \right] \left[\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} (x_1, x_2, \ldots, x_n) \right] \tag{1.93}$$

$$= \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \underbrace{\left[(x_1, x_2, \ldots, x_n) \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \right]}_{=1} (x_1, x_2, \ldots, x_n) \equiv \mathbf{E_x}.$$

In complex vector space, transposition has to be substituted by the conjugate transposition; that is

$$\mathbf{E_x} = \mathbf{x} \otimes \mathbf{x}^\dagger \equiv |\mathbf{x}\rangle\langle\mathbf{x}| \tag{1.94}$$

For two examples, let $\mathbf{x} = (1,0)^T$ and $\mathbf{y} = (1,-1)^T$; then

$$\mathbf{E_x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1,0) = \begin{pmatrix} 1(1,0) \\ 0(1,0) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix},$$

and

$$\mathbf{E_y} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} (1,-1) = \frac{1}{2} \begin{pmatrix} 1(1,-1) \\ -1(1,-1) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}.$$

Note also that

$$\mathbf{E_x}|\mathbf{y}\rangle \equiv \mathbf{E_x}\mathbf{y} = \langle\mathbf{x}|\mathbf{y}\rangle\mathbf{x}, \equiv \langle\mathbf{x}|\mathbf{y}\rangle|\mathbf{x}\rangle, \tag{1.95}$$

which can be directly proven by insertion.

1.13 Change of basis

Let $\mathfrak{V}$ be an n-dimensional vector space and let $\mathfrak{X} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\mathfrak{Y} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ be two bases of $\mathfrak{V}$.

Take an arbitrary vector $\mathbf{z} \in \mathfrak{V}$. In terms of the two bases $\mathfrak{X}$ and $\mathfrak{Y}$, $\mathbf{z}$ can be written as

$$\mathbf{z} = \sum_{i=1}^{n} x_i \mathbf{e}_i = \sum_{i=1}^{n} y_i \mathbf{f}_i, \tag{1.96}$$

where x_i and y_i stand for the coordinates of the vector $\mathbf{z}$ with respect to the bases $\mathfrak{X}$ and $\mathfrak{Y}$, respectively.

The following questions arise:

(i) What is the relation between the "corresponding" basis vectors $\mathbf{e}_i$ and $\mathbf{f}_j$?

(ii) What is the relation between the coordinates x_i (with respect to the basis $\mathfrak{X}$) and y_j (with respect to the basis $\mathfrak{Y}$) of the vector $\mathbf{z}$ in Eq. (1.96)?

(iii) Suppose one fixes an n-touple $v = \left(v_1, v_2, \ldots, v_n\right)$. What is the relation beween $\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i$ and $\mathbf{w} = \sum_{i=1}^{n} v_i \mathbf{f}_i$?

For proofs and additional information see §46 in

Paul Richard Halmos. *Finite-dimensional Vector Spaces*. Undergraduate Texts in Mathematics. Springer, New York, 1958. ISBN 978-1-4612-6387-6,978-0-387-90093-3. DOI: 10.1007/978-1-4612-6387-6. URL http://doi.org/10.1007/978-1-4612-6387-6

1.13.1   Settlement of change of basis vectors by definition

As an *Ansatz* for answering question (i), recall that, just like any other vector in $\mathfrak{V}$, the new basis vectors $\mathbf{f}_i$ contained in the new basis $\mathfrak{Y}$ can be (uniquely) written as a *linear combination* (in quantum physics called *coherent superposition*) of the basis vectors $\mathbf{e}_i$ contained in the old basis $\mathfrak{X}$. This can be defined *via* a linear transformation $\mathbf{A}$ between the corresponding vectors of the bases $\mathfrak{X}$ and $\mathfrak{Y}$ by

$$\mathbf{f}_i = (\mathbf{e}\mathbf{A})_i, \tag{1.97}$$

for all $i = 1, \ldots, n$. More specifically, let a_{ji} be the matrix of the linear transformation $\mathbf{A}$ in the basis $\mathfrak{X} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$, and let us rewrite (1.97) as a matrix equation

$$\mathbf{f}_i = \sum_{j=1}^{n} a_{ji}\mathbf{e}_j = \sum_{j=1}^{n} (a^T)_{ij}\mathbf{e}_j. \tag{1.98}$$

If $\mathbf{A}$ stands for the matrix whose components (with respect to $\mathfrak{X}$) are a_{ji}, and $\mathbf{A}^T$ stands for the transpose of $\mathbf{A}$ whose components (with respect to $\mathfrak{X}$) are a_{ij}, then

$$\begin{pmatrix} \mathbf{f}_1 \\ \mathbf{f}_2 \\ \vdots \\ \mathbf{f}_n \end{pmatrix} = \mathbf{A}^T \begin{pmatrix} \mathbf{e}_1 \\ \mathbf{e}_2 \\ \vdots \\ \mathbf{e}_n \end{pmatrix}. \tag{1.99}$$

That is, very explicitly,

$$\mathbf{f}_1 = (\mathbf{A}\mathbf{e})_1 = \sum_{i=1}^{n} a_{i1}\mathbf{e}_i = a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2 + \cdots + a_{n1}\mathbf{e}_n,$$

$$\mathbf{f}_2 = (\mathbf{A}\mathbf{e})_2 = \sum_{i=1}^{n} a_{i2}\mathbf{e}_i = a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2 + \cdots + a_{n2}\mathbf{e}_n\mathbf{e}_n,$$

$$\vdots \tag{1.100}$$

$$\mathbf{f}_n = (\mathbf{A}\mathbf{e})_n = \sum_{i=1}^{n} a_{in}\mathbf{e}_i = a_{1v}\mathbf{e}_1 + a_{2n}\mathbf{e}_2 + \cdots + a_{nn}\mathbf{e}_n.$$

This *Ansatz* includes a convention; namely the *order of the indices* of the transformation matix. You may have wondered why we have taken the inconvenience of defining $\mathbf{f}_i$ by $\sum_{j=1}^{n} a_{ji}\mathbf{e}_j$ rather than by $\sum_{j=1}^{n} a_{ij}\mathbf{e}_j$. That is, in Eq. (1.98), why not exchange a_{ji} by a_{ij}, so that the summation index j is "next to" $\mathbf{e}_j$? This is because we want to transform the coordinates according to this "more intuitive" rule, and we cannot have both at the same time. More explicitly, suppose that we want to have

$$y_i = \sum_{j=1}^{n} b_{ij}x_j, \tag{1.101}$$

or, in operator notation and the coordinates as n-tuples,

$$y = \mathbf{B}x. \tag{1.102}$$

Then, by insertion of Eqs. (1.98) and (1.101) into (1.96) we obtain

$$\mathbf{z} = \sum_{i=1}^{n} x_i\mathbf{e}_i = \sum_{i=1}^{n} y_i\mathbf{f}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} b_{ij}x_j \right) \left(\sum_{k=1}^{n} a_{ki}\mathbf{e}_k \right) = \sum_{i,j,k=1}^{n} a_{ki}b_{ij}x_j\mathbf{e}_k, \tag{1.103}$$

If, in contrast, we would have started with $\mathbf{f}_i = \sum_{j=1}^{n} a_{ij}\mathbf{e}_j$ and still pretended to define $y_i = \sum_{j=1}^{n} b_{ij}x_j$, then we would have ended up with $\mathbf{z} = \sum_{i=1}^{n} x_i\mathbf{e}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} b_{ij}x_j \right) \left(\sum_{k=1}^{n} a_{ik}\mathbf{e}_k \right) = \sum_{i,j,k=1}^{n} a_{ik}b_{ij}x_j\mathbf{e}_k = \sum_{i=1}^{n} x_i\mathbf{e}_i$ which, in order to represent $\mathbf{B}$ as the inverse of $\mathbf{A}$, would have forced us to take the transpose of either $\mathbf{B}$ or $\mathbf{A}$ anyway.

which, by comparison, can only be satisfied if $\sum_{i=1}^{n} a_{ki} b_{ij} = \delta_{kj}$. Therefore, $\mathbf{AB} = \mathbb{1}_n$ and $\mathbf{B}$ is the inverse of $\mathbf{A}$. This is quite plausible, since any scale basis change needs to be compensated by a reciprokal or inversely proportional scale change of the coordinates.

- Note that the n equalities (1.100) really represent n^2 linear equations for the n^2 unknowns a_{ij}, $1 \le i, j \le n$, since every pair of basis vectors $\{\mathbf{f}_i, \mathbf{e}_i\}$, $1 \le i \le n$ has n components or coefficients.

- If one knows how the basis vectors $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ of $\mathfrak{X}$ transform, then one knows (by linearity) how all other vectors $\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i$ (represented in this basis) transform; namely $\mathbf{A}(\mathbf{v}) = \sum_{i=1}^{n} v_i (\mathbf{eA})_i$.

- Finally note that, if $\mathfrak{X}$ is an orthonormal basis, then the basis transformation has a diagonal form

$$\mathbf{A} = \sum_{i=1}^{n} \mathbf{f}_i \mathbf{e}_i^{\dagger} \equiv \sum_{i=1}^{n} |\mathbf{f}_i\rangle\langle\mathbf{e}_i| \qquad (1.104)$$

because all the off-diagonal components a_{ij}, $i \ne j$ of $\mathbf{A}$ explicitly written down in Eqs.(1.100) vanish. This can be easily checked by applying $\mathbf{A}$ to the elements $\mathbf{e}_i$ of the basis $\mathfrak{X}$. See also Section 1.22.2 on page 48 for a representation of unitary transformations in terms of basis changes. In quantum mechanics, the temporal evolution is represented by nothing but a change of orthonormal bases in Hilbert space.

1.13.2   Scale change of vector components by contra-variation

Having settled question (i) by the *Ansatz* (1.97), we turn to question (ii) next. Since

$$\mathbf{z} = \sum_{j=1}^{n} y_j \mathbf{f}_j = \sum_{j=1}^{n} y_j (\mathbf{eA})_j = \sum_{j=1}^{n} y_j \sum_{i=1}^{n} a_{ij} \mathbf{e}_i = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} y^j \right) \mathbf{e}_i; \quad (1.105)$$

we obtain by comparison of the coefficients in Eq. (1.96),

$$x_i = \sum_{j=1}^{n} a_{ij} y_j. \qquad (1.106)$$

That is, in terms of the "old" coordinates x^i, the "new" coordinates are

$$\sum_{i=1}^{n} (a^{-1})_{ji} x_i = \sum_{i=1}^{n} (a^{-1})_{ji} \sum_{k=1}^{n} a_{ik} y_k$$

$$= \sum_{k=1}^{n} \left[\sum_{i=1}^{n} (a^{-1})_{ji} a_{ik} \right] y_k = \sum_{k=1}^{n} \delta_k^j y_k = y_j. \qquad (1.107)$$

If we prefer to represent the vector coordinates of $\mathbf{x}$ and $\mathbf{y}$ as n-tuples, then Eqs. (1.106) and (1.107) have an interpretation as matrix multiplication; that is,

$$\mathbf{x} = \mathbf{Ay}, \text{ and } \mathbf{y} = (\mathbf{A}^{-1})\mathbf{x}. \qquad (1.108)$$

Finally, let us answer question (iii) – the relation between $\mathbf{v} = \sum_{i=1}^{n} v_i \mathbf{e}_i$ and $\mathbf{w} = \sum_{i=1}^{n} v_i \mathbf{f}_i$ for any n-touple $v = \left(v_1, v_2, \ldots, v_n \right)$ – by substituting the

transformation (1.98) of the basis vectors in $\mathbf{w}$ and comparing it with $\mathbf{v}$; that is,

$$\mathbf{w} = \sum_{j=1}^{n} v_j \mathbf{f}_j = \sum_{j=1}^{n} v_j \left(\sum_{i=1}^{n} a_{ij} \mathbf{e}_i \right) = \sum_{i=1}^{n} \left(\sum_{j=1}^{n} a_{ij} v_j \right) x_i; \text{ or } \mathbf{w} = \mathbf{A}\mathbf{v}. \quad (1.109)$$

Figure 1.3: Basis change by rotation of $\varphi = \frac{\pi}{4}$ around the origin.

1. For the sake of an example consider a change of basis in the plane $\mathbb{R}^2$ by rotation of an angle $\varphi = \frac{\pi}{4}$ around the origin, depicted in Fig. 1.3. According to Eq. (1.97), we have

$$\begin{aligned} \mathbf{f}_1 &= a_{11}\mathbf{e}_1 + a_{21}\mathbf{e}_2, \\ \mathbf{f}_2 &= a_{12}\mathbf{e}_1 + a_{22}\mathbf{e}_2, \end{aligned} \quad (1.110)$$

which amounts to four linear equations in the four unknowns a_{11}, a_{12}, a_{21}, and a_{22}. By inserting the basis vectors $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{f}_1$, and $\mathbf{f}_2$ one obtains for the rotation matrix with respect to the basis $\mathfrak{X}$

$$\begin{aligned} \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= a_{11} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_{21} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \\ \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix} &= a_{12} \begin{pmatrix} 1 \\ 0 \end{pmatrix} + a_{22} \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \end{aligned} \quad (1.111)$$

the first pair of equations yielding $a_{11} = a_{21} = \frac{1}{\sqrt{2}}$, the second pair of equations yielding $a_{12} = -\frac{1}{\sqrt{2}}$ and $a_{22} = \frac{1}{\sqrt{2}}$. Thus,

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \quad (1.112)$$

As both coordinate systems $\mathfrak{X} = \{\mathbf{e}_1, \mathbf{e}_2\}$ and $\mathfrak{Y} = \{\mathbf{f}_1, \mathbf{f}_2\}$ are orthogonal, we might have just computed the diagonal form (1.104)

$$\begin{aligned} \mathbf{A} &= \frac{1}{\sqrt{2}} \left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1, 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 1 \end{pmatrix} \begin{pmatrix} 0, 1 \end{pmatrix} \right] \\ &= \frac{1}{\sqrt{2}} \left[\begin{pmatrix} 1(1,0) \\ 1(1,0) \end{pmatrix} + \begin{pmatrix} -1(0,1) \\ 1(0,1) \end{pmatrix} \right] \\ &= \frac{1}{\sqrt{2}} \left[\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix} \right] = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \end{aligned} \quad (1.113)$$

Note, however that coordinates transform contra-variantly with $\mathbf{A}^{-1}$.

Likewise, the rotation matrix with respect to the basis $\mathfrak{Y}$ is

$$\mathbf{A}' = \frac{1}{\sqrt{2}} \left[\begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} 1, 1 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \begin{pmatrix} -1, 1 \end{pmatrix} \right] = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \quad (1.114)$$

2. By a similar calculation, taking into account the definition for the sine and cosine functions, one obtains the transformation matrix $\mathbf{A}(\varphi)$ associated with an arbitrary angle φ,

$$\mathbf{A} = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \quad (1.115)$$

The coordinates transform as

$$\mathbf{A}^{-1} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}. \quad (1.116)$$

3. Consider the more general rotation depicted in Fig. 1.4. Again, by
    inserting the basis vectors $\mathbf{e}_1, \mathbf{e}_2, \mathbf{f}_1$, and $\mathbf{f}_2$, one obtains

$$
\begin{aligned}
\frac{1}{2}\begin{pmatrix}\sqrt{3}\\1\end{pmatrix} &= a_{11}\begin{pmatrix}1\\0\end{pmatrix} + a_{21}\begin{pmatrix}0\\1\end{pmatrix},\\
\frac{1}{2}\begin{pmatrix}1\\\sqrt{3}\end{pmatrix} &= a_{12}\begin{pmatrix}1\\0\end{pmatrix} + a_{22}\begin{pmatrix}0\\1\end{pmatrix},
\end{aligned}
\tag{1.117}
$$

yielding $a_{11} = a_{22} = \frac{\sqrt{3}}{2}$, the second pair of equations yielding $a_{12} = a_{21} = \frac{1}{2}$.

Thus,

$$
\mathbf{A} = \begin{pmatrix} a & b \\ b & a \end{pmatrix} = \frac{1}{2}\begin{pmatrix}\sqrt{3} & 1 \\ 1 & \sqrt{3}\end{pmatrix}.
\tag{1.118}
$$

The coordinates transform according to the inverse transformation,
which in this case can be represented by

$$
\mathbf{A}^{-1} = \frac{1}{a^2 - b^2}\begin{pmatrix} a & -b \\ -b & a \end{pmatrix} = \begin{pmatrix}\sqrt{3} & -1 \\ -1 & \sqrt{3}\end{pmatrix}.
\tag{1.119}
$$

$\mathbf{e}_2 = (0,1)^T$

$\mathbf{f}_2 = \frac{1}{2}(1,\sqrt{3})^T$

$\varphi = \frac{\pi}{6}$

$\mathbf{f}_1 = \frac{1}{2}(\sqrt{3},1)^T$

$\varphi = \frac{\pi}{6}$

$\mathbf{e}_1 = (1,0)^T$
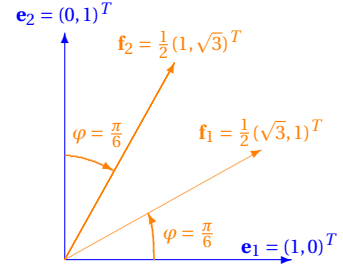
Figure 1.4: More general basis change by
rotation.

1.14   *Mutually unbiased bases*

Two orthonormal bases $\mathfrak{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\mathfrak{B}' = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ are said to be
mutually unbiased if their scalar or inner products are

$$
|\langle \mathbf{e}_i | \mathbf{f}_j \rangle|^2 = \frac{1}{n}
\tag{1.120}
$$

for all $1 \le i, j \le n$. Note without proof – that is, you do not have to be
concerned that you need to understand this from what has been said
so far – that "the elements of two or more mutually unbiased bases are
mutually maximally apart."

In physics, one seeks maximal sets of orthogonal bases who are max-
imally apart [18] . Such maximal sets of bases are used in quantum infor-
mation theory to assure maximal performance of certain protocols used
in quantum cryptography, or for the production of quantum random
sequences by beam splitters. They are essential for the practical exploita-
tions of quantum complementary properties and resources.

Schwinger presented an algorithm (see [19] for a proof) to construct a
new mutually unbiased basis $\mathfrak{B}$ from an existing orthogonal one. The
proof idea is to create a new basis "inbetween" the old basis vectors. by
the following construction steps:

(i)   take the existing orthogonal basis and permute all of its elements by
      "shift-permuting" its elements; that is, by changing the basis vectors
      according to their enumeration $i \to i + 1$ for $i = 1, \ldots, n - 1$, and $n \to 1$;
      or any other nontrivial (i.e., do not consider identity for any basis
      element) permutation;

(ii) consider the *(unitary) transformation* (cf. Sections 1.13 and 1.22.2)
      corresponding to the basis change from the old basis to the new,
      "permutated" basis;

(iii) finally, consider the (orthonormal) *eigenvectors* of this (unitary; cf. page 46) transformation associated with the basis change. These eigenvectors are the elements of a new bases $\mathfrak{B}'$. Together with $\mathfrak{B}$ these two bases – that is, $\mathfrak{B}$ and $\mathfrak{B}'$ – are mutually unbiased.

Consider, for example, the real plane $\mathbb{R}^2$, and the basis

$$\mathfrak{B} = \{\mathbf{e}_1, \mathbf{e}_2\} \equiv \{|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle\} \equiv \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$$

The shift-permutation [step (i)] brings $\mathfrak{B}$ to a new, "shift-permuted" basis $\mathfrak{S}$; that is,

$$\{\mathbf{e}_1, \mathbf{e}_2\} \mapsto \mathfrak{S} = \{\mathbf{f}_1 = \mathbf{e}_2, \mathbf{f}_1 = \mathbf{e}_1\} \equiv \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}.$$

The (unitary) basis transformation [step (ii)] between $\mathfrak{B}$ and $\mathfrak{S}$ can be constructed by a diagonal sum

$$\begin{aligned}
\mathbf{U} &= \mathbf{f}_1 \mathbf{e}_1^\dagger + \mathbf{f}_2 \mathbf{e}_2^\dagger = \mathbf{e}_2 \mathbf{e}_1^\dagger + \mathbf{e}_1 \mathbf{e}_2^\dagger \\
&\equiv |\mathbf{f}_1\rangle\langle\mathbf{e}_1| + |\mathbf{f}_2\rangle\langle\mathbf{e}_2| = |\mathbf{e}_2\rangle\langle\mathbf{e}_1| + |\mathbf{e}_1\rangle\langle\mathbf{e}_2| \\
&\equiv \begin{pmatrix} 0 \\ 1 \end{pmatrix}(1,0) + \begin{pmatrix} 1 \\ 0 \end{pmatrix}(0,1) \\
&\equiv \begin{pmatrix} 0(1,0) \\ 1(1,0) \end{pmatrix} + \begin{pmatrix} 1(0,1) \\ 0(0,1) \end{pmatrix} \\
&\equiv \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.
\end{aligned} \tag{1.121}$$

The set of eigenvectors [step (iii)] of this (unitary) basis transformation $\mathbf{U}$ forms a new basis

$$\begin{aligned}
\mathfrak{B}' &= \{\frac{1}{\sqrt{2}}(\mathbf{f}_1 - \mathbf{e}_1), \frac{1}{\sqrt{2}}(\mathbf{f}_2 + \mathbf{e}_2)\} \\
&= \{\frac{1}{\sqrt{2}}(|\mathbf{f}_1\rangle - |\mathbf{e}_1\rangle), \frac{1}{\sqrt{2}}(|\mathbf{f}_2\rangle + |\mathbf{e}_2\rangle)\} \\
&= \{\frac{1}{\sqrt{2}}(|\mathbf{e}_2\rangle - |\mathbf{e}_1\rangle), \frac{1}{\sqrt{2}}(|\mathbf{e}_1\rangle + |\mathbf{e}_2\rangle)\} \\
&\equiv \left\{ \frac{1}{\sqrt{2}}\left\{ \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\} \right\}.
\end{aligned} \tag{1.122}$$

For a proof of mutually unbiasedness, just form the four inner products of one vector in $\mathfrak{B}$ times one vector in $\mathfrak{B}'$, respectively.

In three-dimensional complex vector space $\mathbb{C}^3$, a similar construction from the Cartesian standard basis $\mathfrak{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\} \equiv \{(1,0,0)^T, (0,1,0)^T, (0,0,1)^T\}$ yields

$$\mathfrak{B}' \equiv \frac{1}{\sqrt{3}} \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2}[\sqrt{3}i - 1] \\ \frac{1}{2}[-\sqrt{3}i - 1] \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{1}{2}[-\sqrt{3}i - 1] \\ \frac{1}{2}[\sqrt{3}i - 1] \\ 1 \end{pmatrix} \right\}. \tag{1.123}$$

So far, nobody has discovered a systematic way to derive and construct a *complete* or *maximal* set of mutually unbiased bases in arbitrary dimensions; in particular, *how many* bases are there in such sets.

1.15 Completeness or resolution of the identity operator in terms of base vectors

The identity $\mathbb{I}_n$ in an n-dimensional vector space $\mathfrak{V}$ can be represented in terms of the sum over all outer (by another naming tensor or dyadic) products of all vectors of an arbitrary orthonormal basis $\mathfrak{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\} \equiv \{|\mathbf{e}_1\rangle, \ldots, |\mathbf{e}_n\rangle\}$; that is,

$$\mathbb{I}_n = \sum_{i=1}^{n} |\mathbf{e}_i\rangle\langle\mathbf{e}_i| \equiv \sum_{i=1}^{n} \mathbf{e}_i \mathbf{e}_i^{\dagger}. \tag{1.124}$$

This is sometimes also referred to as *completeness*.

For a proof, consider an arbitrary vector $|\mathbf{x}\rangle \in \mathfrak{V}$. Then,

$$\mathbb{I}_n|\mathbf{x}\rangle = \left(\sum_{i=1}^{n} |\mathbf{e}_i\rangle\langle\mathbf{e}_i|\right)|\mathbf{x}\rangle = \sum_{i=1}^{n} |\mathbf{e}_i\rangle\langle\mathbf{e}_i|\mathbf{x}\rangle = \sum_{i=1}^{n} |\mathbf{e}_i\rangle x_i = |\mathbf{x}\rangle. \tag{1.125}$$

Consider, for example, the basis $\mathfrak{B} = \{|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle\} \equiv \{(1,0)^T, (0,1)^T\}$. Then the two-dimensional resolution of the identity operator $\mathbb{I}_2$ can be written as

$$\mathbb{I}_2 = |\mathbf{e}_1\rangle\langle\mathbf{e}_1| + |\mathbf{e}_2\rangle\langle\mathbf{e}_2|$$
$$= (1,0)^T(1,0) + (0,1)^T(0,1) = \begin{pmatrix} 1(1,0) \\ 0(1,0) \end{pmatrix} + \begin{pmatrix} 0(0,1) \\ 1(0,1) \end{pmatrix}$$
$$= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.126}$$

Consider, for another example, the basis $\mathfrak{B}' \equiv \{\frac{1}{\sqrt{2}}(-1,1)^T, \frac{1}{\sqrt{2}}(1,1)^T\}$. Then the two-dimensional resolution of the identity operator $\mathbb{I}_2$ can be written as

$$\mathbb{I}_2 = \frac{1}{\sqrt{2}}(-1,1)^T \frac{1}{\sqrt{2}}(-1,1) + \frac{1}{\sqrt{2}}(1,1)^T \frac{1}{\sqrt{2}}(1,1)$$
$$= \frac{1}{2}\begin{pmatrix} -1(-1,1) \\ 1(-1,1) \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1(1,1) \\ 1(1,1) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{1.127}$$

1.16 Rank

The (column or row) *rank*, $\rho(\mathbf{A})$, or rk$(\mathbf{A})$, of a linear transformation $\mathbf{A}$ in an n-dimensional vector space $\mathfrak{V}$ is the maximum number of linearly independent (column or, equivalently, row) vectors of the associated n-by-n square matrix A, represented by its entries a_{ij}.

This definition can be generalized to arbitrary m-by-n matrices A, represented by its entries a_{ij}. Then, the row and column ranks of A are identical; that is,

$$\text{row rk}(A) = \text{column rk}(A) = \text{rk}(A). \tag{1.128}$$

For a proof, consider Mackiw's argument [20]. First we show that row rk$(A) \leq$ column rk(A) for any real (a generalization to complex vector space requires some adjustments) m-by-n matrix A. Let the vectors

[20] George Mackiw. A note on the equality of the column and row rank of a matrix. *Mathematics Magazine*, 68(4):pp. 285–286, 1995. ISSN 0025570X. URL http://www.jstor.org/stable/2690576

$\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r\}$ with $\mathbf{e}_i \in \mathbb{R}^n$, $1 \le i \le r$, be a basis spanning the *row space* of A; that is, all vectors that can be obtained by a linear combination of the m row vectors

$$\begin{pmatrix} (a_{11}, a_{12}, \ldots, a_{1n}) \\ (a_{21}, a_{22}, \ldots, a_{2n}) \\ \vdots \\ (a_{m1}, a_{n2}, \ldots, a_{mn}) \end{pmatrix}$$

of A can also be obtained as a linear combination of $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$. Note that $r \le m$.

Now form the *column vectors* $A\mathbf{e}_i^T$ for $1 \le i \le r$, that is, $A\mathbf{e}_1^T, A\mathbf{e}_2^T, \ldots, A\mathbf{e}_r^T$ *via* the usual rules of matrix multiplication. Let us prove that these resulting column vectors $A\mathbf{e}_i^T$ are linearly independent.

Suppose they were not (proof by contradiction). Then, for some scalars $c_1, c_2, \ldots, c_r \in \mathbb{R}$,

$$c_1 A\mathbf{e}_1^T + c_2 A\mathbf{e}_2^T + \ldots + c_r A\mathbf{e}_r^T = A\left(c_1\mathbf{e}_1^T + c_2\mathbf{e}_2^T + \ldots + c_r\mathbf{e}_r^T\right) = 0$$

without all c_i's vanishing.

That is, $\mathbf{v} = c_1\mathbf{e}_1^T + c_2\mathbf{e}_2^T + \ldots + c_r\mathbf{e}_r^T$, must be in the *null space* of A defined by all vectors $\mathbf{x}$ with $A\mathbf{x} = \mathbf{0}$, and $A(\mathbf{v}) = \mathbf{0}$. (In this case the inner (Euclidean) product of $\mathbf{x}$ with all the rows of A must vanish.) But since the $\mathbf{e}_i$'s form also a basis of the row vectors, $\mathbf{v}^T$ is also some vector in the row space of A. The linear independence of the basis elements $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_r$ of the row space of A guarantees that all the coefficients c_i have to vanish; that is, $c_1 = c_2 = \cdots = c_r = 0$.

At the same time, as for every vector $\mathbf{x} \in \mathbb{R}^n$, $A\mathbf{x}$ is a linear combination of the column vectors

$$\left(\begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{pmatrix}, \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{pmatrix}, \cdots, \begin{pmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{pmatrix} \right),$$

the r linear independent vectors $A\mathbf{e}_1^T, A\mathbf{e}_2^T, \ldots, A\mathbf{e}_r^T$ are all linear combinations of the column vectors of A. Thus, they are in the column space of A. Hence, $r \le \text{column rk}(A)$. And, as $r = \text{row rk}(A)$, we obtain row $\text{rk}(A) \le \text{column rk}(A)$.

By considering the transposed matrix A^T, and by an analogous argument we obtain that row $\text{rk}(A^T) \le \text{column rk}(A^T)$. But row $\text{rk}(A^T) = \text{column rk}(A)$ and column $\text{rk}(A^T) = \text{row rk}(A)$, and thus row $\text{rk}(A^T) = \text{column rk}(A) \le \text{column rk}(A^T) = \text{row rk}(A)$. Finally, by considering both estimates row $\text{rk}(A) \le \text{column rk}(A)$ as well as column $\text{rk}(A) \le \text{row rk}(A)$, we obtain that row $\text{rk}(A) = \text{column rk}(A)$.

1.17 Determinant

1.17.1 Definition

In what follows, the *determinant* of a matrix A will be denoted by $\det A$ or, equivalently, by $|A|$.

Suppose $A = a_{ij}$ is the n-by-n square matrix representation of a linear transformation **A** in an n-dimensional vector space $\mathfrak{V}$. We shall define its *determinant* in two equivalent ways.

The *Leibniz formula* defines the determinant of the n-by-n square matrix $A = a_{ij}$ by

$$\det A = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^{n} a_{\sigma(i), j}, \qquad (1.129)$$

where "sgn" represents the *sign function* of permutations σ in the permutation group S_n on n elements $\{1, 2, \ldots, n\}$, which returns -1 and $+1$ for odd and even permutations, respectively. $\sigma(i)$ stands for the element in position i of $\{1, 2, \ldots, n\}$ *after* permutation σ.

An equivalent (no proof is given here) definition

$$\det A = \varepsilon_{i_1 i_2 \cdots i_n} a_{1 i_1} a_{2 i_2} \cdots a_{n i_n}, \qquad (1.130)$$

makes use of the totally antisymmetric Levi-Civita symbol (2.99) on page 96, and makes use of the Einstein summation convention.

The second, *Laplace formula* definition of the determinant is recursive and expands the determinant in cofactors. It is also called *Laplace expansion*, or *cofactor expansion* . First, a *minor M_{ij}* of an n-by-n square matrix A is defined to be the determinant of the $(n-1) \times (n-1)$ submatrix that remains after the entire ith row and jth column have been deleted from A.

A *cofactor A_{ij}* of an n-by-n square matrix A is defined in terms of its associated minor by

$$A_{ij} = (-1)^{i+j} M_{ij}. \qquad (1.131)$$

The *determinant* of a square matrix A, denoted by $\det A$ or $|A|$, is a scalar rekursively defined by

$$\det A = \sum_{j=1}^{n} a_{ij} A_{ij} = \sum_{i=1}^{n} a_{ij} A_{ij} \qquad (1.132)$$

for any i (row expansion) or j (column expansion), with $i, j = 1, \ldots, n$. For 1×1 matrices (i.e., scalars), $\det A = a_{11}$.

1.17.2   Properties

The following properties of determinants are mentioned (almost) without proof:

(i)   If A and B are square matrices of the same order, then $\det AB = (\det A)(\det B)$.

(ii)   If either two rows or two columns are exchanged, then the determinant is multiplied by a factor "-1."

(iii)   The determinant of the transposed matrix is equal to the determinant of the original matrix; that is, $\det(A^T) = \det A$.

(iv)   The determinant $\det A$ of a matrix A is nonzero if and only if A is invertible. In particular, if A is not invertible, $\det A = 0$. If A has an inverse matrix A^{-1}, then $\det(A^{-1}) = (\det A)^{-1}$.

This is a very important property which we shall use in Eq. (1.202) on page 54 for the determination of nontrivial eigenvalues λ (including the associated eigenvectors)   of a matrix A by solving the secular equation $\det(A - \lambda\mathbb{I}) = 0$.

(v) Multiplication of any row or column with a factor α results in a determinant which is α times the original determinant. Consequently, multiplication of an $n \times n$ matrix with a scalar α results in a determinant which is α^n times the original determinant.

(vi) The determinant of a identity matrix is one; that is, $\det \mathbb{I}_n = 1$. Likewise, the determinat of a diagonal matrix is just the product of the diagonal entries; that is, $\det[\mathrm{diag}(\lambda_1, \ldots, \lambda_n)] = \lambda_1 \cdots \lambda_n$.

(vii) The determinant is not changed if a multiple of an existing row is added to another row.

This can be easily demonstrated by considering the Leibniz formula: suppose a multiple α of the j'th column is added to the k'th column since

$$
\begin{aligned}
\varepsilon_{i_1 i_2 \cdots i_j \cdots i_k \cdots i_n} & a_{1 i_1} a_{2 i_2} \cdots a_{j i_j} \cdots (a_{k i_k} + \alpha\, a_{j i_k}) \cdots a_{n i_n} \\
& = \varepsilon_{i_1 i_2 \cdots i_j \cdots i_k \cdots i_n} a_{1 i_1} a_{2 i_2} \cdots a_{j i_j} \cdots a_{k i_k} \cdots a_{n i_n} + \\
& \quad \alpha \varepsilon_{i_1 i_2 \cdots i_j \cdots i_k \cdots i_n} a_{1 i_1} a_{2 i_2} \cdots a_{j i_j} \cdots a_{j i_k} \cdots a_{n i_n}.
\end{aligned}
\tag{1.133}
$$

The second summation term vanishes, since $a_{j i_j} a_{j i_k} = a_{j i_k} a_{j i_j}$ is totally symmetric in the indices i_j and i_k, and the Levi-Civita symbol $\varepsilon_{i_1 i_2 \cdots i_j \cdots i_k \cdots i_n}$.

(viii) The absolute value of the determinant of a square matrix $A = (\mathbf{e}_1, \ldots \mathbf{e}_n)$ formed by (not necessarily orthogonal) row (or column) vectors of a basis $\mathfrak{B} = \{\mathbf{e}_1, \ldots \mathbf{e}_n\}$ is equal to the *volume* of the parallelepiped $\{\mathbf{x} \mid \mathbf{x} = \sum_{i=1}^{n} t_i \mathbf{e}_i,\ 0 \le t_i \le 1,\ 0 \le i \le n\}$ formed by those vectors.

This can be demonstrated by supposing that the square matrix A consists of all the n row (column) vectors of an orthogonal basis of dimension n. Then $AA^T = A^T A$ is a diagonal matrix which just contains the square of the length of all the basis vectors forming a perpendicular parallelepiped which is just an n dimensional box. Therefore the volume is just the positive square root of $\det(AA^T) = (\det A)(\det A^T) = (\det A)(\det A^T) = (\det A)^2$.

see, for instance, Section 4.3 of Strang's account

Gilbert Strang. *Introduction to linear algebra*. Wellesley-Cambridge Press, Wellesley, MA, USA, fourth edition, 2009. ISBN 0-9802327-1-6. URL http://math.mit.edu/linearalgebra/

For any nonorthogonal basis all we need to employ is a Gram-Schmidt process to obtain a (perpendicular) box of equal volume to the original parallelepiped formed by the nonorthogonal basis vectors – any volume that is cut is compensated by adding the same amount to the new volume. Note that the Gram-Schmidt process operates by adding (subtracting) the projections of already existing orthogonalized vectors from the old basis vectors (to render these sums orthogonal to the existing vectors of the new orthogonal basis); a process which does not change the determinant.

This result can be used for changing the differential volume element in integrals *via* the Jacobian matrix J (2.20), as

$$dx_1' \, dx_2' \cdots dx_n' = |det J| dx_1 \, dx_2 \cdots dx_n$$

$$= \sqrt{\left[\det\left(\frac{dx_i'}{dx_j} \right) \right]^2} \, dx_1 \, dx_2 \cdots dx_n. \tag{1.134}$$

(ix) The *sign* of a determinant of a matrix formed by the row (column) vectors of a basis indicates the orientation of that basis.

1.18   Trace

1.18.1   Definition

The *trace* of an n-by-n square matrix $A = a_{ij}$, denoted by $\text{Tr} A$, is a scalar defined to be the sum of the elements on the main diagonal (the diagonal from the upper left to the lower right) of A; that is (also in Dirac's bra and ket notation),

$$\text{Tr} \, A = a_{11} + a_{22} + \cdots + a_{nn} = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \langle i | A | i \rangle. \tag{1.135}$$

Traces are noninvertible (irreversible) almost by definition: for $n \geq 2$ and for arbitrary values $a_{ii} \in \mathbb{R}, \mathbb{C}$, there are "many" ways to obtain the same value of $\sum_{i=1}^{n} a_{ii}$.

Traces are linear functionals, because, for two arbitrary matrices A, B and two arbitrary scalars α, β,

$$\text{Tr} \, (\alpha A + \beta B) = \sum_{i=1}^{n} (\alpha a_{ii} + \beta b_{ii}) = \alpha \sum_{i=1}^{n} a_{ii} + \beta \sum_{i=1}^{n} b_{ii} = \alpha \text{Tr} \, (A) + \beta \text{Tr} \, (B). \tag{1.136}$$

Traces can be realized *via* some arbitrary orthonormal basis $\mathfrak{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ by "sandwiching" an operator $\mathbf{A}$ between all basis elements – thereby effectively taking the diagonal components of $\mathbf{A}$ with respect to the basis $\mathfrak{B}$ – and summing over all these scalar compontents; that is, with definition (1.82),

$$\text{Tr} \, \mathbf{A} = \sum_{i=1}^{n} \langle \mathbf{e}_i | \mathbf{A} | \mathbf{e}_i \rangle = \sum_{i=1}^{n} \langle \mathbf{e}_i | \mathbf{A} \mathbf{e}_i \rangle$$

$$= \sum_{i=1}^{n} \sum_{l=1}^{n} \langle \mathbf{e}_i | \alpha_{li} \mathbf{e}_l \rangle = \sum_{i=1}^{n} \sum_{l=1}^{n} \alpha_{li} \langle \mathbf{e}_i | \mathbf{e}_l \rangle \tag{1.137}$$

$$= \sum_{i=1}^{n} \sum_{l=1}^{n} \alpha_{li} \delta_{il} = \sum_{i=1}^{n} \alpha_{ii}.$$

This representation is particularly useful in quantum mechanics.

Suppose an operator is defined *via* two arbitrary vectors $|\mathbf{u}\rangle$ and $|\mathbf{v}\rangle$ by $\mathbf{A} = |\mathbf{u}\rangle \langle \mathbf{v}|$. Then its trace can be rewritten as the scalar product of the two vectors (in exchanged order); that is, for some arbitrary orthonormal basis $\mathfrak{B} = \{\langle \mathbf{e}_1 \rangle, \ldots, \langle \mathbf{e}_n \rangle\}$

$$\text{Tr} \, \mathbf{A} = \sum_{i=1}^{n} \langle \mathbf{e}_i | \mathbf{A} | \mathbf{e}_i \rangle = \sum_{i=1}^{n} \langle \mathbf{e}_i | \mathbf{u} \rangle \langle \mathbf{v} | \mathbf{e}_i \rangle$$

$$= \sum_{i=1}^{n} \langle \mathbf{v} | \mathbf{e}_i \rangle \langle \mathbf{e}_i | \mathbf{u} \rangle = \langle \mathbf{v} | \mathbb{I}_n | \mathbf{u} \rangle = \langle \mathbf{v} | \mathbf{u} \rangle. \tag{1.138}$$

In general traces represent noninvertible (irreversible) many-to-one functionals since the same trace value can be obtained from different inputs. More explicitly, consider two nonidentical vectors $|\mathbf{u}\rangle \neq |\mathbf{v}\rangle$ in real Hilbert space. In this case,

$$\text{Tr } \mathbf{A} = \text{Tr } |\mathbf{u}\rangle\langle\mathbf{v}| = \langle\mathbf{v}|\mathbf{u}\rangle = \langle\mathbf{u}|\mathbf{v}\rangle = \text{Tr } |\mathbf{v}\rangle\langle\mathbf{u}| = \text{Tr } \mathbf{A}^T \qquad (1.139)$$

This example shows that the traces of two matrices such as Tr $\mathbf{A}$ and Tr $\mathbf{A}^T$ can be identical although the argument matrices $\mathbf{A} = |\mathbf{u}\rangle\langle\mathbf{v}|$ and $\mathbf{A}^T = |\mathbf{v}\rangle\langle\mathbf{u}|$ need not be.

1.18.2 Properties

The following properties of traces are mentioned without proof:

(i) $\text{Tr}(A + B) = \text{Tr}A + \text{Tr}B$;

(ii) $\text{Tr}(\alpha A) = \alpha \text{Tr}A$, with $\alpha \in \mathbb{C}$;

(iii) $\text{Tr}(AB) = \text{Tr}(BA)$, hence the trace of the commutator vanishes; that is, $\text{Tr}([A, B]) = 0$;

(iv) $\text{Tr}A = \text{Tr}A^T$;

(v) $\text{Tr}(A \otimes B) = (\text{Tr}A)(\text{Tr}B)$;

(vi) the trace is the sum of the eigenvalues of a *normal operator* (cf. page 57);

(vii) $\det(e^A) = e^{\text{Tr}A}$;

(viii) the trace is the derivative of the determinant at the identity;

(ix) the complex conjugate of the trace of an operator is equal to the trace of its adjoint (cf. page 43); that is $\overline{(\text{Tr}A)} = \text{Tr}(A^\dagger)$;

(x)   the trace is invariant under rotations of the basis as well as under cyclic permutations.

(xi) the trace of an $n \times n$ matrix A for which $AA = \alpha A$ for some $\alpha \in \mathbb{R}$ is $\text{Tr}A = \alpha\text{rank}(A)$, where rank is the rank of A defined on page 36. Consequently, the trace of an idempotent (with $\alpha = 1$) operator – that is, a projection – is equal to its rank; and, in particular, the trace of a one-dimensional projection is one.

(xii) Only commutators have trace zero.

A *trace class* operator is a compact operator for which a trace is finite and independent of the choice of basis.

1.18.3 Partial trace

The quantum mechanics of multi-particle (multipartite) systems allows for configurations – actually rather processes – that can be informally described as "beam dump experiments;" in which we start out with entangled states (such as the Bell states on page 25) which carry information

about *joint properties of the constituent quanta* and *choose to disregard* one quantum state entirely; that is, we pretend not to care about, and "look the other way" with regards to the (possible) outcomes of a measurement on this particle. In this case, we have to *trace out* that particle; and as a result we obtain a *reduced state* without this particle we do not care about.

Formally the partial trace with respect to the first particle maps the general density matrix $\rho_{12} = \sum_{i_1 j_1 i_2 j_2} \rho_{i_1 j_1 i_2 j_2} |i_1\rangle\langle j_1| \otimes |i_2\rangle\langle j_2|$ on a composite Hilbert space $\mathfrak{H}_1 \otimes \mathfrak{H}_2$ to a density matrix on the Hilbert space $\mathfrak{H}_2$ of the second particly by

$$
\begin{aligned}
\mathrm{Tr}_1 \rho_{12} = \mathrm{Tr}_1 &\left(\sum_{i_1 j_1 i_2 j_2} \rho_{i_1 j_1 i_2 j_2} |i_1\rangle\langle j_1| \otimes |i_2\rangle\langle j_2| \right) = \\
&= \sum_{k_1} \left\langle \mathbf{e}_{k_1} \left| \left(\sum_{i_1 j_1 i_2 j_2} \rho_{i_1 j_1 i_2 j_2} |i_1\rangle\langle j_1| \otimes |i_2\rangle\langle j_2| \right) \right| \mathbf{e}_{k_1} \right\rangle = \\
&= \sum_{i_1 j_1 i_2 j_2} \rho_{i_1 j_1 i_2 j_2} \left(\sum_{k_1} \langle \mathbf{e}_{k_1} | i_1\rangle\langle j_1 | \mathbf{e}_{k_1}\rangle \right) |i_2\rangle\langle j_2| = \quad (1.140) \\
&= \sum_{i_1 j_1 i_2 j_2} \rho_{i_1 j_1 i_2 j_2} \left(\sum_{k_1} \langle j_1 | \mathbf{e}_{k_1}\rangle\langle \mathbf{e}_{k_1} | i_1\rangle \right) |i_2\rangle\langle j_2| = \\
= \sum_{i_1 j_1 i_2 j_2} \rho_{i_1 j_1 i_2 j_2} \langle j_1 | \mathbb{1} | i_1\rangle |i_2\rangle\langle j_2| = &\sum_{i_1 j_1 i_2 j_2} \rho_{i_1 j_1 i_2 j_2} \langle j_1 | i_1\rangle |i_2\rangle\langle j_2|.
\end{aligned}
$$

Suppose further that the vectors $|i_1\rangle$ and $|j_1\rangle$ by associated with the first particle belong to an orthonormal basis. Then $\langle j_1 | i_1\rangle = \delta_{i_1 j_1}$ and (1.140) reduces to

$$
\mathrm{Tr}_1 \rho_{12} = \sum_{i_1 i_2 j_2} \rho_{i_1 i_1 i_2 j_2} |i_2\rangle\langle j_2|. \quad (1.141)
$$

The partial trace in general corresponds to a noninvertible map corresponding to an irreversible process; that is, it is an m-to-n with $m > n$, or a many-to-one mapping: $\rho_{11 i_2 j_2} = 1, \rho_{12 i_2 j_2} = \rho_{21 i_2 j_2} = \rho_{22 i_2 j_2} = 0$ and $\rho_{22 i_2 j_2} = 1, \rho_{12 i_2 j_2} = \rho_{21 i_2 j_2} = \rho_{11 i_2 j_2} = 0$ are mapped into the same $\sum_{i_1} \rho_{i_1 i_1 i_2 j_2}$. This can be expected, as information about the first particle is "erased."

For an explicit example's sake, consider the Bell state $|\Psi^-\rangle$ defined in Eq. (1.72). Suppose we do not care about the state of the first particle, then we may ask what kind of reduced state results from this pretension. Then the partial trace is just the trace over the first particle; that is, with subscripts referring to the particle number,

$$
\begin{aligned}
\mathrm{Tr}_1 &|\Psi^-\rangle\langle\Psi^-| \\
&= \sum_{i_1=0}^{1} \langle i_1 | \Psi^-\rangle\langle\Psi^- | i_1\rangle \\
&= \langle 0_1 | \Psi^-\rangle\langle\Psi^- | 0_1\rangle + \langle 1_1 | \Psi^-\rangle\langle\Psi^- | 1_1\rangle \\
&= \langle 0_1 | \frac{1}{\sqrt{2}} (|0_1 1_2\rangle - |1_1 0_2\rangle) \frac{1}{\sqrt{2}} (\langle 0_1 1_2| - \langle 1_1 0_2|) |0_1\rangle \\
&\quad + \langle 1_1 | \frac{1}{\sqrt{2}} (|0_1 1_2\rangle - |1_1 0_2\rangle) \frac{1}{\sqrt{2}} (\langle 0_1 1_2| - \langle 1_1 0_2|) |1_1\rangle \\
&= \frac{1}{2} (|1_2\rangle\langle 1_2| + |0_2\rangle\langle 0_2|).
\end{aligned}
\quad (1.142)
$$

The same is true for all elements of the Bell basis.

Be careful here to make the experiment in such a way that in no way you could know the state of the first particle. You may actually think about this as a measurement of the state of the first particle by a degenerate observable with only a single, nondiscriminating measurement outcome.

The resulting state is a *mixed state* defined by the property that its trace is equal to one, but the trace of its square is smaller than one; in this case the trace is $\frac{1}{2}$, because

$$\mathrm{Tr}_2 \frac{1}{2} \left(|1_2\rangle\langle 1_2| + |0_2\rangle\langle 0_2| \right)$$

$$= \frac{1}{2} \langle 0_2| \left(|1_2\rangle\langle 1_2| + |0_2\rangle\langle 0_2| \right) |0_2\rangle + \frac{1}{2} \langle 1_2| \left(|1_2\rangle\langle 1_2| + |0_2\rangle\langle 0_2| \right) |1_2\rangle \qquad (1.143)$$

$$= \frac{1}{2} + \frac{1}{2} = 1;$$

but

$$\mathrm{Tr}_2 \left[\frac{1}{2} \left(|1_2\rangle\langle 1_2| + |0_2\rangle\langle 0_2| \right) \frac{1}{2} \left(|1_2\rangle\langle 1_2| + |0_2\rangle\langle 0_2| \right) \right]$$

$$= \mathrm{Tr}_2 \frac{1}{4} \left(|1_2\rangle\langle 1_2| + |0_2\rangle\langle 0_2| \right) = \frac{1}{2}. \qquad (1.144)$$

This *mixed state* is a 50:50 mixture of the pure particle states $|0_2\rangle$ and $|1_2\rangle$, respectively. Note that this is different from a coherent superposition $|0_2\rangle + |1_2\rangle$ of the pure particle states $|0_2\rangle$ and $|1_2\rangle$, respectively – also formalizing a 50:50 mixture with respect to measurements of property 0 *versus* 1, respectively.

In quantum mechanics, the "inverse" of the partial trace is called *purification*: it is the creation of a pure state from a mixed one, associated with an "enlargement" of Hilbert space (more dimensions). This cannot be done in a unique way (see Section 1.31 below). Some people – members of the "church of the larger Hilbert space" – believe that mixed states are epistemic (that is, associated with our own personal ignorance rather than with any ontic, microphysical property), and are always part of an, albeit unknown, pure state in a larger Hilbert space.

For additional information see page 110, Sect. 2.5 in
  Michael A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2010. 10th Anniversary Edition

1.19   Adjoint or dual transformation

1.19.1   Definition

Let $\mathfrak{V}$ be a vector space and let $\mathbf{y}$ be any element of its dual space $\mathfrak{V}^*$. For any linear transformation $\mathbf{A}$, consider the bilinear functional $\mathbf{y}'(\mathbf{x}) \equiv [\mathbf{x}, \mathbf{y}'] = [\mathbf{A}\mathbf{x}, \mathbf{y}] \equiv \mathbf{y}(\mathbf{A}\mathbf{x})$. Let the *adjoint* (or *dual*) transformation $\mathbf{A}^*$ be defined by $\mathbf{y}'(\mathbf{x}) = \mathbf{A}^*\mathbf{y}(\mathbf{x})$ with

Here $[\cdot, \cdot]$ is the bilinear functional, not the commutator.

$$\mathbf{A}^* \mathbf{y}(\mathbf{x}) \equiv [\mathbf{x}, \mathbf{A}^* \mathbf{y}] = [\mathbf{A}\mathbf{x}, \mathbf{y}] \equiv \mathbf{y}(\mathbf{A}\mathbf{x}). \qquad (1.145)$$

1.19.2   Adjoint matrix notation

In matrix notation and in complex vector space with the dot product, note that there is a correspondence with the inner product (cf. page 20) so that, for all $\mathbf{z} \in \mathfrak{V}$ and for all $\mathbf{x} \in \mathfrak{V}$, there exist a unique $\mathbf{y} \in \mathfrak{V}$ with

Recall that, for $\alpha, \beta \in \mathbb{C}$, $\overline{(\alpha\beta)} = \overline{\alpha}\overline{\beta}$, and $\overline{\left[\overline{(\alpha)} \right]} = \alpha$, and that the Euclidean scalar product is assumed to be linear in its first argument and antilinear in its second argument.

$$[\mathbf{A}\mathbf{x}, \mathbf{z}] = \langle \mathbf{A}\mathbf{x} \mid \mathbf{y} \rangle = \overline{\langle \mathbf{y} \mid \mathbf{A}\mathbf{x} \rangle} =$$

$$= \overline{\left[y_i \overline{(A_{ij} x_j)} \right]} = \overline{\left[y_i \overline{A_{ij}} \, \overline{x_j} \right]} = \overline{y_i} \overline{\overline{A_{ij}}} \, \overline{\overline{x_j}} = \qquad (1.146)$$

$$= \overline{y_i} A_{ij} x_j = x_j A_{ij} \overline{y_i} = x_j (A^T)_{ji} \overline{y_i} = \mathbf{x} \mathbf{A}^T \overline{\mathbf{y}},$$

and another unique vector $\mathbf{y}'$ obtained from $\mathbf{y}$ by some linear operator $\mathbf{A}^*$ such that $\mathbf{y}' = \mathbf{A}^*\mathbf{y}$ with

$$
\begin{aligned}
[\mathbf{x}, \mathbf{A}^*\mathbf{z}] = \langle \mathbf{x} \mid \mathbf{y}' \rangle = \langle \mathbf{x} \mid \mathbf{A}^*\mathbf{y} \rangle = \\
= x_i \overline{\left(A_{ij}^* y_j \right)} = \\
[i \leftrightarrow j] = x_j \overline{A_{ji}^*} \, \overline{y_i} = x\overline{\mathbf{A}^*} \, \overline{y}.
\end{aligned}
\tag{1.147}
$$

Therefore, by comparing Es. (1.147) and /1.146), we obtain $\overline{\mathbf{A}^*} = \mathbf{A}^T$, so that

$$
\mathbf{A}^* = \overline{\mathbf{A}^T} = \overline{\mathbf{A}}^T.
\tag{1.148}
$$

That is, in matrix notation, the adjoint transformation is just the transpose of the complex conjugate of the original matrix.

Accordingly, in real inner product spaces, $\mathbf{A}^* = \overline{\mathbf{A}}^T = \mathbf{A}^T$ is just the transpose of $\mathbf{A}$:

$$
[\mathbf{x}, \mathbf{A}^T\mathbf{y}] = [\mathbf{A}\mathbf{x}, \mathbf{y}].
\tag{1.149}
$$

In complex inner product spaces, define the Hermitian conjugate matrix by $\mathbf{A}^\dagger = \mathbf{A}^* = \overline{\mathbf{A}^T} = \overline{\mathbf{A}}^T$, so that

$$
[\mathbf{x}, \mathbf{A}^\dagger\mathbf{y}] = [\mathbf{A}\mathbf{x}, \mathbf{y}].
\tag{1.150}
$$

1.19.3   Properties

We mention without proof that the adjoint operator is a linear operator. Furthermore, $\mathbf{0}^* = \mathbf{0}$, $\mathbf{1}^* = \mathbf{1}$, $(\mathbf{A} + \mathbf{B})^* = \mathbf{A}^* + \mathbf{B}^*$, $(\alpha\mathbf{A})^* = \alpha\mathbf{A}^*$, $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$, and $(\mathbf{A}^{-1})^* = (\mathbf{A}^*)^{-1}$.

A proof for $(\mathbf{AB})^* = \mathbf{B}^*\mathbf{A}^*$ is $[\mathbf{x}, (\mathbf{AB})^*\mathbf{y}] = [\mathbf{AB}\mathbf{x}, \mathbf{y}] = [\mathbf{B}\mathbf{x}, \mathbf{A}^*\mathbf{y}] = [\mathbf{x}, \mathbf{B}^*\mathbf{A}^*\mathbf{y}]$.

If $\mathbf{A}$ and $\mathbf{B}$ commute, then $(\mathbf{AB})^* = \mathbf{A}^*\mathbf{B}^*$. Therefore, by identifying $\mathbf{B}$ with $\mathbf{A}$, and repeating this, $(\mathbf{A}^n)^* = (\mathbf{A}^*)^n$. In particular, if $\mathbf{E}$ is a projection, then $\mathbf{E}^*$ is a projection, since $(\mathbf{E}^*)^2 = (\mathbf{E}^2)^* = \mathbf{E}^*$.

For finite dimensions,

$$
\mathbf{A}^{**} = \mathbf{A},
\tag{1.151}
$$

as, *per* definition, $[\mathbf{A}\mathbf{x}, \mathbf{y}] = [\mathbf{x}, \mathbf{A}^*\mathbf{y}] = [(\mathbf{A}^*)^*\mathbf{x}, \mathbf{y}]$.

1.20   Self-adjoint transformation

The following definition yields some analogy to real numbers as compared to complex numbers ("a complex number z is real if $\overline{z} = z$"), expressed in terms of operators on a complex vector space.

An operator $\mathbf{A}$ on a linear vector space $\mathfrak{V}$ is called *self-adjoint*, if

$$
\mathbf{A}^* = \mathbf{A}
\tag{1.152}
$$

and if the domains of $\mathbf{A}$ and $\mathbf{A}^*$ – that is, the set of vectors on which they are well defined – coincide.

In finite dimensional *real* inner product spaces, self-adoint operators are called *symmetric*, since they are symmetric with respect to transpositions; that is,

$$
\mathbf{A}^* = \mathbf{A}^T = \mathbf{A}.
\tag{1.153}
$$

In finite dimensional *complex* inner product spaces, self-adoint operators are called *Hermitian,* since they are identical with respect to Hermitian conjugation (transposition of the matrix and complex conjugation of its entries); that is,

$$\mathbf{A}^* = \mathbf{A}^\dagger = \mathbf{A}. \tag{1.154}$$

In what follows, we shall consider only the latter case and identify self-adjoint operators with Hermitian ones. In terms of matrices, a matrix A corresponding to an operator $\mathbf{A}$ in some fixed basis is self-adjoint if

$$A^\dagger \equiv (\overline{A_{ij}})^T = \overline{A_{ji}} = A_{ij} \equiv A. \tag{1.155}$$

That is, suppose A_{ij} is the matrix representation corresponding to a linear transformation $\mathbf{A}$ in some basis $\mathfrak{B}$, then the *Hermitian* matrix $\mathbf{A}^* = \mathbf{A}^\dagger$ to the dual basis $\mathfrak{B}^*$ is $\overline{(A_{ij})}^T$.

For the sake of an example, consider again the *Pauli spin matrices*

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{1.156}$$

$$\sigma_1 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

which, together with the identity, that is, $\mathbb{I}_2 = \mathrm{diag}(1,1)$, are all self-adjoint.

The following operators are not self-adjoint:

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ i & 0 \end{pmatrix}, \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix}. \tag{1.157}$$

Note that the coherent real-valued superposition of a self-adjoint transformations (such as the sum or difference of correlations in the Clauser-Horne-Shimony-Holt expression [21]) is a self-adjoint transformation.

For a direct proof, suppose that $\alpha_i \in \mathbb{R}$ for all $1 \le i \le n$ are n real-valued coefficients and $\mathbf{A}_1, \dots \mathbf{A}_n$ are n self-adjoint operators. Then $\mathbf{B} = \sum_{i=1}^{n} \alpha_i \mathbf{A}_i$ is self-adjoint, since

$$\mathbf{B}^* = \sum_{i=1}^{n} \overline{\alpha_i} \mathbf{A}_i^* = \sum_{i=1}^{n} \alpha_i \mathbf{A}_i = \mathbf{B}. \tag{1.158}$$

1.21 *Positive transformation*

A linear transformation $\mathbf{A}$ on an inner product space $\mathfrak{V}$ is *positive* (or, used synonymously, *nonnegative*), that is, in symbols $\mathbf{A} \ge 0$, if $\langle \mathbf{Ax} \mid \mathbf{x} \rangle \ge 0$ for all $\mathbf{x} \in \mathfrak{V}$. If $\langle \mathbf{Ax} \mid \mathbf{x} \rangle = 0$ implies $\mathbf{x} = 0$, $\mathbf{A}$ is called *strictly positive*.

Positive transformations – indeed, transformations with real inner products such that $\langle \mathbf{Ax}|\mathbf{x}\rangle = \overline{\langle \mathbf{x}|\mathbf{Ax}\rangle} = \langle \mathbf{x}|\mathbf{Ax}\rangle$ for all vectors $\mathbf{x}$ of a complex inner product space $\mathfrak{V}$ – are self-adjoint.

For a direct proof, recall the polarization identity (1.9) in a slightly different form, with the first argument (vector) transformed by $\mathbf{A}$, as well as the definition of the adjoint operator (1.145) on page 43, and write

$$
\begin{aligned}
\langle \mathbf{x} | \mathbf{A}^* \mathbf{y} \rangle = \langle \mathbf{A}\mathbf{x} | \mathbf{y} \rangle &= \frac{1}{4} \left[\langle \mathbf{A}(\mathbf{x}+\mathbf{y}) | \mathbf{x}+\mathbf{y} \rangle - \langle \mathbf{A}(\mathbf{x}-\mathbf{y}) | \mathbf{x}-\mathbf{y} \rangle \right. \\
&\quad + i \langle \mathbf{A}(\mathbf{x}+i\mathbf{y}) | \mathbf{x}+i\mathbf{y} \rangle - i \langle \mathbf{A}(\mathbf{x}-i\mathbf{y}) | \mathbf{x}-i\mathbf{y} \rangle \left. \right] \\
&= \frac{1}{4} \left[\langle \mathbf{x}+\mathbf{y} | \mathbf{A}(\mathbf{x}+\mathbf{y}) \rangle - \langle \mathbf{x}-\mathbf{y} | \mathbf{A}(\mathbf{x}-\mathbf{y}) \rangle \right. \\
&\quad + i \langle \mathbf{x}+i\mathbf{y} | \mathbf{A}(\mathbf{x}+i\mathbf{y}) \rangle - i \langle \mathbf{x}-i\mathbf{y} | \mathbf{A}(\mathbf{x}-i\mathbf{y}) \rangle \left. \right] = \langle \mathbf{x} | \mathbf{A}\mathbf{y} \rangle .
\end{aligned}
\tag{1.159}
$$

1.22 Unitary transformations and isometries

1.22.1 Definition

Note that a complex number z has absolute value one if $z\overline{z} = 1$, or $\overline{z} = 1/z$. In analogy to this "modulus one" behavior, consider *unitary transformations*, or, used synonymuously, *(one-to-one) isometries* $\mathbf{U}$ for which

$$
\mathbf{U}^* = \mathbf{U}^\dagger = \mathbf{U}^{-1}, \text{ or } \mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger \mathbf{U} = \mathbb{I}.
\tag{1.160}
$$

The following conditions are equivalent:

(i) $\mathbf{U}^* = \mathbf{U}^\dagger = \mathbf{U}^{-1}$, or $\mathbf{U}\mathbf{U}^\dagger = \mathbf{U}^\dagger \mathbf{U} = \mathbb{I}$.

(ii) $\langle \mathbf{U}\mathbf{x} | \mathbf{U}\mathbf{y} \rangle = \langle \mathbf{x} | \mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y} \in \mathfrak{V}$;

(iii) $\mathbf{U}$ is an *isometry*; that is, preserving the norm $\|\mathbf{U}\mathbf{x}\| = \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathfrak{V}$.

(iv) $\mathbf{U}$ represents a change of orthonormal basis [22]: Let $\mathfrak{B} = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ be an orthonormal basis. Then $\mathbf{U}\mathfrak{B} = \mathfrak{B}' = \{\mathbf{U}\mathbf{f}_1, \mathbf{U}\mathbf{f}_2, \ldots, \mathbf{U}\mathbf{f}_n\}$ is also an orthonormal basis of $\mathfrak{V}$. Conversely, two arbitrary orthonormal bases $\mathfrak{B}$ and $\mathfrak{B}'$ are connected by a unitary transformation $\mathbf{U}$ *via* the pairs $\mathbf{f}_i$ and $\mathbf{U}\mathbf{f}_i$ for all $1 \le i \le n$, respectively. More explicitly, denote $\mathbf{U}\mathbf{f}_i = \mathbf{e}_i$; then (recall $\mathbf{f}_i$ and $\mathbf{e}_i$ are elements of the orthonormal bases $\mathfrak{B}$ and $\mathbf{U}\mathfrak{B}$, respectively) $\mathbf{U}_{ef} = \sum_{i=1}^{n} \mathbf{e}_i \mathbf{f}_i^\dagger$.

For a direct proof, suppose that (i) holds; that is, $\mathbf{U}^* = \mathbf{U}^\dagger = \mathbf{U}^{-1}$. then, (ii) follows by

$$
\langle \mathbf{U}\mathbf{x} | \mathbf{U}\mathbf{y} \rangle = \langle \mathbf{U}^* \mathbf{U}\mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{U}^{-1} \mathbf{U}\mathbf{x} | \mathbf{y} \rangle = \langle \mathbf{x} | \mathbf{y} \rangle
\tag{1.161}
$$

for all $\mathbf{x}, \mathbf{y}$.

In particular, if $\mathbf{y} = \mathbf{x}$, then

$$
\|\mathbf{U}\mathbf{x}\|^2 = |\langle \mathbf{U}\mathbf{x} | \mathbf{U}\mathbf{x} \rangle| = |\langle \mathbf{x} | \mathbf{x} \rangle| = \|\mathbf{x}\|^2
\tag{1.162}
$$

for all $\mathbf{x}$.

In order to prove (i) from (iii) consider the transformation $\mathbf{A} = \mathbf{U}^* \mathbf{U} - \mathbb{I}$, motivated by (1.162), or, by linearity of the inner product in the first argument,

$$
\begin{aligned}
\|\mathbf{U}\mathbf{x}\| - \|\mathbf{x}\| &= \langle \mathbf{U}\mathbf{x} | \mathbf{U}\mathbf{x} \rangle - \langle \mathbf{x} | \mathbf{x} \rangle = \langle \mathbf{U}^* \mathbf{U}\mathbf{x} | \mathbf{x} \rangle - \langle \mathbf{x} | \mathbf{x} \rangle = \\
&= \langle \mathbf{U}^* \mathbf{U}\mathbf{x} | \mathbf{x} \rangle - \langle \mathbb{I}\mathbf{x} | \mathbf{x} \rangle = \langle (\mathbf{U}^* \mathbf{U} - \mathbb{I})\mathbf{x} | \mathbf{x} \rangle = 0
\end{aligned}
\tag{1.163}
$$

for all $\mathbf{x}$. $\mathbf{A}$ is self-adjoint, since

$$\mathbf{A}^* = \left(\mathbf{U}^*\mathbf{U}\right)^* - \mathbb{I}^* = \mathbf{U}^* \left(\mathbf{U}^*\right)^* - \mathbb{I} = \mathbf{U}^*\mathbf{U} - \mathbb{I} = \mathbf{A}. \qquad (1.164)$$

We need to prove that a necessary and sufficient condition for a self-adjoint linear transformation $\mathbf{A}$ on an inner product space to be 0 is that $\langle \mathbf{Ax} \,|\, \mathbf{x} \rangle = 0$ for all vectors $\mathbf{x}$.

Necessity is easy: whenever $\mathbf{A} = 0$ the scalar product vanishes. A proof of sufficiency first notes that, by linearity allowing the expansion of the first summand on the right side,

$$\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle + \langle \mathbf{Ay} \,|\, \mathbf{x} \rangle = \langle \mathbf{A}(\mathbf{x}+\mathbf{y}) \,|\, \mathbf{x}+\mathbf{y} \rangle - \langle \mathbf{Ax} \,|\, \mathbf{x} \rangle - \langle \mathbf{Ay} \,|\, \mathbf{y} \rangle. \qquad (1.165)$$

Since $\mathbf{A}$ is self-adjoint, the left side is

$$\begin{aligned} \langle \mathbf{Ax} \,|\, \mathbf{y} \rangle + \langle \mathbf{Ay} \,|\, \mathbf{x} \rangle &= \langle \mathbf{Ax} \,|\, \mathbf{y} \rangle + \langle \mathbf{y} \,|\, \mathbf{A}^*\mathbf{x} \rangle = \langle \mathbf{Ax} \,|\, \mathbf{y} \rangle + \langle \mathbf{y} \,|\, \mathbf{Ax} \rangle = \\ &= \langle \mathbf{Ax} \,|\, \mathbf{y} \rangle + \overline{\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle} = 2\Re\left(\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle\right). \end{aligned} \qquad (1.166)$$

Note that our assumption implied that the right hand side of (1.165) vanishes. Thus,

$$2\Re\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle = 0. \qquad (1.167)$$

Since the real part $\Re\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle$ of $\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle$ vanishes, what remains is to show that the imaginary part $\Im\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle$ of $\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle$ vanishes as well.

As long as the Hilbert space is real (and thus the self-adjoint transformation $\mathbf{A}$ is just symmetric) we are almost finished, as $\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle$ is real, with vanishing imaginary part. That is, $\Re\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle = \langle \mathbf{Ax} \,|\, \mathbf{y} \rangle = 0$. In this case, we are free to identify $\mathbf{y} = \mathbf{Ax}$, thus obtaining $\langle \mathbf{Ax} \,|\, \mathbf{Ax} \rangle = 0$ for all vectors $\mathbf{x}$. Because of the positive-definiteness [condition (iii) on page 7] we must have $\mathbf{Ax} = 0$ for all vectors $\mathbf{x}$, and thus finally $\mathbf{A} = \mathbf{U}^*\mathbf{U} - \mathbb{I} = 0$, and $\mathbf{U}^*\mathbf{U} = \mathbb{I}$.

In the case of complex Hilbert space, and thus $\mathbf{A}$ being Hermitian, we can find an unimodular complex number θ such that $|\theta| = 1$, and, in particular, $\theta = \theta(\mathbf{x}, \mathbf{y}) = +i$ for $\Im\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle < 0$ or $\theta(\mathbf{x}, \mathbf{y}) = -i$ for $\Im\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle \geq 0$, such that $\theta\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle = |\Im\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle| = |\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle|$ (recall that the real part of $\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle$ vanishes).

Now we are free to substitute $\theta\mathbf{x}$ for $\mathbf{x}$. We can again start with our assumption (iii), now with $\mathbf{x} \to \theta\mathbf{x}$ and thus rewritten as $0 = \langle \mathbf{A}(\theta\mathbf{x}) \,|\, \mathbf{y} \rangle$, which we have already converted into $0 = \Re\langle \mathbf{A}(\theta\mathbf{x}) \,|\, \mathbf{y} \rangle$ for self-adjoint (Hermitian) $\mathbf{A}$. By linearity in the first argument of the inner product we obtain

$$\begin{aligned} 0 = \Re\langle \mathbf{A}(\theta\mathbf{x}) \,|\, \mathbf{y} \rangle &= \Re\langle \theta\mathbf{Ax} \,|\, \mathbf{y} \rangle = \Re\left(\theta\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle\right) = \\ &= \Re\left(|\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle|\right) = |\langle \mathbf{Ax} \,|\, \mathbf{y} \rangle| = \langle \mathbf{Ax} \,|\, \mathbf{y} \rangle. \end{aligned} \qquad (1.168)$$

Again we can identify $\mathbf{y} = \mathbf{Ax}$, thus obtaining $\langle \mathbf{Ax} \,|\, \mathbf{Ax} \rangle = 0$ for all vectors $\mathbf{x}$. Because of the positive-definiteness [condition (iii) on page 7] we must have $\mathbf{Ax} = 0$ for all vectors $\mathbf{x}$, and thus finally $\mathbf{A} = \mathbf{U}^*\mathbf{U} - \mathbb{I} = 0$, and $\mathbf{U}^*\mathbf{U} = \mathbb{I}$.

A proof of (iv) from (i) can be given as follows: identify $\mathbf{x}$ with $\mathbf{f}_i$ and $\mathbf{y}$ with $\mathbf{f}_j$ with elements of the original basis. Then, since $\mathbf{U}$ is an isometry,

$$\langle \mathbf{Uf}_i \,|\, \mathbf{Uf}_j \rangle = \langle \mathbf{U}^*\mathbf{Uf}_i \,|\, \mathbf{f}_j \rangle = \langle \mathbf{U}^{-1}\mathbf{Uf}_i \,|\, \mathbf{f}_j \rangle = \langle \mathbf{f}_i \,|\, \mathbf{f}_j \rangle = \delta_{ij}. \qquad (1.169)$$

$\Re z$ and $\Im z$ stand for the real real and imaginary parts of the complex number $z = \Re z + i\Im z$.

$\mathbf{U}\mathfrak{B}$ forms a new basis: both $\mathfrak{B}$ as well as $\mathbf{U}\mathfrak{B}$ have the same number of mutually orthonormal elements; furthermore, completeness of $\mathbf{U}\mathfrak{B}$ follows from the completeness of $\mathfrak{B}$; with $\langle \mathbf{x} \mid \mathbf{U}\mathbf{f}_j \rangle = \langle \mathbf{U}^*\mathbf{x} \mid \mathbf{f}_j \rangle = 0$ for all basis elements $\mathbf{f}_j$ implying $\mathbf{U}^*\mathbf{x} = \mathbf{x} = 0$.

Conversely, if both $\mathbf{U}\mathbf{f}_i$ are orthonormal bases with $\mathbf{f}_i \in \mathfrak{B}$ and $\mathbf{U}\mathbf{f}_i \in \mathbf{U}\mathfrak{B} = \mathfrak{B}'$, so that $\langle \mathbf{U}\mathbf{f}_i \mid \mathbf{U}\mathbf{f}_j \rangle = \langle \mathbf{f}_i \mid \mathbf{f}_j \rangle$, then by linearity $\langle \mathbf{U}\mathbf{x} \mid \mathbf{U}\mathbf{y} \rangle = \langle \mathbf{x} \mid \mathbf{y} \rangle$ for all $\mathbf{x}, \mathbf{y}$, thus proving (ii) from (iv).

Note that $\mathbf{U}$ preserves length or distances and thus is an *isometry*, as for all $\mathbf{x}, \mathbf{y}$,

$$\| \mathbf{U}\mathbf{x} - \mathbf{U}\mathbf{y} \| = \| \mathbf{U}\left(\mathbf{x} - \mathbf{y}\right) \| = \| \mathbf{x} - \mathbf{y} \|. \tag{1.170}$$

Note also that $\mathbf{U}$ preserves the angle θ between two nonzero vectors $\mathbf{x}$ and $\mathbf{y}$ defined by

$$\cos\theta = \frac{\langle \mathbf{x} \mid \mathbf{y} \rangle}{\| \mathbf{x} \| \| \mathbf{y} \|} \tag{1.171}$$

as it preserves the inner product and the norm.

Since unitary transformations can also be defined via *one-to-one transformations preserving the scalar product,* functions such as $f : x \mapsto x' = \alpha x$ with $\alpha \neq e^{i\varphi}$, $\varphi \in \mathbb{R}$, do not correspond to a unitary transformation in a one-dimensional Hilbert space, as the scalar product $f : \langle x|y \rangle \mapsto \langle x'|y' \rangle = |\alpha|^2 \langle x|y \rangle$ is not preserved; whereas if α is a modulus of one; that is, with $\alpha = e^{i\varphi}$, $\varphi \in \mathbb{R}$, $|\alpha|^2 = 1$, and the scalar product is preseved. Thus, $u : x \mapsto x' = e^{i\varphi} x$, $\varphi \in \mathbb{R}$, represents a unitary transformation.

1.22.2 Characterization in terms of orthonormal basis

A complex matrix $\mathbf{U}$ is unitary if and only if its row (or column) vectors form an orthonormal basis.

This can be readily verified [23] by writing $\mathbf{U}$ in terms of two orthonormal bases $\mathfrak{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\} \equiv \{|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle, \ldots, |\mathbf{e}_n\rangle\}$ $\mathfrak{B}' = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\} \equiv \{|\mathbf{f}_1\rangle, |\mathbf{f}_2\rangle, \ldots, |\mathbf{f}_n\rangle\}$ as

$$\mathbf{U}_{ef} = \sum_{i=1}^{n} \mathbf{e}_i \mathbf{f}_i^{\dagger} \equiv \sum_{i=1}^{n} |\mathbf{e}_i\rangle\langle\mathbf{f}_i|. \tag{1.172}$$

[23] Julian Schwinger. Unitary operators bases. *Proceedings of the National Academy of Sciences (PNAS)*, 46:570–579, 1960. DOI: 10.1073/pnas.46.4.570. URL http://doi.org/10.1073/pnas.46.4.570

Together with $\mathbf{U}_{fe} = \sum_{i=1}^{n} \mathbf{f}_i \mathbf{e}_i^{\dagger} \equiv \sum_{i=1}^{n} |\mathbf{f}_i\rangle\langle\mathbf{e}_i|$ we form

$$\mathbf{e}_k^{\dagger}\mathbf{U}_{ef} = \mathbf{e}_k^{\dagger} \sum_{i=1}^{n} \mathbf{e}_i \mathbf{f}_i^{\dagger} = \sum_{i=1}^{n} (\mathbf{e}_k^{\dagger}\mathbf{e}_i)\mathbf{f}_i^{\dagger} = \sum_{i=1}^{n} \delta_{ki}\mathbf{f}_i^{\dagger} = \mathbf{f}_k^{\dagger}. \tag{1.173}$$

In a similar way we find that

$$\mathbf{U}_{ef}\mathbf{f}_k = \mathbf{e}_k, \mathbf{f}_k^{\dagger}\mathbf{U}_{fe} = \mathbf{e}_k^{\dagger}, \mathbf{U}_{fe}\mathbf{e}_k = \mathbf{f}_k. \tag{1.174}$$

Moreover,

$$\mathbf{U}_{ef}\mathbf{U}_{fe} = \sum_{i=1}^{n}\sum_{j=1}^{n}(|\mathbf{e}_i\rangle\langle\mathbf{f}_i|)(|\mathbf{f}_j\rangle\langle\mathbf{e}_j|) = \sum_{i=1}^{n}\sum_{j=1}^{n}|\mathbf{e}_i\rangle\delta_{ij}\langle\mathbf{e}_j| = \sum_{i=1}^{n}|\mathbf{e}_i\rangle\langle\mathbf{e}_i| = \mathbb{I}. \tag{1.175}$$

In a similar way we obtain $\mathbf{U}_{fe}\mathbf{U}_{ef} = \mathbb{I}$. Since

$$\mathbf{U}_{ef}^{\dagger} = \sum_{i=1}^{n}(\mathbf{f}_i^{\dagger})^{\dagger}\mathbf{e}_i^{\dagger} = \sum_{i=1}^{n}\mathbf{f}_i\mathbf{e}_i^{\dagger} = \mathbf{U}_{fe}, \tag{1.176}$$

we obtain that $\mathbf{U}_{ef}^{\dagger} = (\mathbf{U}_{ef})^{-1}$ and $\mathbf{U}_{fe}^{\dagger} = (\mathbf{U}_{fe})^{-1}$.

Note also that the *composition* holds; that is, $\mathbf{U}_{ef}\mathbf{U}_{fg} = \mathbf{U}_{eg}$.

If we identify one of the bases $\mathfrak{B}$ and $\mathfrak{B}'$ by the Cartesian standard basis, it becomes clear that, for instance, every unitary operator $\mathbf{U}$ can be written in terms of an orthonormal basis of the dual space $\mathfrak{B}^* = \{\langle\mathbf{f}_1|, \langle\mathbf{f}_2|\ldots, \langle\mathbf{f}_n|\}$ by "stacking" the conjugate transpose vectors of that orthonormal basis "on top of each other;" by "stacking" the conjugate transpose vectors of that orthonormal basis "on top of each other;" that is

$$\mathbf{U} \equiv \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \mathbf{f}_1^{\dagger} + \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} \mathbf{f}_2^{\dagger} + \cdots + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ n \end{pmatrix} \mathbf{f}_n^{\dagger} \equiv \begin{pmatrix} \mathbf{f}_1^{\dagger} \\ \mathbf{f}_2^{\dagger} \\ \vdots \\ \mathbf{f}_n^{\dagger} \end{pmatrix} \equiv \begin{pmatrix} \langle\mathbf{f}_1| \\ \langle\mathbf{f}_2| \\ \vdots \\ \langle\mathbf{f}_n| \end{pmatrix}. \tag{1.177}$$

Thereby the conjugate transpose vectors of the orthonormal basis $\mathfrak{B}$ serve as the rows of $\mathbf{U}$.

In a similar manner, every unitary operator $\mathbf{U}$ can be written in terms of an orthonormal basis $\mathfrak{B} = \{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ by "pasting" the vectors of that orthonormal basis "one after another;" that is

$$\begin{aligned} \mathbf{U} &\equiv \mathbf{f}_1\left(1, 0, \ldots, 0\right) + \mathbf{f}_2\left(0, 1, \ldots, 0\right) + \cdots + \mathbf{f}_n\left(0, 0, \ldots, 1\right) \\ &\equiv \left(\mathbf{f}_1, \mathbf{f}_2, \cdots, \mathbf{f}_n\right) \equiv \left(|\mathbf{f}_1\rangle, |\mathbf{f}_2\rangle, \cdots, |\mathbf{f}_n\rangle\right). \end{aligned} \tag{1.178}$$

Thereby the vectors of the orthonormal basis $\mathfrak{B}$ serve as the columns of $\mathbf{U}$.

Note also that any permutation of vectors in $\mathfrak{B}$ would also yield unitary matrices.

1.23 *Orthonormal (orthogonal) transformations*

Orthonormal (orthogonal) transformations are special cases of unitary transformations restricted to *real* Hilbert space.

An *orthonormal* or *orthogonal transformation* $\mathbf{R}$ is a linear transformation whose corresponding square matrix R has real-valued entries and mutually orthogonal, normalized row (or, equivalently, column) vectors. As a consequence (see the equivalence of definitions of unitary definitions and the proofs mentioned earlier),

$$RR^T = R^T R = \mathbb{I}, \text{ or } R^{-1} = R^T. \tag{1.179}$$

As all unitary transformations, orthonormal transformations $\mathbf{R}$ preserve a symmetric inner product as well as the norm.

If $\det R = 1$, $\mathbf{R}$ corresponds to a *rotation*. If $\det R = -1$, $\mathbf{R}$ corresponds to a rotation and a *reflection*. A reflection is an isometry (a distance preserving map) with a hyperplane as set of fixed points.

For the sake of a two-dimensional example of rotations in the plane $\mathbb{R}^2$, take the rotation matrix in Eq. (1.115) representing a rotation of the basis by an angle φ.

1.24 Permutations

Permutations are "discrete" orthogonal transformations in the sense that they merely allow the entries "0" and "1" in the respective matrix representations. With regards to classical and quantum bits [24] they serve as a sort of "reversible classical analogue" for classical reversible computation, as compared to the more general, continuous unitary transformations of quantum bits introduced earlier.

Permutation matrices are defined by the requirement that they only contain a single nonvanishing entry "1" per row and column; all the other row and column entries vanish; that is, the respective matrix entry is "0." For example, the matrices $\mathbb{I}_n = \mathrm{diag}(\underbrace{1,\ldots,1}_{n \text{ times}})$, or

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \text{ or } \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

are permutation matrices.

Note that from the definition and from matrix multiplication follows that, if P is a permutation matrix, then $PP^T = P^T P = \mathbb{I}_n$. That is, P^T represents the inverse element of P. As P is real-valued, it is a *normal operator* (cf. page 57).

Note further that, as all unitary matrices, any permuation matrix can be interpreted in terms of row and column vectors: The set of all these row and column vectors constitute the Cartesian standard basis of n-dimensional vector space, with permuted elements.

Note also that, if P and Q are permutation matrices, so is PQ and QP. The set of all $n!$ permutation $(n \times n)$−matrices correponding to permutations of n elements of $\{1,2,\ldots,n\}$ form the *symmetric group S_n*, with $\mathbb{I}_n$ being the identity element.

1.25 Orthogonal (perpendicular) projections

Orthogonal, or, used synonymously, *perpendicular* projections are associated with a *direct sum decomposition* of the vector space $\mathfrak{V}$; that is,

$$\mathfrak{M} \oplus \mathfrak{M}^\perp = \mathfrak{V}, \tag{1.180}$$

whereby $\mathfrak{M} = P_\mathfrak{M}(\mathfrak{V})$ is the image of some projector $\mathbf{E} = P_\mathfrak{M}$ along $\mathfrak{M}^\perp$, and $\mathfrak{M}^\perp$ is the kernel of $P_\mathfrak{M}$. That is, $\mathfrak{M}^\perp = \{\mathbf{x} \in \mathfrak{V} \mid P_\mathfrak{M}(\mathbf{x}) = \mathbf{0}\}$ is the subspace of $\mathfrak{V}$ whose elements are mapped to the zero vector $\mathbf{0}$ by $P_\mathfrak{M}$.

Let us, for the sake of concreteness, suppose that, in n-dimensional complex Hilbert space $\mathbb{C}^n$, we are given a k-dimensional subspace

$$\mathfrak{M} = \mathrm{span}(\mathbf{x}_1,\ldots,\mathbf{x}_k) \equiv \mathrm{span}(|\mathbf{x}_1\rangle,\ldots,|\mathbf{x}_k\rangle) \tag{1.181}$$

spanned by $k \le n$ linear independent base vectors $\mathbf{x}_1,\ldots,\mathbf{x}_k$. In addition, we are given another (arbitrary) vector $\mathbf{y} \in \mathbb{C}^n$.

Now consider the following question: how can we project $\mathbf{y}$ onto $\mathfrak{M}$ orthogonally (perpendicularly)? That is, can we find a vector $\mathbf{y}' \in \mathfrak{M}$ so that $\mathbf{y}^\perp = \mathbf{y} - \mathbf{y}'$ is orthogonal (perpendicular) to all of $\mathfrak{M}$?

The orthogonality of $\mathbf{y}^\perp$ on the entire $\mathfrak{M}$ can be rephrased in terms of all the vectors $\mathbf{x}_1, \ldots, \mathbf{x}_k$ spanning $\mathfrak{M}$; that is, for all $\mathbf{x}_i \in \mathfrak{M}$, $1 \le i \le k$ we must have $\langle \mathbf{x}_i | \mathbf{y}^\perp \rangle = 0$. This can be transformed into matrix algebra by considering the $n \times k$ matrix [note that $\mathbf{x}_i$ are column vectors, and recall the construction in Eq. (1.178)]

$$\mathbf{A} = \left(\mathbf{x}_1, \ldots, \mathbf{x}_k \right) \equiv \left(|\mathbf{x}_1\rangle, \ldots, |\mathbf{x}_k\rangle \right), \tag{1.182}$$

and by requiring

$$\mathbf{A}^\dagger |\mathbf{y}^\perp\rangle \equiv \mathbf{A}^\dagger \mathbf{y}^\perp = \mathbf{A}^\dagger \left(\mathbf{y} - \mathbf{y}' \right) = \mathbf{A}^\dagger \mathbf{y} - \mathbf{A}^\dagger \mathbf{y}' = 0, \tag{1.183}$$

yielding

$$\mathbf{A}^\dagger |\mathbf{y}\rangle \equiv \mathbf{A}^\dagger \mathbf{y} = \mathbf{A}^\dagger \mathbf{y}' \equiv \mathbf{A}^\dagger |\mathbf{y}'\rangle. \tag{1.184}$$

On the other hand, $\mathbf{y}'$ must be a linear combination of $\mathbf{x}_1, \ldots, \mathbf{x}_k$ with the k-tuple of coefficients $\mathbf{c}$ defined by

Recall that $(\mathbf{AB})^\dagger = \mathbf{B}^\dagger \mathbf{A}^\dagger$, and $(\mathbf{A}^\dagger)^\dagger = \mathbf{A}$.

$$\mathbf{y}' = c_1 \mathbf{x}_1 + \cdots + c_k \mathbf{x}_k = \left(\mathbf{x}_1, \ldots, \mathbf{x}_k \right) \begin{pmatrix} c_1 \\ \vdots \\ c_k \end{pmatrix} = \mathbf{A}\mathbf{c}. \tag{1.185}$$

Insertion into (1.184) yields

$$\mathbf{A}^\dagger \mathbf{y} = \mathbf{A}^\dagger \mathbf{A} \mathbf{c}. \tag{1.186}$$

Taking the inverse of $\mathbf{A}^\dagger \mathbf{A}$ (this is a $k \times k$ diagonal matrix which is invertible, since the k vectors defining $\mathbf{A}$ are linear independent), and multiplying (1.186) from the left yields

$$\mathbf{c} = \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \mathbf{A}^\dagger \mathbf{y}. \tag{1.187}$$

With (1.185) and (1.187) we find $\mathbf{y}'$ to be

$$\mathbf{y}' = \mathbf{A}\mathbf{c} = \mathbf{A} \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \mathbf{A}^\dagger \mathbf{y}. \tag{1.188}$$

We can define

$$\mathbf{E}_{\mathfrak{M}} = \mathbf{A} \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \mathbf{A}^\dagger \tag{1.189}$$

to be the *projection matrix for the subspace* $\mathfrak{M}$. Note that

$$\mathbf{E}_{\mathfrak{M}}^\dagger = \left[\mathbf{A} \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \mathbf{A}^\dagger \right]^\dagger = \mathbf{A} \left[\left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \right]^\dagger \mathbf{A}^\dagger = \mathbf{A} \left[\mathbf{A}^{-1} \left(\mathbf{A}^\dagger \right)^{-1} \right]^\dagger \mathbf{A}^\dagger$$
$$= \mathbf{A}\mathbf{A}^{-1} \left(\mathbf{A}^{-1} \right)^\dagger \mathbf{A}^\dagger = \mathbf{A}\mathbf{A}^{-1} \left(\mathbf{A}^\dagger \right)^{-1} \mathbf{A}^\dagger = \mathbf{A} \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \mathbf{A}^\dagger = \mathbf{E}_{\mathfrak{M}}, \tag{1.190}$$

that is, $\mathbf{E}_{\mathfrak{M}}$ is self-adjoint and thus normal, as well as idempotent:

$$\mathbf{E}_{\mathfrak{M}}^2 = \left(\mathbf{A} \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \mathbf{A}^\dagger \right) \left(\mathbf{A} \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \mathbf{A}^\dagger \right)$$
$$= \mathbf{A}^\dagger \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \left(\mathbf{A}^\dagger \mathbf{A} \right) \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \mathbf{A} = \mathbf{A}^\dagger \left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \mathbf{A} = \mathbf{E}_{\mathfrak{M}}. \tag{1.191}$$

Conversely, every normal projection operator has a "trivial" spectral decomposition (cf. later Sec. 1.28.1 on page 58) $\mathbf{E}_{\mathfrak{M}} = 1 \cdot \mathbf{E}_{\mathfrak{M}} + 0 \cdot \mathbf{E}_{\mathfrak{M}^\perp} = 1 \cdot \mathbf{E}_{\mathfrak{M}} + 0 \cdot (\mathbf{1} - \mathbf{E}_{\mathfrak{M}})$ associated with the two eigenvalues 0 and 1, and thus must be orthogonal.

If the basis $\mathfrak{B} = \{\mathbf{x}_1, \ldots, \mathbf{x}_k\}$ of $\mathfrak{M}$ is orthonormal, then

$$\mathbf{A}^\dagger \mathbf{A} \equiv \begin{pmatrix} \langle \mathbf{x}_1 | \\ \vdots \\ \langle \mathbf{x}_k | \end{pmatrix} \Big(|\mathbf{x}_1\rangle, \ldots, |\mathbf{x}_k\rangle \Big) = \begin{pmatrix} \langle \mathbf{x}_1 | \mathbf{x}_1\rangle & \ldots & \langle \mathbf{x}_1 | \mathbf{x}_k\rangle \\ \vdots & \vdots & \vdots \\ \langle \mathbf{x}_k | \mathbf{x}_1\rangle & \ldots & \langle \mathbf{x}_k | \mathbf{x}_k\rangle \end{pmatrix} \equiv \mathbb{I}_k \qquad (1.192)$$

represents a k-dimensional resolution of the identity operator. Thus, $\left(\mathbf{A}^\dagger \mathbf{A} \right)^{-1} \equiv (\mathbb{I}_k)^{-1}$ is also a k-dimensional resolution of the identity operator, and the orthogonal projector $\mathbf{E}_\mathfrak{M}$ in Eq. (1.189) reduces to

$$\mathbf{E}_\mathfrak{M} = \mathbf{A}\mathbf{A}^\dagger \equiv \sum_{i=1}^{k} |\mathbf{x}_i\rangle \langle \mathbf{x}_i|. \qquad (1.193)$$

The simplest example of an orthogonal projection onto a one-dimensional subspace of a Hilbert space spanned by some unit vector $|\mathbf{x}\rangle$ is the dyadic or outer product $\mathbf{E}_x = |\mathbf{x}\rangle\langle\mathbf{x}|$.

If two unit vectors $|\mathbf{x}\rangle$ and $|\mathbf{y}\rangle$ are orthogonal; that is, if $\langle\mathbf{x}|\mathbf{y}\rangle = 0$, then $\mathbf{E}_{x,y} = |\mathbf{x}\rangle\langle\mathbf{x}| + |\mathbf{y}\rangle\langle\mathbf{y}|$ is an orthogonal projector onto a two-dimensional subspace spanned by $|\mathbf{x}\rangle$ and $|\mathbf{y}\rangle$.

In general, the orthonormal projection corresponding to some arbitrary subspace of some Hilbert space can be (non-uniquely) constructed by (i) finding an orthonormal basis spanning that subsystem (this is nonunique), if necessary by a Gram-Schmidt process; (ii) forming the projection operators corresponding to the dyadic or outer product of all these vectors; and (iii) summing up all these orthogonal operators.

The following propositions are stated mostly without proof. A linear transformation $\mathbf{E}$ is an orthogonal (perpendicular) projection if and only if is self-adjoint; that is, $\mathbf{E} = \mathbf{E}^2 = \mathbf{E}^*$.

Perpendicular projections are *positive* linear transformations, with $\|\mathbf{E}\mathbf{x}\| \leq \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathfrak{V}$. Conversely, if a linear transformation $\mathbf{E}$ is idempotent; that is, $\mathbf{E}^2 = \mathbf{E}$, and $\|\mathbf{E}\mathbf{x}\| \leq \|\mathbf{x}\|$ for all $\mathbf{x} \in \mathfrak{V}$, then is self-adjoint; that is, $\mathbf{E} = \mathbf{E}^*$.

Recall that for *real* inner product spaces, the self-adjoint operator can be identified with a *symmetric* operator $\mathbf{E} = \mathbf{E}^T$, whereas for *complex* inner product spaces, the self-adjoint operator can be identified with a *Hermitian* operator $\mathbf{E} = \mathbf{E}^\dagger$.

If $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_n$ are (perpendicular) projections, then a necessary and sufficient condition that $\mathbf{E} = \mathbf{E}_1 + \mathbf{E}_2 + \cdots + \mathbf{E}_n$ be a (perpendicular) projection is that $\mathbf{E}_i\mathbf{E}_j = \delta_{ij}\mathbf{E}_i = \delta_{ij}\mathbf{E}_j$; and, in particular, $\mathbf{E}_i\mathbf{E}_j = 0$ whenever $i \neq j$; that is, that all E_i are pairwise orthogonal.

For a start, consider just two projections $\mathbf{E}_1$ and $\mathbf{E}_2$. Then we can assert that $\mathbf{E}_1 + \mathbf{E}_2$ is a projection if and only if $\mathbf{E}_1\mathbf{E}_2 = \mathbf{E}_2\mathbf{E}_1 = 0$.

Because, for $\mathbf{E}_1 + \mathbf{E}_2$ to be a projection, it must be idempotent; that is,

$$(\mathbf{E}_1 + \mathbf{E}_2)^2 = (\mathbf{E}_1 + \mathbf{E}_2)(\mathbf{E}_1 + \mathbf{E}_2) = \mathbf{E}_1^2 + \mathbf{E}_1\mathbf{E}_2 + \mathbf{E}_2\mathbf{E}_1 + \mathbf{E}_2^2 = \mathbf{E}_1 + \mathbf{E}_2. \quad (1.194)$$

As a consequence, the cross-product terms in (1.194) must vanish; that is,

$$\mathbf{E}_1\mathbf{E}_2 + \mathbf{E}_2\mathbf{E}_1 = 0. \qquad (1.195)$$

Multiplication of (1.195) with $\mathbf{E}_1$ from the left and from the right yields

$$\mathbf{E}_1\mathbf{E}_1\mathbf{E}_2 + \mathbf{E}_1\mathbf{E}_2\mathbf{E}_1 = 0,$$

$$\mathbf{E}_1\mathbf{E}_2 + \mathbf{E}_1\mathbf{E}_2\mathbf{E}_1 = 0; \text{ and}$$

$$\mathbf{E}_1\mathbf{E}_2\mathbf{E}_1 + \mathbf{E}_2\mathbf{E}_1\mathbf{E}_1 = 0,$$

$$\mathbf{E}_1\mathbf{E}_2\mathbf{E}_1 + \mathbf{E}_2\mathbf{E}_1 = 0.$$

(1.196)

Subtraction of the resulting pair of equations yields

$$\mathbf{E}_1\mathbf{E}_2 - \mathbf{E}_2\mathbf{E}_1 = [\mathbf{E}_1,\mathbf{E}_2] = 0, \qquad (1.197)$$

or

$$\mathbf{E}_1\mathbf{E}_2 = \mathbf{E}_2\mathbf{E}_1. \qquad (1.198)$$

Hence, in order for the cross-product terms in Eqs. (1.194) and (1.195) to vanish, we must have

$$\mathbf{E}_1\mathbf{E}_2 = \mathbf{E}_2\mathbf{E}_1 = 0. \qquad (1.199)$$

Proving the reverse statement is straightforward, since (1.199) implies (1.194).

A generalisation by induction to more than two projections is straightforward, since, for instance, $(\mathbf{E}_1 + \mathbf{E}_2)\mathbf{E}_3 = 0$ implies $\mathbf{E}_1\mathbf{E}_3 + \mathbf{E}_2\mathbf{E}_3 = 0$. Multiplication with $\mathbf{E}_1$ from the left yields $\mathbf{E}_1\mathbf{E}_1\mathbf{E}_3 + \mathbf{E}_1\mathbf{E}_2\mathbf{E}_3 = \mathbf{E}_1\mathbf{E}_3 = 0$.

Examples for projections which are not orthogonal are

$$\begin{pmatrix} 1 & \alpha \\ 0 & 0 \end{pmatrix}, \text{ or } \begin{pmatrix} 1 & 0 & \alpha \\ 0 & 1 & \beta \\ 0 & 0 & 0 \end{pmatrix},$$

with $\alpha \neq 0$. Such projectors are sometimes called *oblique* projections.

Indeed, for two-dimensional Hilbert space, the solution of idempotence

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

yields the three orthogonal projections

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \text{ and } \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

as well as a continuum of oblique projections

$$\begin{pmatrix} 0 & 0 \\ c & 1 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ c & 0 \end{pmatrix}, \text{ and } \begin{pmatrix} a & b \\ \frac{a(1-a)}{b} & 1-a \end{pmatrix},$$

with $a, b, c \neq 0$.

1.26   Proper value or eigenvalue

1.26.1   Definition

A scalar λ is a *proper value* or *eigenvalue*, and a nonzero vector $\mathbf{x}$ is a *proper vector* or *eigenvector* of a linear transformation $\mathbf{A}$ if

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} = \lambda\mathbb{I}\mathbf{x}. \qquad (1.200)$$

In an n-dimensional vector space $\mathfrak{V}$ The set of the set of eigenvalues and the set of the associated eigenvectors $\{\{\lambda_1,\ldots,\lambda_k\}, \{\mathbf{x}_1,\ldots,\mathbf{x}_n\}\}$ of a linear transformation $\mathbf{A}$ form an *eigensystem* of $\mathbf{A}$.

1.26.2 Determination

Since the eigenvalues and eigenvectors are those scalars λ vectors $\mathbf{x}$ for which $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, this equation can be rewritten with a zero vector on the right side of the equation; that is ($\mathbb{I} = \mathrm{diag}(1,\ldots,1)$ stands for the identity matrix),

$$(\mathbf{A} - \lambda\mathbb{I})\mathbf{x} = \mathbf{0}. \tag{1.201}$$

Suppose that $\mathbf{A} - \lambda\mathbb{I}$ is invertible. Then we could formally write $\mathbf{x} = (\mathbf{A} - \lambda\mathbb{I})^{-1}\mathbf{0}$; hence $\mathbf{x}$ must be the zero vector.

We are not interested in this trivial solution of Eq. (1.201). Therefore, suppose that, contrary to the previous assumption, $\mathbf{A} - \lambda\mathbb{I}$ is *not* invertible. We have mentioned earlier (without proof) that this implies that its determinant vanishes; that is,

$$\det(\mathbf{A} - \lambda\mathbb{I}) = |\mathbf{A} - \lambda\mathbb{I}| = 0. \tag{1.202}$$

This determinant is often called the *secular determinant*; and the corresponding equation after expansion of the determinant is called the *secular equation* or *characteristic equation*. Once the eigenvalues, that is, the roots (i.e., the solutions) of this equation are determined, the eigenvectors can be obtained one-by-one by inserting these eigenvalues one-by-one into Eq. (1.201).

For the sake of an example, consider the matrix

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}. \tag{1.203}$$

The secular equation is

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

yielding the characteristic equation $(1-\lambda)^3 - (1-\lambda) = (1-\lambda)[(1-\lambda)^2 - 1] = (1-\lambda)[\lambda^2 - 2\lambda] = -\lambda(1-\lambda)(2-\lambda) = 0$, and therefore three eigenvalues $\lambda_1 = 0$, $\lambda_2 = 1$, and $\lambda_3 = 2$ which are the roots of $\lambda(1-\lambda)(2-\lambda) = 0$.

Next let us determine the eigenvectors of A, based on the eigenvalues. Insertion $\lambda_1 = 0$ into Eq. (1.201) yields

$$\left[\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right] \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \tag{1.204}$$

therefore $x_1 + x_3 = 0$ and $x_2 = 0$. We are free to choose any (nonzero) $x_1 = -x_3$, but if we are interested in normalized eigenvectors, we obtain $\mathbf{x}_1 = (1/\sqrt{2})(1, 0, -1)^T$.

Insertion $\lambda_2 = 1$ into Eq. (1.201) yields

$$\left[\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right] \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \tag{1.205}$$

therefore $x_1 = x_3 = 0$ and x_2 is arbitrary. We are again free to choose any (nonzero) x_2, but if we are interested in normalized eigenvectors, we obtain $\mathbf{x}_2 = (0, 1, 0)^T$.

Insertion $\lambda_3 = 2$ into Eq. (1.201) yields

$$\left[\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} \right] \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \quad (1.206)$$

therefore $-x_1 + x_3 = 0$ and $x_2 = 0$. We are free to choose any (nonzero) $x_1 = x_3$, but if we are once more interested in normalized eigenvectors, we obtain $\mathbf{x}_3 = (1/\sqrt{2})(1, 0, 1)^T$.

Note that the eigenvectors are mutually orthogonal. We can construct the corresponding orthogonal projections by the outer (dyadic or tensor) product of the eigenvectors; that is,

$$\mathbf{E}_1 = \mathbf{x}_1 \otimes \mathbf{x}_1^T = \frac{1}{2}(1, 0, -1)^T (1, 0, -1) = \frac{1}{2} \begin{pmatrix} 1(1,0,-1) \\ 0(1,0,-1) \\ -1(1,0,-1) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$\mathbf{E}_2 = \mathbf{x}_2 \otimes \mathbf{x}_2^T = (0, 1, 0)^T (0, 1, 0) = \begin{pmatrix} 0(0,1,0) \\ 1(0,1,0) \\ 0(0,1,0) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{E}_3 = \mathbf{x}_3 \otimes \mathbf{x}_3^T = \frac{1}{2}(1, 0, 1)^T (1, 0, 1) = \frac{1}{2} \begin{pmatrix} 1(1,0,1) \\ 0(1,0,1) \\ 1(1,0,1) \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$
$$(1.207)$$

Note also that A can be written as the sum of the products of the eigenvalues with the associated projections; that is (here, $\mathbf{E}$ stands for the corresponding matrix), $A = 0\mathbf{E}_1 + 1\mathbf{E}_2 + 2\mathbf{E}_3$. Also, the projections are mutually orthogonal – that is, $\mathbf{E}_1\mathbf{E}_2 = \mathbf{E}_1\mathbf{E}_3 = \mathbf{E}_2\mathbf{E}_3 = 0$ – and add up to the identity; that is, $\mathbf{E}_1 + \mathbf{E}_2 + \mathbf{E}_3 = \mathbb{I}$.

If the eigenvalues obtained are not distinct und thus some eigenvalues are *degenerate*, the associated eigenvectors traditionally – that is, by convention and not necessity – are chosen to be *mutually orthogonal*. A more formal motivation will come from the spectral theorem below.

For the sake of an example, consider the matrix

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}. \quad (1.208)$$

The secular equation yields

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

which yields the characteristic equation $(2 - \lambda)(1 - \lambda)^2 + [-(2 - \lambda)] = (2 - \lambda)[(1 - \lambda)^2 - 1] = -\lambda(2 - \lambda)^2 = 0$, and therefore just two eigenvalues $\lambda_1 = 0$, and $\lambda_2 = 2$ which are the roots of $\lambda(2 - \lambda)^2 = 0$.

Let us now determine the eigenvectors of B, based on the eigenvalues. Insertion $\lambda_1 = 0$ into Eq. (1.201) yields

$$\left[\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}\right]\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \qquad (1.209)$$

therefore $x_1 + x_3 = 0$ and $x_2 = 0$. Again we are free to choose any (nonzero) $x_1 = -x_3$, but if we are interested in normalized eigenvectors, we obtain $\mathbf{x}_1 = (1/\sqrt{2})(1, 0, -1)^T$.

Insertion $\lambda_2 = 2$ into Eq. (1.201) yields

$$\left[\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} - \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}\right]\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}; \quad (1.210)$$

therefore $x_1 = x_3$; x_2 is arbitrary. We are again free to choose any values of x_1, x_3 and x_2 as long $x_1 = x_3$ as well as x_2 are satisfied. Take, for the sake of choice, the orthogonal normalized eigenvectors $\mathbf{x}_{2,1} = (0, 1, 0)^T$ and $\mathbf{x}_{2,2} = (1/\sqrt{2})(1, 0, 1)^T$, which are also orthogonal to $\mathbf{x}_1 = (1/\sqrt{2})(1, 0, -1)^T$.

Note again that we can find the corresponding orthogonal projections by the outer (dyadic or tensor) product of the eigenvectors; that is, by

$$\mathbf{E}_1 = \mathbf{x}_1 \otimes \mathbf{x}_1^T = \frac{1}{2}(1, 0, -1)^T(1, 0, -1) = \frac{1}{2}\begin{pmatrix} 1(1, 0, -1) \\ 0(1, 0, -1) \\ -1(1, 0, -1) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$$\mathbf{E}_{2,1} = \mathbf{x}_{2,1} \otimes \mathbf{x}_{2,1}^T = (0, 1, 0)^T(0, 1, 0) = \begin{pmatrix} 0(0, 1, 0) \\ 1(0, 1, 0) \\ 0(0, 1, 0) \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

$$\mathbf{E}_{2,2} = \mathbf{x}_{2,2} \otimes \mathbf{x}_{2,2}^T = \frac{1}{2}(1, 0, 1)^T(1, 0, 1) = \frac{1}{2}\begin{pmatrix} 1(1, 0, 1) \\ 0(1, 0, 1) \\ 1(1, 0, 1) \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

$$(1.211)$$

Note also that B can be written as the sum of the products of the eigenvalues with the associated projections; that is (here, $\mathbf{E}$ stands for the corresponding matrix), $B = 0\mathbf{E}_1 + 2(\mathbf{E}_{2,1} + \mathbf{E}_{2,2})$. Again, the projections are mutually orthogonal – that is, $\mathbf{E}_1\mathbf{E}_{2,1} = \mathbf{E}_1\mathbf{E}_{2,2} = \mathbf{E}_{2,1}\mathbf{E}_{2,2} = 0$ – and add up to the identity; that is, $\mathbf{E}_1 + \mathbf{E}_{2,1} + \mathbf{E}_{2,2} = \mathbb{I}$. This leads us to the much more general spectral theorem.

Another, extreme, example would be the unit matrix in n dimensions; that is, $\mathbb{I}_n = \mathrm{diag}(\underbrace{1, \ldots, 1}_{n \text{ times}})$, which has an n-fold degenerate eigenvalue 1 corresponding to a solution to $(1 - \lambda)^n = 0$. The corresponding projection operator is $\mathbb{I}_n$. [Note that $(\mathbb{I}_n)^2 = \mathbb{I}_n$ and thus $\mathbb{I}_n$ is a projection.] If one (somehow arbitrarily but conveniently) chooses a resolution of the identity operator $\mathbb{I}_n$ into projections corresponding to the standard basis (any

other orthonormal basis would do as well), then

$$\mathbb{I}_n = \mathrm{diag}(1,0,0,\ldots,0) + \mathrm{diag}(0,1,0,\ldots,0) + \cdots + \mathrm{diag}(0,0,0,\ldots,1)$$

$$
\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
  &   &   & \vdots &   \\
0 & 0 & 0 & \cdots & 1
\end{pmatrix}
=
\begin{pmatrix}
1 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
  &   &   & \vdots &   \\
0 & 0 & 0 & \cdots & 0
\end{pmatrix}
+
$$

$$
+
\begin{pmatrix}
0 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
  &   &   & \vdots &   \\
0 & 0 & 0 & \cdots & 0
\end{pmatrix}
+ \cdots +
\begin{pmatrix}
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
0 & 0 & 0 & \cdots & 0 \\
  &   &   & \vdots &   \\
0 & 0 & 0 & \cdots & 1
\end{pmatrix},
\tag{1.212}
$$

where all the matrices in the sum carrying one nonvanishing entry "1" in their diagonal are projections. Note that

$$
\begin{aligned}
\mathbf{e}_i &= |\mathbf{e}_i\rangle \\
&\equiv \Big(\underbrace{0,\ldots,0}_{i-1 \text{ times}}, 1, \underbrace{0,\ldots,0}_{n-i \text{ times}} \Big)^T \\
&\equiv \mathrm{diag}(\underbrace{0,\ldots,0}_{i-1 \text{ times}}, 1, \underbrace{0,\ldots,0}_{n-i \text{ times}}) \\
&\equiv \mathbf{E}_i.
\end{aligned}
\tag{1.213}
$$

The following theorems are enumerated without proofs.

If $\mathbf{A}$ is a self-adjoint transformation on an inner product space, then every proper value (eigenvalue) of $\mathbf{A}$ is real. If $\mathbf{A}$ is positive, or stricly positive, then every proper value of $\mathbf{A}$ is positive, or stricly positive, respectively

Due to their idempotence $\mathbf{EE} = \mathbf{E}$, projections have eigenvalues 0 or 1.

Every eigenvalue of an isometry has absolute value one.

If $\mathbf{A}$ is either a self-adjoint transformation or an isometry, then proper vectors of $\mathbf{A}$ belonging to distinct proper values are orthogonal.

1.27   Normal transformation

A transformation $\mathbf{A}$ is called *normal* if it commutes with its adjoint; that is,

$$[\mathbf{A}, \mathbf{A}^*] = \mathbf{A}\mathbf{A}^* - \mathbf{A}^*\mathbf{A} = 0. \tag{1.214}$$

It follows from their definition that Hermitian and unitary transformations are normal. That is, $\mathbf{A}^* = \mathbf{A}^\dagger$, and for Hermitian operators, $\mathbf{A} = \mathbf{A}^\dagger$, and thus $[\mathbf{A}, \mathbf{A}^\dagger] = \mathbf{A}\mathbf{A} - \mathbf{A}\mathbf{A} = (\mathbf{A})^2 - (\mathbf{A})^2 = 0$. For unitary operators, $\mathbf{A}^\dagger = \mathbf{A}^{-1}$, and thus $[\mathbf{A}, \mathbf{A}^\dagger] = \mathbf{A}\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{A} = \mathbb{I} - \mathbb{I} = 0$.

We mention without proof that a normal transformation on a finite-dimensional unitary space is (i) Hermitian, (ii) positive, (iii) strictly positive, (iv) unitary, (v) invertible, (vi) idempotent if and only if all its proper values are (i) real, (ii) positive, (iii) strictly positive, (iv) of absolute value one, (v) different from zero, (vi) equal to zero or one.

1.28 Spectrum

1.28.1 Spectral theorem

Let $\mathfrak{V}$ be an n-dimensional linear vector space. The *spectral theorem* states that to every self-adjoint (more general, normal) transformation **A** on an n-dimensional inner product space there correspond real numbers, the *spectrum* $\lambda_1, \lambda_2, \ldots, \lambda_k$ of all the eigenvalues of **A**, and their associated orthogonal projections $\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_k$ where $0 < k \leq n$ is a strictly positive integer so that

(i) the λ_i are pairwise distinct;

(ii) the $\mathbf{E}_i$ are pairwise orthogonal and different from **0**;

(iii) the set of projectors is complete in the sense that their sum $\sum_{i=1}^{k} \mathbf{E}_i = \mathbb{I}_n$ is a resolution of the identity operator; and

(iv) $\mathbf{A} = \sum_{i=1}^{k} \lambda_i \mathbf{E}_i$ is the *spectral form* of **A**.

Rather than proving the spectral theorem in its full generality, we suppose that the spectrum of a Hermitian (self-adjoint) operator **A** is *nondegenerate*; that is, all n eigenvalues of **A** are pairwise distinct. That is, we are assuming a strong form of (i), which is satisfied by definition.

This distinctness of the eigenvalues then translates into mutual orthogonality of all the eigenvectors of **A**. Thereby, the set of n eigenvectors forms some orthogonal (orthonormal) basis of the n-dimensional linear vector space $\mathfrak{V}$. The respective normalized eigenvectors can then be represented by perpendicular projections which can be summed up to yield the identity (iii).

More explicitly, suppose, for the sake of a proof (by contradiction) of the pairwise orthogonality of the eigenvectors (ii), that two distinct eigenvalues λ_1 and $\lambda_2 \neq \lambda_1$ belong to two respective eigenvectors $|\mathbf{x}_1\rangle$ and $|\mathbf{x}_2\rangle$ which are not orthogonal. But then, because **A** is self-adjoint with real eigenvalues,

$$\lambda_1 \langle \mathbf{x}_1 | \mathbf{x}_2 \rangle = \langle \lambda_1 \mathbf{x}_1 | \mathbf{x}_2 \rangle = \langle \mathbf{A} \mathbf{x}_1 | \mathbf{x}_2 \rangle$$
$$= \langle \mathbf{x}_1 | \mathbf{A}^* \mathbf{x}_2 \rangle = \langle \mathbf{x}_1 | \mathbf{A} \mathbf{x}_2 \rangle = \langle \mathbf{x}_1 | \lambda_2 \mathbf{x}_2 \rangle = \lambda_2 \langle \mathbf{x}_1 | \mathbf{x}_2 \rangle, \tag{1.215}$$

which implies that

$$(\lambda_1 - \lambda_2) \langle \mathbf{x}_1 | \mathbf{x}_2 \rangle = 0. \tag{1.216}$$

Equation (1.216) is satisfied by either $\lambda_1 = \lambda_2$ – which is in contradiction to our assumption that λ_1 and λ_2 are distinct – or by $\langle \mathbf{x}_1 | \mathbf{x}_2 \rangle = 0$ (thus allowing $\lambda_1 \neq \lambda_2$) – which is in contradiction to our assumption that $|\mathbf{x}_1\rangle$ and $|\mathbf{x}_2\rangle$ are nonzero and *not* orthogonal. Hence, if we maintain the distinctness of λ_1 and λ_2, the associated eigenvectors need to be orthogonal, thereby assuring (ii).

Since by our assumption there are n distinct eigenvalues, this implies that, associated with these, there are n orthonormal eigenvectors. These n mutually orthonormal eigenvectors span the entire n-dimensional vector space $\mathfrak{V}$; and hence their union $\{|\mathbf{x}_i\rangle, \ldots, |\mathbf{x}_n\rangle\}$ forms an orthonormal basis. Consequently, the sum of the associated perpendicular projections

$\mathbf{E}_i = \frac{|\mathbf{x}_i\rangle\langle\mathbf{x}_i|}{\langle\mathbf{x}_i|\mathbf{x}_i\rangle}$ is a resolution of the identity operator $\mathbb{I}_n$ (cf. section 1.15 on page 36); thereby justifying (iii).

In the last step, let us define the i'th projection of an arbitrary vector $|\mathbf{z}\rangle \in \mathfrak{V}$ by $|\xi_i\rangle = \mathbf{E}_i|\mathbf{z}\rangle = |\mathbf{x}_i\rangle\langle\mathbf{x}_i|\mathbf{z}\rangle = \alpha_i|\mathbf{x}_i\rangle$ with $\alpha_i = \langle\mathbf{x}_i|\mathbf{z}\rangle$, thereby keeping in mind that this vector is an eigenvector of $\mathbf{A}$ with the associated eigenvalue λ_i; that is,

$$\mathbf{A}|\xi_i\rangle = \mathbf{A}\alpha_i|\mathbf{x}_i\rangle = \alpha_i\mathbf{A}|\mathbf{x}_i\rangle = \alpha_i\lambda_i|\mathbf{x}_i\rangle = \lambda_i\alpha_i|\mathbf{x}_i\rangle = \lambda_i|\xi_i\rangle. \qquad (1.217)$$

Then,

$$\mathbf{A}|\mathbf{z}\rangle = \mathbf{A}\mathbb{I}_n|\mathbf{z}\rangle = \mathbf{A}\left(\sum_{i=1}^{n}\mathbf{E}_i\right)|\mathbf{z}\rangle = \mathbf{A}\left(\sum_{i=1}^{n}\mathbf{E}_i|\mathbf{z}\rangle\right) =$$

$$= \mathbf{A}\left(\sum_{i=1}^{n}|\xi_i\rangle\right) = \sum_{i=1}^{n}\mathbf{A}|\xi_i\rangle = \sum_{i=1}^{n}\lambda_i|\xi_i\rangle = \sum_{i=1}^{n}\lambda_i\mathbf{E}_i|\mathbf{z}\rangle = \left(\sum_{i=1}^{n}\lambda_i\mathbf{E}_i\right)|\mathbf{z}\rangle,$$

$$(1.218)$$

which is the spectral form of $\mathbf{A}$.

1.28.2   Composition of the spectral form

If the spectrum of a Hermitian (or, more general, normal) operator $\mathbf{A}$ is nondegenerate, that is, $k = n$, then the ith projection can be written as the outer (dyadic or tensor) product $\mathbf{E}_i = \mathbf{x}_i \otimes \mathbf{x}_i^T$ of the ith normalized eigenvector $\mathbf{x}_i$ of $\mathbf{A}$. In this case, the set of all normalized eigenvectors $\{\mathbf{x}_1,\ldots,\mathbf{x}_n\}$ is an orthonormal basis of the vector space $\mathfrak{V}$. If the spectrum of $\mathbf{A}$ is degenerate, then the projection can be chosen to be the orthogonal sum of projections corresponding to orthogonal eigenvectors, associated with the same eigenvalues.

Furthermore, for a Hermitian (or, more general, normal) operator $\mathbf{A}$, if $1 \le i \le k$, then there exist polynomials with real coefficients, such as, for instance,

$$p_i(t) = \prod_{\substack{1 \le j \le k \\ j \ne i}} \frac{t - \lambda_j}{\lambda_i - \lambda_j} \qquad (1.219)$$

so that $p_i(\lambda_j) = \delta_{ij}$; moreover, for every such polynomial, $p_i(\mathbf{A}) = \mathbf{E}_i$.

For a proof, it is not too difficult to show that $p_i(\lambda_i) = 1$, since in this case in the product of fractions all numerators are equal to denominators, and $p_i(\lambda_j) = 0$ for $j \ne i$, since some numerator in the product of fractions vanishes.

Now, substituting for t the spectral form $\mathbf{A} = \sum_{i=1}^{k}\lambda_i\mathbf{E}_i$ of $\mathbf{A}$, as well as resolution of the identity operator in terms of the projections $\mathbf{E}_i$ in the spectral form of $\mathbf{A}$; that is, $\mathbb{I}_n = \sum_{i=1}^{k}\mathbf{E}_i$, yields

$$p_i(\mathbf{A}) = \prod_{1 \le j \le k,\, j \ne i} \frac{\mathbf{A} - \lambda_j\mathbb{I}_n}{\lambda_i - \lambda_j} = \prod_{1 \le j \le k,\, j \ne i} \frac{\sum_{l=1}^{k}\lambda_l\mathbf{E}_l - \lambda_j\sum_{l=1}^{k}\mathbf{E}_l}{\lambda_i - \lambda_j}, \qquad (1.220)$$

and, because of the idempotence and pairwise orthogonality of the projections $\mathbf{E}_l$,

$$= \prod_{1 \le j \le k,\, j \ne i} \frac{\sum_{l=1}^{k}\mathbf{E}_l(\lambda_l - \lambda_j)}{\lambda_i - \lambda_j}$$

$$= \sum_{l=1}^{k}\mathbf{E}_l \prod_{1 \le j \le k,\, j \ne i} \frac{\lambda_l - \lambda_j}{\lambda_i - \lambda_j} = \sum_{l=1}^{k}\mathbf{E}_l\delta_{li} = \mathbf{E}_i.$$

$$(1.221)$$

With the help of the polynomial $p_i(t)$ defined in Eq. (1.219), which requires knowledge of the eigenvalues, the spectral form of a Hermitian (or, more general, normal) operator $\mathbf{A}$ can thus be rewritten as

$$\mathbf{A} = \sum_{i=1}^{k} \lambda_i p_i(\mathbf{A}) = \sum_{i=1}^{k} \lambda_i \prod_{1 \le j \le k,\ j \ne i} \frac{\mathbf{A} - \lambda_j \mathbb{1}_n}{\lambda_i - \lambda_j}. \qquad (1.222)$$

That is, knowledge of all the eigenvalues entails construction of all the projections in the spectral decomposition of a normal transformation.

For the sake of an example, consider again the matrix

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad (1.223)$$

and the associated Eigensystem

$$\{\{\lambda_1, \lambda_2, \lambda_3\}, \{\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3\}\}$$

$$= \left\{ \{0, 1, 2\}, \left\{ \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \right\} \right\}. \qquad (1.224)$$

The projections associated with the eigenvalues, and, in particular, $\mathbf{E}_1$, can be obtained from the set of eigenvalues $\{0, 1, 2\}$ by

$$p_1(A) = \left(\frac{A - \lambda_2 \mathbb{1}}{\lambda_1 - \lambda_2} \right) \left(\frac{A - \lambda_3 \mathbb{1}}{\lambda_1 - \lambda_3} \right)$$

$$= \frac{\left(\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} - 1 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right)}{(0 - 1)} \cdot \frac{\left(\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right)}{(0 - 2)} \qquad (1.225)$$

$$= \frac{1}{2} \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} -1 & 0 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \mathbf{E}_1.$$

For the sake of another, degenerate, example consider again the matrix

$$B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad (1.226)$$

Again, the projections $\mathbf{E}_1, \mathbf{E}_2$ can be obtained from the set of eigenvalues $\{0, 2\}$ by

$$p_1(A) = \frac{A - \lambda_2 \mathbb{1}}{\lambda_1 - \lambda_2} = \frac{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}{(0 - 2)} = \frac{1}{2} \begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{pmatrix} = \mathbf{E}_1,$$

$$p_2(A) = \frac{A - \lambda_1 \mathbb{1}}{\lambda_2 - \lambda_1} = \frac{\begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} - 0 \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}{(2 - 0)} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 1 \end{pmatrix} = \mathbf{E}_2.$$

$$(1.227)$$

Note that, in accordance with the spectral theorem, $\mathbf{E}_1\mathbf{E}_2 = 0$, $\mathbf{E}_1 + \mathbf{E}_2 = \mathbb{I}$ and $0 \cdot \mathbf{E}_1 + 2 \cdot \mathbf{E}_2 = B$.

1.29 Functions of normal transformations

Suppose $\mathbf{A} = \sum_{i=1}^{k} \lambda_i \mathbf{E}_i$ is a normal transformation in its spectral form. If f is an arbitrary complex-valued function defined at least at the eigenvalues of $\mathbf{A}$, then a linear transformation $f(\mathbf{A})$ can be defined by

$$f(\mathbf{A}) = f\left(\sum_{i=1}^{k} \lambda_i \mathbf{E}_i\right) = \sum_{i=1}^{k} f(\lambda_i)\mathbf{E}_i. \tag{1.228}$$

Note that, if f has a polynomial expansion such as analytic functions, then orthogonality and idempotence of the projections $\mathbf{E}_i$ in the spectral form guarantees this kind of "linearization."

For the definition of the "square root" for every positve operator $\mathbf{A}$), consider

$$\sqrt{\mathbf{A}} = \sum_{i=1}^{k} \sqrt{\lambda_i}\mathbf{E}_i. \tag{1.229}$$

With this definition, $\left(\sqrt{\mathbf{A}}\right)^2 = \sqrt{\mathbf{A}}\sqrt{\mathbf{A}} = \mathbf{A}$.

Consider, for instance, the "square root" of the **not** operator

$$\mathbf{not} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}. \tag{1.230}$$

The denomination "not" for **not** can be motivated by enumerating its performance at the two "classical bit states" $|0\rangle \equiv (1,0)^T$ and $|1\rangle \equiv (0,1)^T$: $\mathbf{not}|0\rangle = |1\rangle$ and $\mathbf{not}|1\rangle = |0\rangle$.

To enumerate $\sqrt{\mathbf{not}}$ we need to find the *spectral form* of **not** first. The eigenvalues of **not** can be obtained by solving the secular equation

$$\det\left(\mathbf{not} - \lambda\mathbb{I}_2\right) = \det\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - \lambda\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) = \det\begin{pmatrix} -\lambda & 1 \\ 1 & -\lambda \end{pmatrix} = \lambda^2 - 1 = 0. \tag{1.231}$$

$\lambda^2 = 1$ yields the two eigenvalues $\lambda_1 = 1$ and $\lambda_1 = -1$. The associated eigenvectors $\mathbf{x}_1$ and $\mathbf{x}_2$ can be derived from either the equations $\mathbf{not}\,\mathbf{x}_1 = \mathbf{x}_1$ and $\mathbf{not}\,\mathbf{x}_2 = -\mathbf{x}_2$, or inserting the eigenvalues into the polynomial (1.219).

We choose the former method. Thus, for $\lambda_1 = 1$,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_{1,1} \\ x_{1,2} \end{pmatrix} = \begin{pmatrix} x_{1,1} \\ x_{1,2} \end{pmatrix}, \tag{1.232}$$

which yields $x_{1,1} = x_{1,2}$, and thus, by normalizing the eigenvector, $\mathbf{x}_1 = (1/\sqrt{2})(1,1)^T$. The associated projection is

$$\mathbf{E}_1 = \mathbf{x}_1\mathbf{x}_1^T = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}. \tag{1.233}$$

Likewise, for $\lambda_2 = -1$,

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} x_{2,1} \\ x_{2,2} \end{pmatrix} = -\begin{pmatrix} x_{2,1} \\ x_{2,2} \end{pmatrix}, \tag{1.234}$$

which yields $x_{2,1} = -x_{2,2}$, and thus, by normalizing the eigenvector, $\mathbf{x}_2 = (1/\sqrt{2})(1,-1)^T$. The associated projection is

$$\mathbf{E}_2 = \mathbf{x}_2\mathbf{x}_2^T = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{1.235}$$

Thus we are finally able to calculate $\sqrt{\textbf{not}}$ through its spectral form

$$\sqrt{\textbf{not}} = \sqrt{\lambda_1}\textbf{E}_1 + \sqrt{\lambda_2}\textbf{E}_2 =$$

$$= \sqrt{1}\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \sqrt{-1}\frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \qquad (1.236)$$

$$= \frac{1}{2}\begin{pmatrix} 1+i & 1-i \\ 1-i & 1+i \end{pmatrix} = \frac{1}{1-i}\begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix}.$$

It can be readily verified that $\sqrt{\textbf{not}}\sqrt{\textbf{not}} = \textbf{not}$. Note that $\pm_1\sqrt{\lambda_1}\textbf{E}_1 + \pm_2\sqrt{\lambda_2}\textbf{E}_2$, where $\pm_1$ and $\pm_2$ represent separate cases, yield alternative expressions of $\sqrt{\textbf{not}}$.

1.30 Decomposition of operators

1.30.1 Standard decomposition

In analogy to the decomposition of every imaginary number $z = \Re z + i\Im z$ with $\Re z, \Im z \in \mathbb{R}$, every arbitrary transformation $\textbf{A}$ on a finite-dimensional vector space can be decomposed into two Hermitian operators $\textbf{B}$ and $\textbf{C}$ such that

$$\textbf{A} = \textbf{B} + i\textbf{C}; \text{ with}$$

$$\textbf{B} = \frac{1}{2}(\textbf{A} + \textbf{A}^\dagger),$$

$$\textbf{C} = \frac{1}{2i}(\textbf{A} - \textbf{A}^\dagger). \qquad (1.237)$$

Proof by insertion; that is,

$$\textbf{A} = \textbf{B} + i\textbf{C}$$

$$= \frac{1}{2}(\textbf{A} + \textbf{A}^\dagger) + i\left[\frac{1}{2i}(\textbf{A} - \textbf{A}^\dagger)\right],$$

$$\textbf{B}^\dagger = \left[\frac{1}{2}(\textbf{A} + \textbf{A}^\dagger)\right]^\dagger = \frac{1}{2}\left[\textbf{A}^\dagger + (\textbf{A}^\dagger)^\dagger\right]$$

$$= \frac{1}{2}\left[\textbf{A}^\dagger + \textbf{A}\right] = \textbf{B}, \qquad (1.238)$$

$$\textbf{C}^\dagger = \left[\frac{1}{2i}(\textbf{A} - \textbf{A}^\dagger)\right]^\dagger = -\frac{1}{2i}\left[\textbf{A}^\dagger - (\textbf{A}^\dagger)^\dagger\right]$$

$$= -\frac{1}{2i}\left[\textbf{A}^\dagger - \textbf{A}\right] = \textbf{C}.$$

1.30.2 Polar representation

In analogy to the polar representation of every imaginary number $z = Re^{i\varphi}$ with $R, \varphi \in \mathbb{R}$, $R > 0$, $0 \le \varphi < 2\pi$, every arbitrary transformation $\textbf{A}$ on a finite-dimensional inner product space can be decomposed into a unique positive transform $\textbf{P}$ and an isometry $\textbf{U}$, such that $\textbf{A} = \textbf{UP}$. If $\textbf{A}$ is invertible, then $\textbf{U}$ is uniquely determined by $\textbf{A}$. A necessary and sufficient condition that $\textbf{A}$ is normal is that $\textbf{UP} = \textbf{PU}$.

1.30.3 Decomposition of isometries

Any unitary or orthogonal transformation in finite-dimensional inner product space can be composed from a succession of two-parameter

unitary transformations in two-dimensional subspaces, and a multiplication of a single diagonal matrix with elements of modulus one in an algorithmic, constructive and tractable manner. The method is similar to Gaussian elimination and facilitates the parameterization of elements of the unitary group in arbitrary dimensions (e.g., Ref. [25], Chapter 2).

It has been suggested to implement these group theoretic results by realizing interferometric analogues of any discrete unitary and Hermitian operator in a unified and experimentally feasible way by "generalized beam splitters" [26].

1.30.4   Singular value decomposition

The *singular value decomposition* (SVD) of an $(m \times n)$ matrix **A** is a factorization of the form

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}, \tag{1.239}$$

where **U** is a unitary $(m \times m)$ matrix (i.e. an isometry), **V** is a unitary $(n \times n)$ matrix, and Σ is a unique $(m \times n)$ diagonal matrix with nonnegative real numbers on the diagonal; that is,

$$\Sigma = \begin{pmatrix} \sigma_1 & & & | & & \vdots & \\ & \ddots & & | & \cdots & 0 & \cdots \\ & & \sigma_r & | & & \vdots & \\ - & - & - & & - & - & - \\ & \vdots & & | & & \vdots & \\ \cdots & 0 & \cdots & | & \cdots & 0 & \cdots \\ & \vdots & & | & & \vdots & \end{pmatrix}. \tag{1.240}$$

The entries $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_r > 0$ of Σ are called *singular values* of **A**. No proof is presented here.

1.30.5   Schmidt decomposition of the tensor product of two vectors

Let $\mathfrak{U}$ and $\mathfrak{V}$ be two linear vector spaces of dimension $n \geq m$ and m, respectively. Then, for any vector $\mathbf{z} \in \mathfrak{U} \otimes \mathfrak{V}$ in the tensor product space, there exist orthonormal basis sets of vectors $\{\mathbf{u}_1, \ldots, \mathbf{u}_n\} \subset \mathfrak{U}$ and $\{\mathbf{v}_1, \ldots, \mathbf{v}_m\} \subset \mathfrak{V}$ such that $\mathbf{z} = \sum_{i=1}^m \sigma_i \mathbf{u}_i \otimes \mathbf{v}_i$, where the σ_is are nonnegative scalars and the set of scalars is uniquely determined by $\mathbf{z}$.

Equivalently [27], suppose that $|\mathbf{z}\rangle$ is some tensor product contained in the set of all tensor products of vectors $\mathfrak{U} \otimes \mathfrak{V}$ of two linear vector spaces $\mathfrak{U}$ and $\mathfrak{V}$. Then there exist orthonormal vectors $|\mathbf{u}_i\rangle \in \mathfrak{U}$ and $|\mathbf{v}_j\rangle \in \mathfrak{V}$ so that

$$|\mathbf{z}\rangle = \sum_i \sigma_i |\mathbf{u}_i\rangle |\mathbf{v}_i\rangle, \tag{1.241}$$

where the σ_is are nonnegative scalars; if $|\mathbf{z}\rangle$ is normalized, then the σ_is are satisfying $\sum_i \sigma_i^2 = 1$; they are called the *Schmidt coefficients*.

For a proof by reduction to the singular value decomposition, let $|i\rangle$ and $|j\rangle$ be any two fixed orthonormal bases of $\mathfrak{U}$ and $\mathfrak{V}$, respectively. Then, $|\mathbf{z}\rangle$ can be expanded as $|\mathbf{z}\rangle = \sum_{ij} a_{ij}|i\rangle|j\rangle$, where the a_{ij}s can be interpreted as the components of a matrix **A**. **A** can then be

[25] Francis D. Murnaghan. *The Unitary and Rotation Groups*, volume 3 of *Lectures on Applied Mathematics*. Spartan Books, Washington, D.C., 1962

[26] Michael Reck, Anton Zeilinger, Herbert J. Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. *Physical Review Letters*, 73:58–61, 1994. DOI: 10.1103/PhysRevLett.73.58. URL http://doi.org/10.1103/PhysRevLett.73.58; and Michael Reck and Anton Zeilinger. Quantum phase tracing of correlated photons in optical multiports. In F. De Martini, G. Denardo, and Anton Zeilinger, editors, *Quantum Interferometry*, pages 170–177, Singapore, 1994. World Scientific

[27] Michael A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, Cambridge, 2000

subjected to a singular value decomposition $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}$, or, written in index form [note that $\Sigma = \mathrm{diag}(\sigma_1,\ldots,\sigma_n)$ is a diagonal matrix], $a_{ij} = \sum_l u_{il}\sigma_l v_{lj}$; and hence $|\mathbf{z}\rangle = \sum_{ijl} u_{il}\sigma_l v_{lj}|i\rangle|j\rangle$. Finally, by identifying $|\mathbf{u}_l\rangle = \sum_i u_{il}|i\rangle$ as well as $|\mathbf{v}_l\rangle = \sum_l v_{lj}|j\rangle$ one obtains the Schmidt decompsition (1.241). Since u_{il} and v_{lj} represent unitary martices, and because $|i\rangle$ as well as $|j\rangle$ are orthonormal, the newly formed vectors $|\mathbf{u}_l\rangle$ as well as $|\mathbf{v}_l\rangle$ form orthonormal bases as well. The sum of squares of the σ_i's is one if $|\mathbf{z}\rangle$ is a unit vector, because (note that σ_is are real-valued) $\langle\mathbf{z}|\mathbf{z}\rangle = 1 = \sum_{lm}\sigma_l\sigma_m\langle\mathbf{u}_l|\mathbf{u}_m\rangle\langle\mathbf{v}_l|\mathbf{v}_m\rangle = \sum_{lm}\sigma_l\sigma_m\delta_{lm} = \sum_l\sigma_l^2$.

Note that the Schmidt decomposition cannot, in general, be extended for more factors than two. Note also that the Schmidt decomposition needs not be unique [28]; in particular if some of the Schmidt coefficients σ_i are equal. For the sake of an example of nonuniqueness of the Schmidt decomposition, take, for instance, the representation of the *Bell state* with the two bases

[28] Artur Ekert and Peter L. Knight. Entangled quantum systems and the Schmidt decomposition. *American Journal of Physics*, 63(5):415–423, 1995. DOI: 10.1119/1.17904. URL http://doi.org/10.1119/1.17904

$$\left\{|\mathbf{e}_1\rangle \equiv (1,0)^T, |\mathbf{e}_2\rangle \equiv (0,1)^T\right\} \text{ and}$$
$$\left\{|\mathbf{f}_1\rangle \equiv \frac{1}{\sqrt{2}}(1,1)^T, |\mathbf{f}_2\rangle \equiv \frac{1}{\sqrt{2}}(-1,1)^T\right\}. \tag{1.242}$$

as follows:

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}\left(|\mathbf{e}_1\rangle|\mathbf{e}_2\rangle - |\mathbf{e}_2\rangle|\mathbf{e}_1\rangle\right)$$
$$\equiv \frac{1}{\sqrt{2}}\left[(1(0,1),0(0,1))^T - (0(1,0),1(1,0))^T\right] = \frac{1}{\sqrt{2}}(0,1,-1,0)^T;$$
$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}\left(|\mathbf{f}_1\rangle|\mathbf{f}_2\rangle - |\mathbf{f}_2\rangle|\mathbf{f}_1\rangle\right) \tag{1.243}$$
$$\equiv \frac{1}{2\sqrt{2}}\left[(1(-1,1),1(-1,1))^T - (-1(1,1),1(1,1))^T\right]$$
$$\equiv \frac{1}{2\sqrt{2}}\left[(-1,1,-1,1)^T - (-1,-1,1,1)^T\right] = \frac{1}{\sqrt{2}}(0,1,-1,0)^T.$$

1.31   Purification

In general, quantum states $\boldsymbol{\rho}$ satisfy two criteria [29]: they are (i) of trace class one: $\mathrm{Tr}(\boldsymbol{\rho}) = 1$; and (ii) positive (or, by another term nonnegative): $\langle\mathbf{x}|\boldsymbol{\rho}|\mathbf{x}\rangle = \langle\mathbf{x}|\boldsymbol{\rho}\mathbf{x}\rangle \geq 0$ for all vectors $\mathbf{x}$ of the Hilbert space.

For additional information see page 110, Sect. 2.5 in

Michael A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2010. 10th Anniversary Edition

[29] L. E. Ballentine. *Quantum Mechanics.* Prentice Hall, Englewood Cliffs, NJ, 1989

With finite dimension n it follows immediately from (ii) that $\boldsymbol{\rho}$ is self-adjoint; that is, $\boldsymbol{\rho}^\dagger = \boldsymbol{\rho}$), and normal, and thus has a spectral decomposition

$$\boldsymbol{\rho} = \sum_{i=1}^n \rho_i|\psi_i\rangle\langle\psi_i| \tag{1.244}$$

into orthogonal projections $|\psi_i\rangle\langle\psi_i|$, with (i) yielding $\sum_{i=1}^n \rho_i = 1$ (hint: take a trace with the orthonormal basis corresponding to all the $|\psi_i\rangle$); (ii) yielding $\overline{\rho_i} = \rho_i$; and (iii) implying $\rho_i \geq 0$, and hence [with (i)] $0 \leq \rho_i \leq 1$ for all $1 \leq i \leq n$.

As has been pointed out earlier, quantum mechanics differentiates between "two sorts of states," namely pure states and mixed ones:

(i)   Pure states $\boldsymbol{\rho}_p$ are represented by one-dimensional orthogonal projections; or, equivalently as one-dimensional linear subspaces by

some (unit) vector. They can be written as $\boldsymbol{\rho}_p = |\psi\rangle\langle\psi|$ for some unit vector $|\psi\rangle$ (discussed in Sec. 1.12), and satisfy $(\boldsymbol{\rho}_p)^2 = \boldsymbol{\rho}_p$.

(ii) General, mixed states $\boldsymbol{\rho}_m$, are ones that are no projections and therefore satisfy $(\boldsymbol{\rho}_m)^2 \neq \boldsymbol{\rho}_m$. They can be composed from projections by their spectral form (1.244).

The question arises: is it possible to "purify" any mixed state by (maybe somewhat superficially) "enlarging" its Hilbert space, such that the resulting state "living in a larger Hilbert space" is pure? This can indeed be achieved by a rather simple procedure: By considering the spectral form (1.244) of a general mixed state $\boldsymbol{\rho}$, define a new, "enlarged," pure state $|\Psi\rangle\langle\Psi|$, with

$$|\Psi\rangle = \sum_{i=1}^{n} \sqrt{\rho_i}|\psi_i\rangle|\psi_i\rangle. \tag{1.245}$$

That $|\Psi\rangle\langle\Psi|$ is pure can be tediously verified by proving that it is idempotent:

$$(|\Psi\rangle\langle\Psi|)^2$$

$$= \left\{\left[\sum_{i=1}^{n} \sqrt{\rho_i}|\psi_i\rangle|\psi_i\rangle\right]\left[\sum_{j=1}^{n} \sqrt{\rho_j}\langle\psi_j|\langle\psi_j|\right]\right\}^2$$

$$= \left[\sum_{i_1=1}^{n} \sqrt{\rho_{i_1}}|\psi_{i_1}\rangle|\psi_{i_1}\rangle\right]\left[\sum_{j_1=1}^{n} \sqrt{\rho_{j_1}}\langle\psi_{j_1}|\langle\psi_{j_1}|\right] \times$$

$$\left[\sum_{i_2=1}^{n} \sqrt{\rho_{i_2}}|\psi_{i_2}\rangle|\psi_{i_2}\rangle\right]\left[\sum_{j_2=1}^{n} \sqrt{\rho_{j_2}}\langle\psi_{j_2}|\langle\psi_{j_2}|\right]$$

$$= \left[\sum_{i_1=1}^{n} \sqrt{\rho_{i_1}}|\psi_{i_1}\rangle|\psi_{i_1}\rangle\right]\underbrace{\left[\sum_{j_1=1}^{n}\sum_{i_2=1}^{n} \sqrt{\rho_{j_1}}\sqrt{\rho_{i_2}}(\delta_{i_2 j_1})^2\right]}_{\sum_{j_1=1}^{n} \rho_{j_1} = 1}\left[\sum_{j_2=1}^{n} \sqrt{\rho_{j_2}}\langle\psi_{j_2}|\langle\psi_{j_2}|\right]$$

$$= \left[\sum_{i_1=1}^{n} \sqrt{\rho_{i_1}}|\psi_{i_1}\rangle|\psi_{i_1}\rangle\right]\left[\sum_{j_2=1}^{n} \sqrt{\rho_{j_2}}\langle\psi_{j_2}|\langle\psi_{j_2}|\right]$$

$$= |\Psi\rangle\langle\Psi|. \tag{1.246}$$

Note that this construction is not unique – any construction $|\Psi'\rangle = \sum_{i=1}^{n} \sqrt{\rho_i}|\psi_i\rangle|\phi_i\rangle$ involving auxiliary components $|\phi_i\rangle$ representing the elements of some orthonormal basis $\{|\phi_1\rangle,\ldots,|\phi_n\rangle\}$ would suffice.

The original mixed state $\boldsymbol{\rho}$ is obtained from the pure state (1.245) corresponding to the unit vector $|\Psi\rangle = |\psi\rangle|\psi^a\rangle = |\psi\psi^a\rangle$ – we might say that "the superscript a stands for auxiliary" – by a partial trace (cf. Sec. 1.18.3) over one of its components, say $|\psi^a\rangle$.

For the sake of a proof let us "trace out of the auxiliary components $|\psi^a\rangle$," that is, take the trace

$$\mathrm{Tr}_a(|\Psi\rangle\langle\Psi|) = \sum_{k=1}^{n} \langle\psi_k^a|(|\Psi\rangle\langle\Psi|)|\psi_k^a\rangle \tag{1.247}$$

of $|\Psi\rangle\langle\Psi|$ with respect to one of its components $|\psi^a\rangle$:

$$
\begin{aligned}
&\mathrm{Tr}_a \left(|\Psi\rangle\langle\Psi|\right) \\
&= \mathrm{Tr}_a \left(\left[\sum_{i=1}^n \sqrt{\rho_i}|\psi_i\rangle|\psi_i^a\rangle\right]\left[\sum_{j=1}^n \sqrt{\rho_j}\langle\psi_j^a|\langle\psi_j|\right]\right) \\
&= \sum_{k=1}^n \left\langle\psi_k^a\left|\left[\sum_{i=1}^n \sqrt{\rho_i}|\psi_i\rangle|\psi_i^a\rangle\right]\left[\sum_{j=1}^n \sqrt{\rho_j}\langle\psi_j^a|\langle\psi_j|\right]\right|\psi_k^a\right\rangle \\
&= \sum_{k=1}^n \sum_{i=1}^n \sum_{j=1}^n \delta_{ki}\delta_{kj}\sqrt{\rho_i}\sqrt{\rho_j}|\psi_i\rangle\langle\psi_j| \\
&= \sum_{k=1}^n \rho_k|\psi_k\rangle\langle\psi_k| = \boldsymbol{\rho}.
\end{aligned}
\tag{1.248}
$$

1.32   Commutativity

If $\mathbf{A} = \sum_{i=1}^k \lambda_i \mathbf{E}_i$ is the spectral form of a self-adjoint transformation $\mathbf{A}$ on a finite-dimensional inner product space, then a necessary and sufficient condition ("if and only if = iff") that a linear transformation $\mathbf{B}$ commutes with $\mathbf{A}$ is that it commutes with each $\mathbf{E}_i$, $1 \le i \le k$.

Sufficiency is derived easily: whenever $\mathbf{B}$ commutes with all the projectors $\mathbf{E}_i$, $1 \le i \le k$ in the spectral decomposition $\mathbf{A} = \sum_{i=1}^k \lambda_i \mathbf{E}_i$ of $\mathbf{A}$, then it commutes with $\mathbf{A}$; that is,

$$
\begin{aligned}
\mathbf{BA} &= \mathbf{B}\left(\sum_{i=1}^k \lambda_i \mathbf{E}_i\right) = \sum_{i=1}^k \lambda_i \mathbf{BE}_i = \\
&= \sum_{i=1}^k \lambda_i \mathbf{E}_i \mathbf{B} = \left(\sum_{i=1}^k \lambda_i \mathbf{E}_i\right)\mathbf{B} = \mathbf{AB}.
\end{aligned}
\tag{1.249}
$$

Necessity follows from the fact that, if $\mathbf{B}$ commutes with $\mathbf{A}$ then it also commutes with every polynomial of $\mathbf{A}$, since in this case $\mathbf{AB} = \mathbf{BA}$, and thus $\mathbf{A}^m\mathbf{B} = \mathbf{A}^{m-1}\mathbf{AB} = \mathbf{A}^{m-1}\mathbf{BA} = \ldots = \mathbf{BA}^m$. In particular, it commutes with the polynomial $p_i(\mathbf{A}) = \mathbf{E}_i$ defined by Eq. (1.219).

If $\mathbf{A} = \sum_{i=1}^k \lambda_i \mathbf{E}_i$ and $\mathbf{B} = \sum_{j=1}^l \mu_j \mathbf{F}_j$ are the spectral forms of a self-adjoint transformations $\mathbf{A}$ and $\mathbf{B}$ on a finite-dimensional inner product space, then a necessary and sufficient condition ("if and only if = iff") that $\mathbf{A}$ and $\mathbf{B}$ commute is that the projections $\mathbf{E}_i$, $1 \le i \le k$ and $\mathbf{F}_j$, $1 \le j \le l$ commute with each other; i.e., $[\mathbf{E}_i, \mathbf{F}_j] = \mathbf{E}_i\mathbf{F}_j - \mathbf{F}_j\mathbf{E}_i = 0$.

Again, sufficiency can be derived as follows: suppose all projection operators $\mathbf{F}_j$, $1 \le j \le l$ occurring in the spectral decomposition of $\mathbf{B}$ commute with all projection operators $\mathbf{E}_i$, $1 \le i \le k$ in the spectral composition of $\mathbf{A}$, then

$$
\begin{aligned}
\mathbf{BA} &= \left(\sum_{j=1}^l \mu_j \mathbf{F}_j\right)\left(\sum_{i=1}^k \lambda_i \mathbf{E}_i\right) = \sum_{i=1}^k \sum_{j=1}^l \lambda_i \mu_j \mathbf{F}_j \mathbf{E}_i = \\
&= \sum_{j=1}^l \sum_{i=1}^k \mu_j \lambda_i \mathbf{E}_i \mathbf{F}_j = \left(\sum_{i=1}^k \lambda_i \mathbf{E}_i\right)\left(\sum_{j=1}^l \mu_j \mathbf{F}_j\right) = \mathbf{AB}.
\end{aligned}
\tag{1.250}
$$

Necessity follows from the fact that, if $\mathbf{F}_j$, $1 \le j \le l$ commutes with $\mathbf{A}$ then, by the same argument as mentioned earlier, it also commutes with every polynomial of $\mathbf{A}$; and hence also with $p_i(\mathbf{A}) = \mathbf{E}_i$ defined by

Eq. (1.219). Conversely, if $\mathbf{E}_i$, $1 \leq i \leq k$ commutes with $\mathbf{B}$ then it also commutes with every polynomial of $\mathbf{B}$; and hence also with the associated polynomial $q_j(\mathbf{A}) = \mathbf{E}_j$ defined by Eq. (1.219).

If $\mathbf{E_x} = |\mathbf{x}\rangle\langle\mathbf{x}|$ and $\mathbf{E_y} = |\mathbf{y}\rangle\langle\mathbf{y}|$ are two commuting projections (into one-dimensional subspaces of $\mathfrak{V}$) corresponding to the normalized vectors $\mathbf{x}$ and $\mathbf{y}$, respectively; that is, if $[\mathbf{E_x}, \mathbf{E_y}] = \mathbf{E_x}\mathbf{E_y} - \mathbf{E_y}\mathbf{E_x} = 0$, then they are either identical (the vectors are collinear) or orthogonal (the vectors $\mathbf{x}$ is orthogonal to $\mathbf{y}$).

For a proof, note that if $\mathbf{E_x}$ and $\mathbf{E_y}$ commute, then $\mathbf{E_x}\mathbf{E_y} = \mathbf{E_y}\mathbf{E_x}$; and hence $|\mathbf{x}\rangle\langle\mathbf{x}|\mathbf{y}\rangle\langle\mathbf{y}| = |\mathbf{y}\rangle\langle\mathbf{y}|\mathbf{x}\rangle\langle\mathbf{x}|$. Thus, $(\langle\mathbf{x}|\mathbf{y}\rangle)|\mathbf{x}\rangle\langle\mathbf{y}| = (\overline{\langle\mathbf{x}|\mathbf{y}\rangle})|\mathbf{y}\rangle\langle\mathbf{x}|$, which, applied to arbitrary vectors $|\nu\rangle \in \mathfrak{V}$, is only true if either $\mathbf{x} = \pm\mathbf{y}$, or if $\mathbf{x} \perp \mathbf{y}$ (and thus $\langle\mathbf{x}|\mathbf{y}\rangle = 0$).

Suppose some set $\mathbf{M} = \{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k\}$ of self-adjoint transformations on a finite-dimensional inner product space. These transformations $\mathbf{A}_i \in \mathbf{M}$, $1 \leq i \leq k$ are mutually commuting – that is, $[\mathbf{A}_i, \mathbf{A}_j] = 0$ for all $1 \leq i, j \leq k$ – if and only if there exists a *maximal* (with respect to the set $\mathbf{M}$) self-adjoint transformation $\mathbf{R}$ and a set of real-valued functions $F = \{f_1, f_2, \ldots, f_k\}$ of a real variable so that $\mathbf{A}_1 = f_1(\mathbf{R})$, $\mathbf{A}_2 = f_2(\mathbf{R})$, $\ldots$, $\mathbf{A}_k = f_k(\mathbf{R})$. If such a *maximal operator* $\mathbf{R}$ exists, then it can be written as a function of all transformations in the set $\mathbf{M}$; that is, $\mathbf{R} = G(\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k)$, where G is a suitable real-valued function of n variables (cf. Ref. [30], Satz 8).

For a proof involving two operators $\mathbf{A}_1$ and $\mathbf{A}_2$ we note that sufficiency can be derived from commutativity, which follows from $\mathbf{A}_1\mathbf{A}_2 = f_1(\mathbf{R})f_2(\mathbf{R}) = f_1(\mathbf{R})f_2(\mathbf{R}) = \mathbf{A}_2\mathbf{A}_1$.

Necessity follows by first noticing that, as derived earlier, the projection operators $\mathbf{E}_i$ and $\mathbf{F}_j$ in the spectral forms of $\mathbf{A}_1 = \sum_{i=1}^{k} \lambda_i \mathbf{E}_i$ and $\mathbf{A}_2 = \sum_{j=1}^{l} \mu_j \mathbf{F}_j$ mutually commute; that is, $\mathbf{E}_i\mathbf{F}_j = \mathbf{F}_j\mathbf{E}_i$.

For the sake of construction, design $g(x, y) \in \mathbb{R}$ to be any real-valued function (which can be a polynomial) of two real variables $x, y \in \mathbb{R}$ with the property that all the coefficients $c_{ij} = g(\lambda_i, \mu_j)$ are distinct. Next, define the maximal operator $\mathbf{R}$ by

$$\mathbf{R} = g(\mathbf{A}_1, \mathbf{A}_2) = \sum_{i=1}^{k} \sum_{j=1}^{l} c_{ij} \mathbf{E}_i \mathbf{F}_j, \qquad (1.251)$$

and the two functions f_1 and f_2 such that $f_1(c_{ij}) = \lambda_i$, as well as $f_2(c_{ij}) = \mu_j$, which result in

$$f_1(\mathbf{R}) = \sum_{i=1}^{k} \sum_{j=1}^{j} f_1(c_{ij}) \mathbf{E}_i \mathbf{F}_j = \sum_{i=1}^{k} \sum_{j=1}^{j} \lambda_i \mathbf{E}_i \mathbf{F}_j = \left(\sum_{i=1}^{k} \lambda_i \mathbf{E}_i\right)\underbrace{\left(\sum_{j=1}^{l} \mathbf{F}_j\right)}_{\mathbb{I}} = \mathbf{A}_1$$

$$f_2(\mathbf{R}) = \sum_{i=1}^{k} \sum_{j=1}^{j} f_2(c_{ij}) \mathbf{E}_i \mathbf{F}_j = \sum_{i=1}^{k} \sum_{j=1}^{j} \mu_j \mathbf{E}_i \mathbf{F}_j = \underbrace{\left(\sum_{i=1}^{k} \mathbf{E}_i\right)}_{\mathbb{I}}\left(\sum_{j=1}^{l} \mu_j \mathbf{F}_j\right) = \mathbf{A}_2.$$

$$(1.252)$$

The maximal operator $\mathbf{R}$ can be interpreted as encoding or containing all the information of a collection of commuting operators at once. Stated pointedly, rather than to enumerate all the k operators in $\mathbf{M}$ sepa-

rately, a single maximal operator **R** represents **M**; in this sense, the operators $\mathbf{A}_i \in \mathbf{M}$ are all just (most likely incomplete) *aspects* of – or individual, "lossy" (i.e., one-to-many) functional views on – the maximal operator **R**.

Let us demonstrate the machinery developed so far by an example. Consider the normal matrices

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{B} = \begin{pmatrix} 2 & 3 & 0 \\ 3 & 2 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \mathbf{C} = \begin{pmatrix} 5 & 7 & 0 \\ 7 & 5 & 0 \\ 0 & 0 & 11 \end{pmatrix},$$

which are mutually commutative; that is, $[\mathbf{A}, \mathbf{B}] = \mathbf{A}\mathbf{B} - \mathbf{B}\mathbf{A} = [\mathbf{A}, \mathbf{C}] = \mathbf{A}\mathbf{C} - \mathbf{B}\mathbf{C} = [\mathbf{B}, \mathbf{C}] = \mathbf{B}\mathbf{C} - \mathbf{C}\mathbf{B} = 0$.

The eigensystems – that is, the set of the set of eigenvalues and the set of the associated eigenvectors – of **A**, **B** and **C** are

$$\{\{1, -1, 0\}, \{(1, 1, 0)^T, (-1, 1, 0)^T, (0, 0, 1)^T\}\},$$
$$\{\{5, -1, 0\}, \{(1, 1, 0)^T, (-1, 1, 0)^T, (0, 0, 1)^T\}\}, \tag{1.253}$$
$$\{\{12, -2, 11\}, \{(1, 1, 0)^T, (-1, 1, 0)^T, (0, 0, 1)^T\}\}.$$

They share a common orthonormal set of eigenvectors

$$\left\{ \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

which form an orthonormal basis of $\mathbb{R}^3$ or $\mathbb{C}^3$. The associated projections are obtained by the outer (dyadic or tensor) products of these vectors; that is,

$$\mathbf{E}_1 = \frac{1}{2} \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$\mathbf{E}_2 = \frac{1}{2} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \tag{1.254}$$

$$\mathbf{E}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Thus the spectral decompositions of **A**, **B** and **C** are

$$\mathbf{A} = \mathbf{E}_1 - \mathbf{E}_2 + 0\mathbf{E}_3,$$
$$\mathbf{B} = 5\mathbf{E}_1 - \mathbf{E}_2 + 0\mathbf{E}_3, \tag{1.255}$$
$$\mathbf{C} = 12\mathbf{E}_1 - 2\mathbf{E}_2 + 11\mathbf{E}_3,$$

respectively.

One way to define the maximal operator **R** for this problem would be

$$\mathbf{R} = \alpha\mathbf{E}_1 + \beta\mathbf{E}_2 + \gamma\mathbf{E}_3,$$

with $\alpha, \beta, \gamma \in \mathbb{R} - 0$ and $\alpha \neq \beta \neq \gamma \neq \alpha$. The functional coordinates $f_i(\alpha)$, $f_i(\beta)$, and $f_i(\gamma)$, $i \in \{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$, of the three functions $f_\mathbf{A}(\mathbf{R})$, $f_\mathbf{B}(\mathbf{R})$, and $f_\mathbf{C}(\mathbf{R})$

chosen to match the projection coefficients obtained in Eq. (1.255); that is,

$$\mathbf{A} = f_{\mathbf{A}}(\mathbf{R}) = \mathbf{E}_1 - \mathbf{E}_2 + 0\mathbf{E}_3,$$
$$\mathbf{B} = f_{\mathbf{B}}(\mathbf{R}) = 5\mathbf{E}_1 - \mathbf{E}_2 + 0\mathbf{E}_3, \qquad (1.256)$$
$$\mathbf{C} = f_{\mathbf{C}}(\mathbf{R}) = 12\mathbf{E}_1 - 2\mathbf{E}_2 + 11\mathbf{E}_3.$$

As a consequence, the functions $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ need to satisfy the relations

$$f_{\mathbf{A}}(\alpha) = 1, \ f_{\mathbf{A}}(\beta) = -1, \ f_{\mathbf{A}}(\gamma) = 0,$$
$$f_{\mathbf{B}}(\alpha) = 5, \ f_{\mathbf{B}}(\beta) = -1, \ f_{\mathbf{B}}(\gamma) = 0, \qquad (1.257)$$
$$f_{\mathbf{C}}(\alpha) = 12, \ f_{\mathbf{C}}(\beta) = -2, \ f_{\mathbf{C}}(\gamma) = 11.$$

It is no coincidence that the projections in the spectral forms of $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are identical. Indeed it can be shown that mutually commuting *normal operators* always share the same eigenvectors; and thus also the same projections.

Let the set $\mathbf{M} = \{\mathbf{A}_1, \mathbf{A}_2, \ldots, \mathbf{A}_k\}$ be mutually commuting normal (or Hermitian, or self-adjoint) transformations on an n-dimensional inner product space. Then there exists an orthonormal basis $\mathfrak{B} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ such that every $\mathbf{f}_j \in \mathfrak{B}$ is an eigenvector of each of the $\mathbf{A}_i \in \mathbf{M}$. Equivalently, there exist n orthogonal projections (let the vectors $\mathbf{f}_j$ be represented by the coordinates which are column vectors) $\mathbf{E}_j = \mathbf{f}_j \otimes \mathbf{f}_j^\dagger$ such that every $\mathbf{E}_j$, $1 \le j \le n$ occurs in the spectral form of each of the $\mathbf{A}_i \in \mathbf{M}$.

Informally speaking, a "generic" maximal operator $\mathbf{R}$ on an n-dimensional Hilbert space $\mathfrak{V}$ can be interpreted in terms of a particular orthonormal basis $\{\mathbf{f}_1, \mathbf{f}_2, \ldots, \mathbf{f}_n\}$ of $\mathfrak{V}$ – indeed, the n elements of that basis would have to correspond to the projections occurring in the spectral decomposition of the self-adjoint operators generated by $\mathbf{R}$.

Likewise, the "maximal knowledge" about a quantized physical system – in terms of empirical operational quantities – would correspond to such a single maximal operator; or to the orthonormal basis corresponding to the spectral decomposition of it. Thus it might not be unreasonable to speculate that a particular (pure) physical state is best characterized by a particular orthonomal basis.

1.33   *Measures on closed subspaces*

In what follows we shall assume that all *(probability) measures* or *states* behave quasi-classically on sets of mutually commuting self-adjoint operators, and, in particular, on orthogonal projections. One could call this property *subclassicality*.

This can be formalized as follows. Consider some set $\{|\mathbf{x}_1\rangle, |\mathbf{x}_2\rangle, \ldots, |\mathbf{x}_k\rangle\}$ of mutually orthogonal, normalized vectors, so that $\langle \mathbf{x}_i | \mathbf{x}_j \rangle = \delta_{ij}$; and associated with it, the set $\{\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_k\}$ of mutually orthogonal (and thus commuting) one-dimensional projections $\mathbf{E}_i = |\mathbf{x}_i\rangle\langle \mathbf{x}_i|$ on a finite-dimensional inner product space $\mathfrak{V}$.

We require that probability measures μ on such mutually commuting sets of observables behave quasi-classically. Therefore, they should be

additive; that is,

$$\mu\left(\sum_{i=1}^{k}\mathbf{E}_i\right) = \sum_{i=1}^{k}\mu(\mathbf{E}_i). \tag{1.258}$$

Such a measure is determined by its values on the one-dimensional projections.

Stated differently, we shall assume that, for any two orthogonal projections $\mathbf{E}$ and $\mathbf{F}$ if $\mathbf{EF} = \mathbf{FE} = 0$, their sum $\mathbf{G} = \mathbf{E} + \mathbf{F}$ has expectation value

$$\mu(\mathbf{G}) \equiv \langle \mathbf{G} \rangle = \langle \mathbf{E} \rangle + \langle \mathbf{F} \rangle \equiv \mu(\mathbf{E}) + \mu(\mathbf{F}). \tag{1.259}$$

Any such measure μ satisfying (1.258) can be expressed in terms of a (positive) real valued function f on the unit vectors in $\mathfrak{V}$ by

$$\mu(\mathbf{E}_x) = f(|\mathbf{x}\rangle) \equiv f(\mathbf{x}), \tag{1.260}$$

(where $\mathbf{E}_x = |\mathbf{x}\rangle\langle\mathbf{x}|$ for all unit vectors $|\mathbf{x}\rangle \in \mathfrak{V}$) by requiring that, for every orthonormal basis $\mathfrak{B} = \{|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle, \ldots, |\mathbf{e}_n\rangle\}$, the sum of all basis vectors yields 1; that is,

$$\sum_{i=1}^{n} f(|\mathbf{e}_i\rangle) \equiv \sum_{i=1}^{n} f(\mathbf{e}_i) = 1. \tag{1.261}$$

f is called a (positive) *frame function* of weight 1.

1.33.1 Gleason's theorem

From now on we shall mostly consider vector spaces of dimension three or greater, since only in these cases two orthonormal bases intertwine in a common vector, making possible some arguments involving multiple intertwining bases – in two dimensions, distinct orthonormal bases contain distinct basis vectors.

Gleason's theorem [31] states that, for a Hilbert space of dimension three or greater, every frame function defined in (1.261) is of the form of the inner product

$$f(\mathbf{x}) \equiv f(|\mathbf{x}\rangle) = \langle\mathbf{x}|\rho\mathbf{x}\rangle = \sum_{i=1}^{k\le n}\rho_i\langle\mathbf{x}|\psi_i\rangle\langle\psi_i|\mathbf{x}\rangle = \sum_{i=1}^{k\le n}\rho_i|\langle\mathbf{x}|\psi_i\rangle|^2, \tag{1.262}$$

where (i) ρ is a positive operator (and therefore self-adjoint; see Section 1.21 on page 45), and (ii) ρ is of the trace class, meaning its trace (cf. Section 1.18 on page 40) is one. That is, $\rho = \sum_{i=1}^{k\le n}\rho_i|\psi_i\rangle\langle\psi_i|$ with $\rho_i \in \mathbb{R}$, $\rho_i \ge 0$, and $\sum_{i=1}^{k\le n}\rho_i = 1$. No proof is given here.

In terms of projections [cf. Eqs.(1.63) on page 22], (1.262) can be written as

$$\mu(\mathbf{E}_x) = \text{Tr}(\rho\mathbf{E}_x) \tag{1.263}$$

Therefore, for a Hilbert space of dimension three or greater, the spectral theorem suggests that the only possible form of the expectation value of a self-adjoint operator $\mathbf{A}$ has the form

$$\langle\mathbf{A}\rangle = \text{Tr}(\rho\mathbf{A}). \tag{1.264}$$

In quantum physical terms, in the formula (1.264) above the trace is taken over the operator product of the density matrix [which represents

[31] Andrew M. Gleason. Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics (now Indiana University Mathematics Journal)*, 6(4):885–893, 1957. ISSN 0022-2518. DOI: 10.1512/iumj.1957.6.56050". URL http://doi.org/10.1512/iumj.1957.6.56050; Anatolij Dvurečenskij. *Gleason's Theorem and Its Applications*, volume 60 of *Mathematics and its Applications*. Kluwer Academic Publishers, Springer, Dordrecht, 1993. ISBN 9048142091,978-90-481-4209-5,978-94-015-8222-3. DOI: 10.1007/978-94-015-8222-3. URL http://doi.org/10.1007/978-94-015-8222-3; Itamar Pitowsky. Infinite and finite Gleason's theorems and the logic of indeterminacy. *Journal of Mathematical Physics*, 39(1):218–228, 1998. DOI: 10.1063/1.532334. URL http://doi.org/10.1063/1.532334; Fred Richman and Douglas Bridges. A constructive proof of Gleason's theorem. *Journal of Functional Analysis*, 162:287–312, 1999. DOI: 10.1006/jfan.1998.3372. URL http://doi.org/10.1006/jfan.1998.3372; Asher Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic Publishers, Dordrecht, 1993; and Jan Hamhalter. *Quantum Measure Theory. Fundamental Theories of Physics, Vol. 134*. Kluwer Academic Publishers, Dordrecht, Boston, London, 2003. ISBN 1-4020-1714-6

a positive (and thus self-adjoint) operator of the trace class] ρ with the observable $\mathbf{A} = \sum_{i=1}^{k} \lambda_i \mathbf{E}_i$.

In particular, if $\mathbf{A}$ is a projection $\mathbf{E} = |\mathbf{e}\rangle\langle\mathbf{e}|$ corresponding to an elementary yes-no proposition *"the system has property Q,"* then $\langle\mathbf{E}\rangle = \text{Tr}(\rho\mathbf{E}) = |\langle\mathbf{e}|\rho\rangle|^2$ corresponds to the probability of that property Q if the system is in state $\rho = |\rho\rangle\langle\rho|$ [for a motivation, see again Eqs. (1.63) on page 22].

Indeed, as already observed by Gleason, even for two-dimensional Hilbert spaces, a straightforward *Ansatz* yields a probability measure satisfying (1.258) as follows. Suppose some unit vector $|\rho\rangle$ corresponding to a pure quantum state (preparation) is selected. For each one-dimensional closed subspace corresponding to a one-dimensional orthogonal projection observable (interpretable as an elementary yes-no proposition) $E = |\mathbf{e}\rangle\langle\mathbf{e}|$ along the unit vector $|\mathbf{e}\rangle$, define $w_\rho(|\mathbf{e}\rangle) = |\langle\mathbf{e}|\rho\rangle|^2$ to be the square of the length $|\langle\rho|\mathbf{e}\rangle|$ of the projection of $|\rho\rangle$ onto the subspace spanned by $|\mathbf{e}\rangle$.

The reason for this is that an orthonormal basis $\{|\mathbf{e}_i\rangle\}$ "induces" an *ad hoc* probability measure w_ρ on any such context (and thus basis). To see this, consider the length of the orthogonal (with respect to the basis vectors) projections of $|\rho\rangle$ onto all the basis vectors $|\mathbf{e}_i\rangle$, that is, the norm of the resulting vector projections of $|\rho\rangle$ onto the basis vectors, respectively. This amounts to computing the absolute value of the Euclidean scalar products $\langle\mathbf{e}_i|\rho\rangle$ of the state vector with all the basis vectors.

In order that all such absolute values of the scalar products (or the associated norms) sum up to one and yield a probability measure as required in Eq. (1.258), recall that $|\rho\rangle$ is a unit vector and note that, by the Pythagorean theorem, these absolute values of the individual scalar products – or the associated norms of the vector projections of $|\rho\rangle$ onto the basis vectors – must be squared. Thus the value $w_\rho(|\mathbf{e}_i\rangle)$ must be the square of the scalar product of $|\rho\rangle$ with $|\mathbf{e}_i\rangle$, corresponding to the square of the length (or norm) of the respective projection vector of $|\rho\rangle$ onto $|\mathbf{e}_i\rangle$. For complex vector spaces one has to take the absolute square of the scalar product; that is, $f_\rho(|\mathbf{e}_i\rangle) = |\langle\mathbf{e}_i|\rho\rangle|^2$.

Pointedly stated, from this point of view the probabilities $w_\rho(|\mathbf{e}_i\rangle)$ are just the (absolute) squares of the coordinates of a unit vector $|\rho\rangle$ with respect to some orthonormal basis $\{|\mathbf{e}_i\rangle\}$, representable by the square $|\langle\mathbf{e}_i|\rho\rangle|^2$ of the length of the vector projections of $|\rho\rangle$ onto the basis vectors $|\mathbf{e}_i\rangle$ – one might also say that each orthonormal basis allows "a view" on the pure state $|\rho\rangle$. In two dimensions this is illustrated for two bases in Fig. 1.5. The squares come in because the absolute values of the individual components do not add up to one; but their squares do. These considerations apply to Hilbert spaces of any, including two, finite dimensions. In this non-general, *ad hoc* sense the Born rule for a system in a pure state and an elementary proposition observable (quantum encodable by a one-dimensional projection operator) can be motivated by the requirement of additivity for arbitrary finite dimensional Hilbert space.
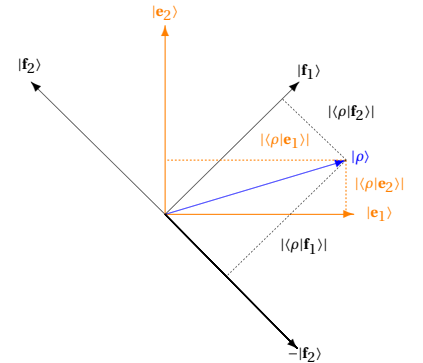
Figure 1.5: Different orthonormal bases $\{|\mathbf{e}_1\rangle, |\mathbf{e}_2\rangle\}$ and $\{|\mathbf{f}_1\rangle, |\mathbf{f}_2\rangle\}$ offer different "views" on the pure state $|\rho\rangle$. As $|\rho\rangle$ is a unit vector it follows from the Pythagorean theorem that $|\langle\rho|\mathbf{e}_1\rangle|^2 + |\langle\rho|\mathbf{e}_2\rangle|^2 = |\langle\rho|\mathbf{f}_1\rangle|^2 + |\langle\rho|\mathbf{f}_2\rangle|^2 = 1$, thereby motivating the use of the absolute value (modulus) squared of the amplitude for quantum probabilities on pure states.

1.33.2 Kochen-Specker theorem

For a Hilbert space of dimension three or greater, there does not exist any two-valued probability measures interpretable as consistent, overall truth assignment [32]. As a result of the nonexistence of two-valued states, the classical strategy to construct probabilities by a convex combination of all two-valued states fails entirely.

In *Greechie diagram* [33], points represent basis vectors. If they belong to the same basis, they are connected by smooth curves.
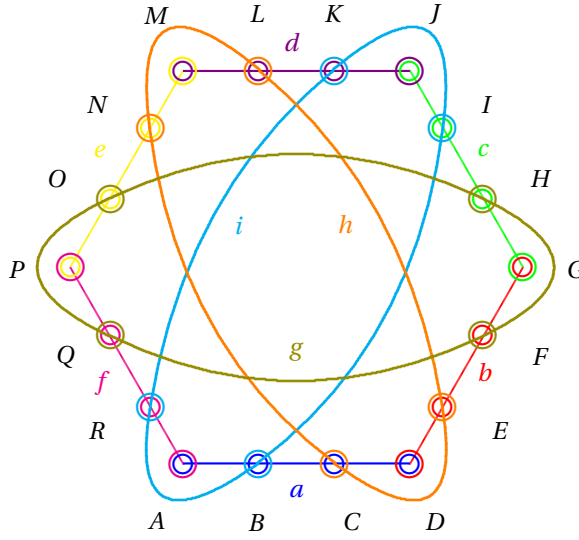
The most compact way of deriving the Kochen-Specker theorem in four dimensions has been given by Cabello [34]. For the sake of demonstration, consider a Greechie (orthogonality) diagram of a finite subset of the continuum of blocks or contexts embeddable in four-dimensional real Hilbert space without a two-valued probability measure. The proof of the Kochen-Specker theorem uses nine tightly interconnected contexts $a = \{A, B, C, D\}$, $b = \{D, E, F, G\}$, $c = \{G, H, I, J\}$, $d = \{J, K, L, M\}$, $e = \{M, N, O, P\}$, $f = \{P, Q, R, A\}$, $g = \{F, H, O, Q\}$ $h = \{C, E, L, N\}$, $i = \{B, I, K, R\}$, consisting of the 18 projections associated with the one dimensional subspaces spanned by the vectors from the origin $(0, 0, 0, 0)$ to $A = (0, 0, 1, -1)$, $B = (1, -1, 0, 0)$, $C = (1, 1, -1, -1)$, $D = (1, 1, 1, 1)$, $E = (1, -1, 1, -1)$, $F = (1, 0, -1, 0)$, $G = (0, 1, 0, -1)$, $H = (1, 0, 1, 0)$, $I = (1, 1, -1, 1)$, $J = (-1, 1, 1, 1)$, $K = (1, 1, 1, -1)$, $L = (1, 0, 0, 1)$, $M = (0, 1, -1, 0)$, $N = (0, 1, 1, 0)$, $O = (0, 0, 0, 1)$, $P = (1, 0, 0, 0)$, $Q = (0, 1, 0, 0)$, $R = (0, 0, 1, 1)$, respectively. Greechie diagrams represent atoms by points, and contexts by maximal smooth, unbroken curves.

In a proof by contradiction, note that, on the one hand, every observable proposition occurs in exactly *two* contexts. Thus, in an enumeration of the four observable propositions of each of the nine contexts, there appears to be an *even* number of true propositions, provided that the value of an observable does not depend on the context (i.e. the assignment is *noncontextual*). Yet, on the other hand, as there is an *odd* number (actually nine) of contexts, there should be an *odd* number (actually nine) of true propositions.

[32] Ernst Specker. Die Logik nicht gleichzeitig entscheidbarer Aussagen. *Dialectica*, 14(2-3):239–246, 1960. DOI: 10.1111/j.1746-8361.1960.tb00422.x. URL http://doi.org/10.1111/j.1746-8361.1960.tb00422.x; and Simon Kochen and Ernst P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics (now Indiana University Mathematics Journal)*, 17(1):59–87, 1967. ISSN 0022-2518. DOI: 10.1512/iumj.1968.17.17004. URL http://doi.org/10.1512/iumj.1968.17.17004

[33] Richard J. Greechie. Orthomodular lattices admitting no states. *Journal of Combinatorial Theory. Series A*, 10: 119–132, 1971. DOI: 10.1016/0097-3165(71)90015-X. URL http://doi.org/10.1016/0097-3165(71)90015-X

[34] Adán Cabello, José M. Estebaranz, and G. García-Alcaine. Bell-Kochen-Specker theorem: A proof with 18 vectors. *Physics Letters A*, 212(4):183–187, 1996. DOI: 10.1016/0375-9601(96)00134-X. URL http://doi.org/10.1016/0375-9601(96)00134-X; and Adán Cabello. Kochen-Specker theorem and experimental test on hidden variables. *International Journal of Modern Physics A*, 15(18):2813–2820, 2000. DOI: 10.1142/S0217751X00002020. URL http://doi.org/10.1142/S0217751X00002020

2

Multilinear Algebra and Tensors

What follows is a multilinear extension of the previous chapter. Tensors will be introduced as multilinear forms, and their transformation properties will be derived.

For many physicists the following derivations might appear confusing and overly formalistic as they might have difficulties to "see the forest for the trees." For those, a brief overview sketching the most important aspects of tensors might serve as a first orientation.

Let us start by defining, or rather *declaring*, the scales of vectors of a basis of the (base) vector space – the reference axes of this base space – to "(co-)vary." This is just a "fixation," a designation of notation.

Based on this declaration or rather convention, and relative to the behaviour with respect to changes of scale of the reference axes (the basis vectors) in the base vector space, there exist two important categories: entities which co-vary, and entities which vary inversely, that is, contra-vary, with such changes.

- Contravariant entities such as vectors in the base vector space: These vectors of the base vector space are called contravariant because their *components* contra-vary (that is, vary inversely) with respect to variations of the basis vectors. By identification the components of contravariant vectors (or tensors) are also contravariant. In general, a multilinear form on a vector space is called contravariant if its components (coordinates) are contravariant; that is, they contra-vary with respect to variations of the basis vectors.

- Covariant entities such as vectors in the dual space: The vectors of the dual space are called covariant because their *components* contra-vary with respect to variations of the basis vectors of the dual space, which in turn contra-vary with respect to variations of the basis vectors of the base space. Thereby the double contra-variations (inversions) cancel out, so that effectively the vectors of the dual space co-vary with the vectors of the basis of the base vector space. By identification the components of covariant vectors (or tensors) are also covariant. In general, a multilinear form on a vector space is called covariant if its components (coordinates) are covariant; that is, they co-vary with respect to variations of the basis vectors of the base vector space.

The dual space is spanned by all linear functionals on that vector space (cf. Section 1.8 on page 15).

- Covariant and contravariant indices will be denoted by subscripts (lower indices) and superscripts (upper indices), respectively.

- Covariant and contravariant entities transform inversely. Informally, this is due to the fact that their changes must compensate each other, as covariant and contravariant entities are "tied together" by some invariant (id)entities such as vector encoding and dual basis formation.

- Covariant entities can be transformed into contravariant ones by the application of metric tensors, and, *vice versa,* by the inverse of metric tensors.

2.1   Notation

In what follows, vectors and tensors will be encoded in terms of indexed coordinates or components (with respect to a specific basis). The biggest advantage is that such coordinates or components are scalars which can be exchanged and rearranged according to commutativity, associativity and distributivity, as well as differentiated.

Let us consider the vector space $\mathfrak{V} = \mathbb{R}^n$ of dimension n. A covariant basis $\mathfrak{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ of $\mathfrak{V}$ consists of n covariant basis vectors $\mathbf{e}_i$. A contravariant basis $\mathfrak{B}^* = \{\mathbf{e}_1^*, \mathbf{e}_2^*, \ldots, \mathbf{e}_n^*\} = \{\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^n\}$ of the dual space $\mathfrak{V}^*$ (cf. Section 1.8.1 on page 16) consists of n basis vectors $\mathbf{e}_i^*$, where $\mathbf{e}_i^* = \mathbf{e}^i$ is just a different notation.

Every contravariant vector $\mathbf{x} \in \mathfrak{V}$ can be coded by, or expressed in terms of, its contravariant vector components $X^1, X^2, \ldots, X^n \in \mathbb{R}$ by $\mathbf{x} = \sum_{i=1}^n X^i \mathbf{e}_i$. Likewise, every covariant vector $\mathbf{x} \in \mathfrak{V}^*$ can be coded by, or expressed in terms of, its covariant vector components $X_1, X_2, \ldots, X_n \in \mathbb{R}$ by $\mathbf{x} = \sum_{i=1}^n X_i \mathbf{e}_i^* = \sum_{i=1}^n X_i \mathbf{e}^i$.

Suppose that there are k arbitrary contravariant vectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ in $\mathfrak{V}$ which are indexed by a subscript (lower index). This lower index should not be confused with a covariant lower index. Every such vector $\mathbf{x}_j$, $1 \le j \le k$ has contravariant vector components $X_j^{1_j}, X_j^{2_j}, \ldots, X_j^{n_j} \in \mathbb{R}$ with respect to a particular basis $\mathfrak{B}$ such that

$$\mathbf{x}_j = \sum_{i_j=1}^n X_j^{i_j} \mathbf{e}_{i_j}. \tag{2.1}$$

Likewise, suppose that there are k arbitrary covariant vectors $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k$ in the dual space $\mathfrak{V}^*$ which are indexed by a superscript (upper index). This upper index should not be confused with a contravariant upper index. Every such vector $\mathbf{x}^j$, $1 \le j \le k$ has covariant vector components $X_{1_j}^j, X_{2_j}^j, \ldots, X_{n_j}^j \in \mathbb{R}$ with respect to a particular basis $\mathfrak{B}^*$ such that

$$\mathbf{x}^j = \sum_{i_j=1}^n X_{i_j}^j \mathbf{e}^{i_j}. \tag{2.2}$$

Tensors are constant with respect to variations of points of $\mathbb{R}^n$. In contradistinction, *tensor fields* depend on points of $\mathbb{R}^n$ in a nontrivial (nonconstant) way. Thus, the components of a tensor field depend on the coordinates. For example, the contravariant vector defined by the coordinates $\left(5.5, 3.7, \ldots, 10.9\right)^T$ with respect to a particular basis $\mathfrak{B}$ is a tensor;

For a more systematic treatment, see for instance Klingbeil's   or Dirschmid's introductions.

Ebergard Klingbeil. *Tensorrechnung für Ingenieure.* Bibliographisches Institut, Mannheim, 1966; and Hans Jörg Dirschmid. *Tensoren und Felder.* Springer, Vienna, 1996

For a detailed explanation of covariance and contravariance, see Section 2.2 on page 77.

Note that in both covariant and contravariant cases the upper-lower pairings " $\cdot_i \cdot^i$ " and " $\cdot^i \cdot_i$ "of the indices match.

This notation " $X_j^{i_j}$ " for the ith component of the jth vector is redundant as it requires two indices j; we could have just denoted it by " X^{i_j} ." The lower index j does *not* correspond to any covariant entity but just indexes the jth vextor $\mathbf{x}_j$.

Again, this notation " $X_{i_j}^j$ " for the ith component of the jth vector is redundant as it requires two indices j; we could have just denoted it by " X_{i_j} ." The upper index j does *not* correspond to any contravariant entity but just indexes the jth vextor $\mathbf{x}^j$.

while, again with respect to a particular basis $\mathfrak{B}$, $\left(\sin X_1, \cos X_2, \ldots, e^{X_n}\right)^T$ or $\left(X_1, X_2, \ldots, X_n\right)^T$, which depend on the coordinates $X_1, X_2, \ldots, X_n \in \mathbb{R}$, are tensor fields.

We adopt Einstein's summation convention to sum over equal indices. If not explained otherwise (that is, for orthonormal bases) those pairs have exactly one lower and one upper index.

In what follows, the notations "$x \cdot y$", "(x, y)" and "$\langle x \mid y \rangle$" will be used synonymously for the *scalar product* or *inner product*. Note, however, that the "dot notation $x \cdot y$" may be a little bit misleading; for example, in the case of the "pseudo-Euclidean" metric represented by the matrix $\mathrm{diag}(+, +, +, \cdots, +, -)$, it is no more the standard Euclidean dot product $\mathrm{diag}(+, +, +, \cdots, +, +)$.

2.2   Change of Basis

2.2.1   Transformation of the covariant basis

Let $\mathfrak{B}$ and $\mathfrak{B}'$ be two arbitrary bases of $\mathbb{R}^n$. Then every vector $\mathbf{f}_i$ of $\mathfrak{B}'$ can be represented as linear combination of basis vectors of $\mathfrak{B}$ [see also Eqs. (1.97) and (1.98)]:

$$\mathbf{f}_i = \sum_{j=1}^n a^j{}_i \mathbf{e}_j, \qquad i = 1, \ldots, n. \tag{2.3}$$

The matrix

$$\mathbf{A} \equiv a^j{}_i \equiv \begin{pmatrix} a^1{}_1 & a^1{}_2 & \cdots & a^1{}_n \\ a^2{}_1 & a^2{}_2 & \cdots & a^2{}_n \\ \vdots & \vdots & \ddots & \vdots \\ a^n{}_1 & a^n{}_2 & \cdots & a^n{}_n \end{pmatrix}. \tag{2.4}$$

is called the *transformation matrix*. As defined in (1.3) on page 4, the second (from the left to the right), rightmost (in this case lower) index i varying in row vectors is the column index; and, the first, leftmost (in this case upper) index j varying in columns is the row index, respectively.

Note that, as discussed earlier, it is necessary to fix a convention for the transformation of the covariant basis vectors discussed on page 31. This then specifies the exact form of the (inverse, contravariant) transformation of the components or coordinates of vectors.

Perhaps not very surprisingly, compared to the transformation (2.3) yielding the "new" basis $\mathfrak{B}'$ in terms of elements of the "old" basis $\mathfrak{B}$, a transformation yielding the "old" basis $\mathfrak{B}$ in terms of elements of the "new" basis $\mathfrak{B}'$ turns out to be just the inverse "back" transformation of the former: substitution of (2.3) yields

$$\mathbf{e}_i = \sum_{j=1}^n a'^j{}_i \mathbf{f}_j = \sum_{j=1}^n a'^j{}_i \sum_{k=1}^n a^k{}_j \mathbf{e}_k = \sum_{k=1}^n \left(\sum_{j=1}^n a'^j{}_i a^k{}_j \right) \mathbf{e}_k, \tag{2.5}$$

which, due to the linear independence of the basis vectors $\mathbf{e}_i$ of $\mathfrak{B}$, can only be satisfied if

$$a^k{}_j a'^j{}_i = \delta^k_i \qquad \text{or} \qquad \mathbf{A}\mathbf{A}' = \mathbb{I}. \tag{2.6}$$

Thus $\mathbf{A}'$ is the *inverse matrix* $\mathbf{A}^{-1}$ of $\mathbf{A}$. In index notation,

$$a'^j{}_i = (a^{-1})^j{}_i, \tag{2.7}$$

and

$$\mathbf{e}_i = \sum_{j=1}^n (a^{-1})^j{}_i \mathbf{f}_j. \tag{2.8}$$

2.2.2   Transformation of the contravariant coordinates

Consider an arbitrary contravariant vector $\mathbf{x} \in \mathbb{R}^n$ in two basis representations: (i) with contravariant components X^i with respect to the basis $\mathfrak{B}$, and (ii) with Y^i with respect to the basis $\mathfrak{B}'$. Then, because both coordinates with respect to the two different bases have to encode the same vector, there has to be a "compensation-of-scaling" such that

$$\mathbf{x} = \sum_{i=1}^n X^i \mathbf{e}_i = \sum_{i=1}^n Y^i \mathbf{f}_i. \tag{2.9}$$

Insertion of the basis transformation (2.3) and relabelling of the indices $i \leftrightarrow j$ yields

$$\mathbf{x} = \sum_{i=1}^n X^i \mathbf{e}_i = \sum_{i=1}^n Y^i \mathbf{f}_i = \sum_{i=1}^n Y^i \sum_{j=1}^n a^j{}_i \mathbf{e}_j$$

$$= \sum_{i=1}^n \sum_{j=1}^n a^j{}_i Y^i \mathbf{e}_j = \sum_{j=1}^n \left[\sum_{i=1}^n a^j{}_i Y^i \right] \mathbf{e}_j = \sum_{i=1}^n \left[\sum_{j=1}^n a^i{}_j Y^j \right] \mathbf{e}_i. \tag{2.10}$$

A comparison of coefficients yields the transformation laws of vector components [see also Eq. (1.106)]

$$X^i = \sum_{j=1}^n a^i{}_j Y^j. \tag{2.11}$$

In the matrix notation introduced in Eq. (1.19) on page 12, (2.11) can be written as

$$X = \mathbf{A} Y. \tag{2.12}$$

A similar "compensation-of-scaling" argument using (2.8) yields the transformation laws for

$$Y^j = \sum_{i=1}^n (a^{-1})^j{}_i X^i \tag{2.13}$$

with respect to the covariant basis vectors. In the matrix notation introduced in Eq. (1.19) on page 12, (2.13) can simply be written as

$$Y = \left(\mathbf{A}^{-1} \right) X. \tag{2.14}$$

If the basis transformations involve nonlinear coordinate changes – such as from the Cartesian to the polar or spherical coordinates discussed later – we have to employ differentials

$$dX^j = \sum_{i=1}^n a^j{}_i \, dY^i, \tag{2.15}$$

so that, by partial diffentiation,

$$a^j{}_i = \frac{\partial X^j}{\partial Y^i}. \tag{2.16}$$

By assuming that the coordinate transformations are linear, $a_i{}^j$ can be expressed in terms of the coordinates X^j

$$a^j{}_i = \frac{X^j}{Y^i}. \tag{2.17}$$

Likewise,

$$dY^j = \sum_{i=1}^{n} (a^{-1})^j{}_i \, dX^i, \tag{2.18}$$

so that, by partial diffentiation,

$$(a^{-1})^j{}_i = \frac{\partial Y^j}{\partial X^i} = J_{ij}, \tag{2.19}$$

where J_{ij} stands for the *Jacobian matrix*

$$J \equiv J_{ij} = \frac{\partial Y^i}{\partial X^j} \equiv \begin{pmatrix} \frac{\partial Y^1}{\partial X^1} & \cdots & \frac{\partial Y^1}{\partial X^n} \\ \vdots & \ddots & \vdots \\ \frac{\partial Y^n}{\partial X^1} & \cdots & \frac{\partial Y^n}{\partial X^n} \end{pmatrix}. \tag{2.20}$$

2.2.3 *Transformation of the contravariant basis*

Consider again, as a starting point, a covariant basis $\mathfrak{B} = \{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ consisting of n basis vectors $\mathbf{e}_i$. A *contravariant basis* can be defined by identifying it with the dual basis introduced earlier in Section 1.8.1 on page 16, in particular, Eq. (1.39). Thus a *contravariant* basis $\mathfrak{B}^* = \{\mathbf{e}^1, \mathbf{e}^2, \ldots, \mathbf{e}^n\}$ is a set of n covariant basis vectors $\mathbf{e}^i$ which satisfy Eqs. (1.39)-(1.41)

$$\mathbf{e}^j(\mathbf{e}_i) = \left[\mathbf{e}_i, \mathbf{e}^j\right] = \left[\mathbf{e}_i, \mathbf{e}_j^*\right] = \delta_i^j = \delta_{ij}. \tag{2.21}$$

In terms of the bra-ket notation, (2.21) somewhat superficially transforms into (a formal justification for this identification is the Riesz representation theorem)

$$\left[|\mathbf{e}_i\rangle, \langle\mathbf{e}^j|\right] = \langle\mathbf{e}^j|\mathbf{e}_i\rangle = \delta_{ij}. \tag{2.22}$$

Furthermore, the resolution of identity (1.124) can be rewritten as

$$\mathbb{I}_n = \sum_{i=1}^{n} |\mathbf{e}^i\rangle\langle\mathbf{e}_i|. \tag{2.23}$$

As demonstrated earlier in Eq. (1.42) the vectors $\mathbf{e}_i^* = \mathbf{e}^i$ of the dual basis can be used to "retreive" the components of arbitrary vectors $\mathbf{x} = \sum_j X^j \mathbf{e}_j$ through

$$\mathbf{e}^i(\mathbf{x}) = \mathbf{e}^i\left(\sum_i X^j \mathbf{e}_j\right) = \sum_i X^j \mathbf{e}^i(\mathbf{e}_j) = \sum_i X^j \delta_j^i = X^i. \tag{2.24}$$

Likewise, the basis vectors $\mathbf{e}_i$ of the "base space" can be used to obtain the coordinates of any dual vector $\mathbf{x} = \sum_j X_j \mathbf{e}^j$ through

$$\mathbf{e}_i(\mathbf{x}) = \mathbf{e}_i\left(\sum_i X_j \mathbf{e}^j\right) = \sum_i X_j \mathbf{e}_i\left(\mathbf{e}^j\right) = \sum_i X_j \delta_i^j = X_i. \tag{2.25}$$

As also noted earlier, for orthonormal bases and Euclidean scalar (dot) products (the coordinates of) the dual basis vectors of an orthonormal

basis can be coded identically as (the coordinates of) the original basis vectors; that is, in this case, (the coordinates of) the dual basis vectors are just rearranged as the transposed form of the original basis vectors.

In the same way as argued for changes of covariant bases (2.3), that is, because every vector in the new basis of the dual space can be represented as a linear combination of the vectors of the original dual basis – we can make the formal *Ansatz*:

$$\mathbf{f}^j = \sum_i b^j{}_i \mathbf{e}^i, \tag{2.26}$$

where $\mathbf{B} \equiv b^j{}_i$ is the transformation matrix associated with the contravariant basis. How is b, the transformation of the contravariant basis, related to a, the transformation of the covariant basis?

Before answering this question, note that, again – and just as the necessity to fix a convention for the transformation of the covariant basis vectors discussed on page 31 – we have to choose by *convention* the way transformations are represented. In particular, if in (2.26) we would have reversed the indices $b^j{}_i \leftrightarrow b_i{}^j$, thereby effectively transposing the transformation matrix $\mathbf{B}$, this would have resulted in a changed (transposed) form of the transformation laws, as compared to both the transformation a of the covariant basis, and of the transformation of covariant vector components.

By exploiting (2.21) twice we can find the connection between the transformation of covariant and contravariant basis elements and thus tensor components; that is (by assuming Einstein's summation convention we are omitting to write sums explicitly),

$$\delta_i^j = \delta_{ij} = \mathbf{f}^j(\mathbf{f}_i) = \left[\mathbf{f}_i, \mathbf{f}^j\right] = \left[a^k{}_i \mathbf{e}_k, b^j{}_l \mathbf{e}^l\right] =$$
$$= a^k{}_i b^j{}_l \left[\mathbf{e}_k, \mathbf{e}^l\right] = a^k{}_i b^j{}_l \delta_k^l = a^k{}_i b^j{}_l \delta_{kl} = b^j{}_k a^k{}_i. \tag{2.27}$$

Therefore,

$$\mathbf{B} = (\mathbf{A})^{-1}, \text{ or } b^j{}_i = \left(a^{-1}\right)^j{}_i, \tag{2.28}$$

and

$$\mathbf{f}^j = \sum_i \left(a^{-1}\right)^j{}_i \mathbf{e}^i. \tag{2.29}$$

In short, by comparing (2.29) with (2.13), we find that the vectors of the contravariant dual basis transform just like the components of contravariant vectors.

2.2.4   *Transformation of the covariant coordinates*

For the same, compensatory, reasons yielding the "contra-varying" transformation of the contravariant coordinates with respect to variations of the covariant bases [reflected in Eqs. (2.3), (2.13), and (2.19)] the coordinates with respect to the dual, contravariant, basis vectors, transform *covariantly*. We may therefore say that "basis vectors $\mathbf{e}_i$, as well as dual components (coordinates) X_i vary covariantly." Likewise, "vector components (coordinates) X^i, as well as dual basis vectors $\mathbf{e}_i^* = \mathbf{e}^i$ vary contra-variantly."

A similar calculation as for the contravariant components (2.10) yields a transformation for the covariant components:

$$\mathbf{x} = \sum_{i=1}^{n} X_j \mathbf{e}^j = \sum_{i=1}^{n} Y_i \mathbf{f}^i = \sum_{i=1}^{n} Y_i \sum_{j=1}^{n} b^i{}_j \mathbf{e}^j = \sum_{j=1}^{n} \left(\sum_{i=1}^{n} b^i{}_j Y_i \right) \mathbf{e}^j. \qquad (2.30)$$

Thus, by comparison we obtain

$$X_i = \sum_{j=1}^{n} b^j{}_i Y_j = \sum_{j=1}^{n} \left(a^{-1} \right)^j{}_i Y_j, \text{ and}$$

$$Y_i = \sum_{j=1}^{n} \left(b^{-1} \right)^j{}_i X_j = \sum_{j=1}^{n} a^j{}_i X_j. \qquad (2.31)$$

In short, by comparing (2.31) with (2.3), we find that the components of covariant vectors transform just like the vectors of the covariant basis vectors of "base space."

2.2.5 Orthonormal bases

For orthonormal bases of n-dimensional Hilbert space,

$$\delta^j_i = \mathbf{e}_i \cdot \mathbf{e}^j \text{ if and only if } \mathbf{e}_i = \mathbf{e}^i \text{ for all } 1 \le i, j \le n. \qquad (2.32)$$

Therefore, the vector space and its dual vector space are "identical" in the sense that the coordinate tuples representing their bases are identical (though relatively transposed). That is, besides transposition, the two bases are identical

$$\mathfrak{B} \equiv \mathfrak{B}^* \qquad (2.33)$$

and formally any distinction between covariant and contravariant vectors becomes irrelevant. Conceptually, such a distinction persists, though. In this sense, we might "forget about the difference between covariant and contravariant orders."

2.3 Tensor as multilinear form

A *multilinear form* $\alpha : \mathfrak{V}^k \mapsto \mathbb{R}$ or $\mathbb{C}$ is a map from (multiple) arguments $\mathbf{x}_i$ which are elements of some vector space $\mathfrak{V}$ into some scalars in $\mathbb{R}$ or $\mathbb{C}$, satisfying

$$\alpha(\mathbf{x}_1, \mathbf{x}_2, \ldots, A\mathbf{y} + B\mathbf{z}, \ldots, \mathbf{x}_k) = A\alpha(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{y}, \ldots, \mathbf{x}_k)$$
$$+ B\alpha(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{z}, \ldots, \mathbf{x}_k) \qquad (2.34)$$

for every one of its (multi-)arguments.

Note that linear functionals on $\mathfrak{V}$, which constitute the elements of the dual space $\mathfrak{V}^*$ (cf. Section 1.8 on page 15) is just a particular example of a multilinear form – indeed rather a linear form – with just one argument, a vector in $\mathfrak{V}$.

In what follows we shall concentrate on *real-valued* multilinear forms which map k vectors in $\mathbb{R}^n$ into $\mathbb{R}$.

2.4   Covariant tensors

Mind the notation introduced earlier; in particular in Eqs. (2.1) and (2.2). A covariant tensor of rank k

$$\alpha : \mathfrak{V}^k \mapsto \mathbb{R} \tag{2.35}$$

is a multilinear form

$$\alpha(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} X_1^{i_1} X_2^{i_2} \ldots X_k^{i_k} \alpha(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_k}). \tag{2.36}$$

The

$$A_{i_1 i_2 \cdots i_k} \overset{\text{def}}{=} \alpha(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_k}) \tag{2.37}$$

are the *covariant components* or *covariant coordinates* of the tensor α with respect to the basis $\mathfrak{B}$.

Note that, as each of the k arguments of a tensor of type (or rank) k has to be evaluated at each of the n basis vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ in an n-dimensional vector space, $A_{i_1 i_2 \cdots i_k}$ has n^k coordinates.

To prove that tensors are multilinear forms, insert

$$\alpha(\mathbf{x}_1, \mathbf{x}_2, \ldots, A\mathbf{x}_j^1 + B\mathbf{x}_j^2, \ldots, \mathbf{x}_k)$$

$$= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} X_1^{i_i} X_2^{i_2} \ldots [A(X^1)_j^{i_j} + B(X^2)_j^{i_j}] \ldots X_k^{i_k} \alpha(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_j}, \ldots, \mathbf{e}_{i_k})$$

$$= A \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} X_1^{i_i} X_2^{i_2} \ldots (X^1)_j^{i_j} \ldots X_k^{i_k} \alpha(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_j}, \ldots, \mathbf{e}_{i_k})$$

$$+ B \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} X_1^{i_i} X_2^{i_2} \ldots (X^2)_j^{i_j} \ldots X_k^{i_k} \alpha(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_j}, \ldots, \mathbf{e}_{i_k})$$

$$= A\alpha(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j^1, \ldots, \mathbf{x}_k) + B\alpha(\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_j^2, \ldots, \mathbf{x}_k)$$

2.4.1   Transformation of covariant tensor components

Because of multilinearity and by insertion into (2.3),

$$\alpha(\mathbf{f}_{j_1}, \mathbf{f}_{j_2}, \ldots, \mathbf{f}_{j_k}) = \alpha\left(\sum_{i_1=1}^{n} a^{i_1}{}_{j_1} \mathbf{e}_{i_1}, \sum_{i_2=1}^{n} a^{i_2}{}_{j_2} \mathbf{e}_{i_2}, \ldots, \sum_{i_k=1}^{n} a^{i_k}{}_{j_k} \mathbf{e}_{i_k} \right)$$

$$= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} a^{i_1}{}_{j_1} a^{i_2}{}_{j_2} \cdots a^{i_k}{}_{j_k} \alpha(\mathbf{e}_{i_1}, \mathbf{e}_{i_2}, \ldots, \mathbf{e}_{i_k}) \tag{2.38}$$

or

$$A'_{j_1 j_2 \cdots j_k} = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} a^{i_1}{}_{j_1} a^{i_2}{}_{j_2} \cdots a^{i_k}{}_{j_k} A_{i_1 i_2 \ldots i_k}. \tag{2.39}$$

2.5   Contravariant tensors

Recall the inverse scaling of contravariant vector coordinates with respect to covariantly varying basis vectors. Recall further that the dual base vectors are defined in terms of the base vectors by a kind of "inversion" of the latter, as expressed by $[\mathbf{e}_i, \mathbf{e}_j^*] = \delta_{ij}$ in Eq. 1.39. Thus, by analogy it can be expected that similar considerations apply to the scaling of dual base vectors with respect to the scaling of covariant base

vectors: in order to compensate those scale changes, dual basis vectors should contra-vary, and, again analogously, their respective dual coordinates, as well as the dual vectors, should vary covariantly. Thus, both vectors in the dual space, as well as their components or coordinates, will be called covariant vectors, as well as covariant coordinates, respectively.

2.5.1　*Definition of contravariant tensors*

The entire tensor formalism developed so far can be transferred and applied to define *contravariant* tensors as multinear forms with contravariant components

$$\beta : \mathfrak{V}^{*\,k} \mapsto \mathbb{R} \tag{2.40}$$

by

$$\beta(\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^k) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} X_{i_1}^1 X_{i_2}^2 \ldots X_{i_k}^k \beta(\mathbf{e}^{i_1}, \mathbf{e}^{i_2}, \ldots, \mathbf{e}^{i_k}). \tag{2.41}$$

By definition

$$B^{i_1 i_2 \cdots i_k} = \beta(\mathbf{e}^{i_1}, \mathbf{e}^{i_2}, \ldots, \mathbf{e}^{i_k}) \tag{2.42}$$

are the contravariant *components* of the contravariant tensor β with respect to the basis $\mathfrak{B}^*$.

2.5.2　*Transformation of contravariant tensor components*

The argument concerning transformations of covariant tensors and components can be carried through to the contravariant case. Hence, the contravariant components transform as

$$\beta(\mathbf{f}^{j_1}, \mathbf{f}^{j_2}, \ldots, \mathbf{f}^{j_k}) = \beta\left(\sum_{i_1=1}^{n} b^{j_1}{}_{i_1} \mathbf{e}^{i_1}, \sum_{i_2=1}^{n} b^{j_2}{}_{i_2} \mathbf{e}^{i_2}, \ldots, \sum_{i_k=1}^{n} b^{j_k}{}_{i_k} \mathbf{e}^{i_k} \right)$$

$$= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} b^{j_1}{}_{i_1} b^{j_2}{}_{i_2} \cdots b^{j_k}{}_{i_k} \beta(\mathbf{e}^{i_1}, \mathbf{e}^{i_2}, \ldots, \mathbf{e}^{i_k}) \tag{2.43}$$

or

$$B'^{j_1 j_2 \cdots j_k} = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_k=1}^{n} b^{j_1}{}_{i_1} b^{j_2}{}_{i_2} \cdots b^{j_k}{}_{i_k} B^{i_1 i_2 \cdots i_k}. \tag{2.44}$$

Note that, by Eq. (2.28), $b^j{}_i = \left(a^{-1}\right)^j{}_i$.

2.6　*General tensor*

A (general) Tensor T can be defined as a multilinear form on the r-fold product of a vector space $\mathfrak{V}$, times the s-fold product of the dual vector space $\mathfrak{V}^*$; that is,

$$T : (\mathfrak{V})^r \times \left(\mathfrak{V}^*\right)^s = \underbrace{\mathfrak{V} \times \cdots \times \mathfrak{V}}_{r \text{ copies}} \times \underbrace{\mathfrak{V}^* \times \cdots \times \mathfrak{V}^*}_{s \text{ copies}} \mapsto \mathbb{F}, \tag{2.45}$$

where, most commonly, the scalar field $\mathbb{F}$ will be identified with the set $\mathbb{R}$ of reals, or with the set $\mathbb{C}$ of complex numbers. Thereby, r is called the *covariant order*, and s is called the *contravariant order* of T. A tensor of covariant order r and contravariant order s is then pronounced a tensor

of *type* (or *rank*) (r, s). By convention, covariant indices are denoted by *subscripts*, whereas the contravariant indices are denoted by *superscripts*.

With the standard, "inherited" addition and scalar multiplication, the set $\mathfrak{T}_r^s$ of all tensors of type (r, s) forms a linear vector space.

Note that a tensor of type $(1, 0)$ is called a *covariant vector*, or just a *vector*. A tensor of type $(0, 1)$ is called a *contravariant vector*.

Tensors can change their type by the invocation of the *metric tensor*. That is, a covariant tensor (index) i can be made into a contravariant tensor (index) j by summing over the index i in a product involving the tensor and g^{ij}. Likewise, a contravariant tensor (index) i can be made into a covariant tensor (index) j by summing over the index i in a product involving the tensor and g_{ij}.

Under basis or other linear transformations, covariant tensors with index i transform by summing over this index with (the transformation matrix) $a_i{}^j$. Contravariant tensors with index i transform by summing over this index with the inverse (transformation matrix) $(a^{-1})_i{}^j$.

2.7   Metric

A *metric* or *metric tensor g* is a measure of *distance* between two points in a vector space.

2.7.1   Definition

Formally, a metric, or metric tensor, can be defined as a functional g : $\mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}$ which maps two vectors (directing from the origin to the two points) into a scalar with the following properties:

- g is symmetric; that is, $g(\mathbf{x}, \mathbf{y}) = g(\mathbf{y}, \mathbf{x})$;

- g is bilinear; that is, $g(\alpha\mathbf{x} + \beta\mathbf{y}, \mathbf{z}) = \alpha g(\mathbf{x}, \mathbf{z}) + \beta g(\mathbf{y}, \mathbf{z})$ (due to symmetry g is also bilinear in the second argument);

- g is nondegenerate; that is, for every $\mathbf{x} \in \mathfrak{V}$, $\mathbf{x} \neq 0$, there exists a $\mathbf{y} \in \mathfrak{V}$ such that $g(\mathbf{x}, \mathbf{y}) \neq 0$.

2.7.2   Construction from a scalar product

In real Hilbert spaces the *metric* tensor can be defined *via* the scalar product by

$$g_{ij} = \langle \mathbf{e}_i \mid \mathbf{e}_j \rangle. \tag{2.46}$$

and

$$g^{ij} = \langle \mathbf{e}^i \mid \mathbf{e}^j \rangle. \tag{2.47}$$

For orthonormal bases, the metric tensor can be represented as a Kronecker delta function, and thus remains form invariant. Moreover, its covariant and contravariant components are identical; that is, $g_{ij} = \delta_{ij} = \delta^i_j = \delta^j_i = \delta^{ij} = g^{ij}$.

2.7.3   What can the metric tensor do for you?

We shall see that with the help of the metric tensor we can "raise and lower indices;" that is, we can transfor lower (covariant) indices into upper (contravariant) indices, and *vice versa*. This can be seen as follows. Because of linearity, any contravariant basis vector $\mathbf{e}^i$ can be written as a linear sum of covariant (transposed, but we do not mark transposition here) basis vectors:

$$\mathbf{e}^i = A^{ij}\mathbf{e}_j. \tag{2.48}$$

Then,

$$g^{ik} = \langle \mathbf{e}^i | \mathbf{e}^k \rangle = \langle A^{ij}\mathbf{e}_j | \mathbf{e}^k \rangle = A^{ij}\langle \mathbf{e}_j | \mathbf{e}^k \rangle = A^{ij}\delta_j^k = A^{ik} \tag{2.49}$$

and thus

$$\mathbf{e}^i = g^{ij}\mathbf{e}_j \tag{2.50}$$

and, by a similar argument,

$$\mathbf{e}_i = g_{ij}\mathbf{e}^j. \tag{2.51}$$

This property can also be used to raise or lower the indices not only of basis vectors, but also of tensor components; that is, to change from contravariant to covariant and conversely from covariant to contravariant. For example,

$$\mathbf{x} = X^i\mathbf{e}_i = X^i g_{ij}\mathbf{e}^j = X_j\mathbf{e}^j, \tag{2.52}$$

and hence $X_j = X^i g_{ij}$.

What is $g^i{}_j$? A straightforward calculation yields, through insertion of Eqs. (2.46) and (2.47), as well as the resulution of unity (in a modified form involving upper and lower indices; cf. Section 1.15 on page 36),

$$g^i{}_j = g^{ik}g_{kj} = \langle \mathbf{e}^i | \underbrace{\mathbf{e}^k \rangle \langle \mathbf{e}_k}_{\mathbb{I}} | \mathbf{e}_j \rangle = \langle \mathbf{e}^i | \mathbf{e}_j \rangle = \delta_j^i = \delta_{ij}. \tag{2.53}$$

A similar calculation yields $g_i{}^j = \delta_{ij}$.

The metric tensor has been defined in terms of the scalar product. The converse can be true as well. (Note, however, that the metric need not be positive.) In Euclidean space with the dot (scalar, inner) product the metric tensor represents the scalar product between vectors: let $\mathbf{x} = X^i\mathbf{e}_i \in \mathbb{R}^n$ and $\mathbf{y} = Y^j\mathbf{e}_j \in \mathbb{R}^n$ be two vectors. Then ("T" stands for the transpose),

$$\mathbf{x}\cdot\mathbf{y} \equiv (\mathbf{x}, \mathbf{y}) \equiv \langle \mathbf{x} | \mathbf{y} \rangle = X^i\mathbf{e}_i \cdot Y^j\mathbf{e}_j = X^i Y^j \mathbf{e}_i\cdot\mathbf{e}_j = X^i Y^j g_{ij} = X^T g Y. \tag{2.54}$$

It also characterizes the length of a vector: in the above equation, set $\mathbf{y} = \mathbf{x}$. Then,

$$\mathbf{x}\cdot\mathbf{x} \equiv (\mathbf{x}, \mathbf{x}) \equiv \langle \mathbf{x} | \mathbf{x} \rangle = X^i X^j g_{ij} \equiv X^T g X, \tag{2.55}$$

and thus

$$||x|| = \sqrt{X^i X^j g_{ij}} = \sqrt{X^T g X}. \tag{2.56}$$

The square of an infinitesimal vector $ds = ||d\mathbf{x}||$ is

$$(ds)^2 = g_{ij}dX^i dX^j = d\mathbf{x}^T g d\mathbf{x}. \tag{2.57}$$

2.7.4   Transformation of the metric tensor

Insertion into the definitions and coordinate transformations (2.8) as well as (2.7) yields

$$g_{ij} = \mathbf{e}_i \cdot \mathbf{e}_j = a'^l{}_i \mathbf{e}'_l \cdot a'^m{}_j \mathbf{e}'_m = a'^l{}_i a'^m{}_j \mathbf{e}'_l \cdot \mathbf{e}'_m$$
$$= a'^l{}_i a'^m{}_j g'_{lm} = \frac{\partial Y^l}{\partial X^i} \frac{\partial Y^m}{\partial X^j} g'_{lm}. \tag{2.58}$$

Conversely, (2.3) as well as (2.17) yields

$$g'_{ij} = \mathbf{f}_i \cdot \mathbf{f}_j = a^l{}_i \mathbf{e}_l \cdot a^m{}_j \mathbf{e}_m = a^l{}_i a^m{}_j \mathbf{e}_l \cdot \mathbf{e}_m$$
$$= a^l{}_i a^m{}_j g_{lm} = \frac{\partial X^l}{\partial Y^i} \frac{\partial X^m}{\partial Y^j} g_{lm}. \tag{2.59}$$

If the geometry (i.e., the basis) is locally orthonormal, $g_{lm} = \delta_{lm}$, then $g'_{ij} = \frac{\partial X^l}{\partial Y^i} \frac{\partial X_l}{\partial Y^j}$.

Just to check consistency with Eq. (2.53) we can compute, for suitable differentiable coordinates X and Y,

$$g_i'^j = \mathbf{f}_i \cdot \mathbf{f}^j = a^l{}_i \mathbf{e}_l \cdot (a^{-1})_m{}^j \mathbf{e}^m = a^l{}_i (a^{-1})_m{}^j \mathbf{e}_l \cdot \mathbf{e}^m$$
$$= a^l{}_i (a^{-1})_m{}^j \delta_l^m = a^l{}_i (a^{-1})_l{}^j \tag{2.60}$$
$$= \frac{\partial X^l}{\partial Y^i} \frac{\partial Y_l}{\partial X_j} = \frac{\partial X^l}{\partial Y^i} \frac{\partial Y_l}{\partial Y_i} = \delta^{lj} \delta_{li} = \delta_i^l.$$

In terms of the *Jacobian matrix* defined in Eq. (2.20) the metric tensor in Eq. (2.58) can be rewritten as

$$g = J^T g' J \equiv g_{ij} = J_{li} J_{mj} g'_{lm}. \tag{2.61}$$

The metric tensor and the Jacobian (determinant) are thus related by

$$\det g = (\det J^T)(\det g')(\det J). \tag{2.62}$$

If the manifold is embedded into an Euclidean space, then $g'_{lm} = \delta_{lm}$ and $g = J^T J$.

2.7.5   Examples

In what follows a few metrics are enumerated and briefly commented. For a more systematic treatment, see, for instance, Snapper and Troyer's *Metric Affine geometry* [1].

Note also that due to the properties of the metric tensor, its coordinate representation has to be a *symmetric matrix* with *nonzero diagonals.* For the symmetry $g(\mathbf{x}, \mathbf{y}) = g(\mathbf{y}, \mathbf{x})$ implies that $g_{ij} x_i y^j = g_{ij} y^i x_j = g_{ij} x_j y^i = g_{ji} x_i y^j$ for all coordinate tuples x^i and y^j. And for any zero diagonal entry (say, in the k'th position of the diagonal we can choose a nonzero vector $\mathbf{z}$ whose coordinates are all zero except the k'th coordinate. Then $g(\mathbf{z}, \mathbf{x}) = 0$ for all $\mathbf{x}$ in the vector space.

n-dimensional Euclidean space

$$g \equiv \{g_{ij}\} = \text{diag}(\underbrace{1, 1, \ldots, 1}_{n \text{ times}}) \tag{2.63}$$

[1] Ernst Snapper and Robert J. Troyer. *Metric Affine Geometry.* Academic Press, New York, 1971

One application in physics is quantum mechanics, where n stands for the dimension of a complex Hilbert space. Some definitions can be easily adopted to accommodate the complex numbers. E.g., axiom 5 of the scalar product becomes $(x, y) = \overline{(x, y)}$, where "$\overline{(x, y)}$" stands for complex conjugation of (x, y). Axiom 4 of the scalar product becomes $(x, \alpha y) = \overline{\alpha}(x, y)$.

Lorentz plane

$$g \equiv \{g_{ij}\} = \text{diag}(1, -1) \tag{2.64}$$

Minkowski space of dimension n

In this case the metric tensor is called the *Minkowski metric* and is often denoted by "η":

$$\eta \equiv \{\eta_{ij}\} = \text{diag}(\underbrace{1, 1, \ldots, 1}_{n-1 \text{ times}}, -1) \tag{2.65}$$

One application in physics is the theory of special relativity, where $D = 4$. Alexandrov's theorem states that the mere requirement of the preservation of zero distance (i.e., lightcones), combined with bijectivity (one-to-oneness) of the transformation law yields the Lorentz transformations [2].

Negative Euclidean space of dimension n

$$g \equiv \{g_{ij}\} = \text{diag}(\underbrace{-1, -1, \ldots, -1}_{n \text{ times}}) \tag{2.66}$$

Artinian four-space

$$g \equiv \{g_{ij}\} = \text{diag}(+1, +1, -1, -1) \tag{2.67}$$

General relativity

In general relativity, the metric tensor g is linked to the energy-mass distribution. There, it appears as the primary concept when compared to the scalar product. In the case of zero gravity, g is just the Minkowski metric (often denoted by "η") $\text{diag}(1, 1, 1, -1)$ corresponding to "flat" space-time.

The best known non-flat metric is the Schwarzschild metric

$$g \equiv \begin{pmatrix} (1 - 2m/r)^{-1} & 0 & 0 & 0 \\ 0 & r^2 & 0 & 0 \\ 0 & 0 & r^2 \sin^2 \theta & 0 \\ 0 & 0 & 0 & -(1 - 2m/r) \end{pmatrix} \tag{2.68}$$

with respect to the spherical space-time coordinates r, θ, ϕ, t.

Computation of the metric tensor of the circle of radius r

Consider the transformation from the standard orthonormal threedimensional "Cartesian" coordinates $X_1 = x$, $X_2 = y$, into polar coordinates

$X'_1 = r$, $X'_2 = \varphi$. In terms of r and φ, the Cartesian coordinates can be written as

$$
\begin{aligned}
X_1 &= r\cos\varphi \equiv X'_1 \cos X'_2, \\
X_2 &= r\sin\varphi \equiv X'_1 \sin X'_2.
\end{aligned}
\tag{2.69}
$$

Furthermore, since the basis we start with is the Cartesian orthonormal basis, $g_{ij} = \delta_{ij}$; therefore,

$$
g'_{ij} = \frac{\partial X^l}{\partial Y^i}\frac{\partial X_k}{\partial Y^j} g_{lk} = \frac{\partial X^l}{\partial Y^i}\frac{\partial X_k}{\partial Y^j}\delta_{lk} = \frac{\partial X^l}{\partial Y^i}\frac{\partial X_l}{\partial Y^j}.
\tag{2.70}
$$

More explicity, we obtain for the coordinates of the transformed metric tensor g'

$$
\begin{aligned}
g'_{11} &= \frac{\partial X^l}{\partial Y^1}\frac{\partial X_l}{\partial Y^1} \\
&= \frac{\partial(r\cos\varphi)}{\partial r}\frac{\partial(r\cos\varphi)}{\partial r} + \frac{\partial(r\sin\varphi)}{\partial r}\frac{\partial(r\sin\varphi)}{\partial r} \\
&= (\cos\varphi)^2 + (\sin\varphi)^2 = 1, \\
g'_{12} &= \frac{\partial X^l}{\partial Y^1}\frac{\partial X_l}{\partial Y^2} \\
&= \frac{\partial(r\cos\varphi)}{\partial r}\frac{\partial(r\cos\varphi)}{\partial \varphi} + \frac{\partial(r\sin\varphi)}{\partial r}\frac{\partial(r\sin\varphi)}{\partial \varphi} \\
&= (\cos\varphi)(-r\sin\varphi) + (\sin\varphi)(r\cos\varphi) = 0, \\
g'_{21} &= \frac{\partial X^l}{\partial Y^2}\frac{\partial X_l}{\partial Y^1} \\
&= \frac{\partial(r\cos\varphi)}{\partial \varphi}\frac{\partial(r\cos\varphi)}{\partial r} + \frac{\partial(r\sin\varphi)}{\partial \varphi}\frac{\partial(r\sin\varphi)}{\partial r} \\
&= (-r\sin\varphi)(\cos\varphi) + (r\cos\varphi)(\sin\varphi) = 0, \\
g'_{22} &= \frac{\partial X^l}{\partial Y^2}\frac{\partial X_l}{\partial Y^2} \\
&= \frac{\partial(r\cos\varphi)}{\partial \varphi}\frac{\partial(r\cos\varphi)}{\partial \varphi} + \frac{\partial(r\sin\varphi)}{\partial \varphi}\frac{\partial(r\sin\varphi)}{\partial \varphi} \\
&= (-r\sin\varphi)^2 + (r\cos\varphi)^2 = r^2;
\end{aligned}
\tag{2.71}
$$

that is, in matrix notation,

$$
g' = \begin{pmatrix} 1 & 0 \\ 0 & r^2 \end{pmatrix},
\tag{2.72}
$$

and thus

$$
(ds')^2 = g'_{ij}\, d\mathbf{x}'^i d\mathbf{x}'^j = (dr)^2 + r^2(d\varphi)^2.
\tag{2.73}
$$

Computation of the metric tensor of the ball

Consider the transformation from the standard orthonormal threedimensional "Cartesian" coordinates $X_1 = x$, $X_2 = y$, $X_3 = z$, into spherical coordinates $X'_1 = r$, $X'_2 = \theta$, $X'_3 = \varphi$. In terms of r, θ, φ, the Cartesian coordinates can be written as

$$
\begin{aligned}
X_1 &= r\sin\theta\cos\varphi \equiv X'_1 \sin X'_2 \cos X'_3, \\
X_2 &= r\sin\theta\sin\varphi \equiv X'_1 \sin X'_2 \sin X'_3, \\
X_3 &= r\cos\theta \equiv X'_1 \cos X'_2.
\end{aligned}
\tag{2.74}
$$

Furthermore, since the basis we start with is the Cartesian orthonormal basis, $g_{ij} = \delta_{ij}$; hence finally

$$g'_{ij} = \frac{\partial X^l}{\partial Y^i} \frac{\partial X_l}{\partial Y^j} \equiv \mathrm{diag}(1, r^2, r^2 \sin^2 \theta), \tag{2.75}$$

and

$$(ds')^2 = (dr)^2 + r^2 (d\theta)^2 + r^2 \sin^2 \theta (d\varphi)^2. \tag{2.76}$$

The expression $(ds)^2 = (dr)^2 + r^2 (d\varphi)^2$ for polar coordinates in two dimensions (i.e., $n = 2$) of Eq. (2.73) is recovered by setting $\theta = \pi/2$ and $d\theta = 0$.

Computation of the metric tensor of the Moebius strip

The parameter representation of the Moebius strip is

$$\Phi(u, v) = \begin{pmatrix} (1 + v \cos \frac{u}{2}) \sin u \\ (1 + v \cos \frac{u}{2}) \cos u \\ v \sin \frac{u}{2} \end{pmatrix}, \tag{2.77}$$

where $u \in [0, 2\pi]$ represents the position of the point on the circle, and where $2a > 0$ is the "width" of the Moebius strip, and where $v \in [-a, a]$.

$$\Phi_v = \frac{\partial \Phi}{\partial v} = \begin{pmatrix} \cos \frac{u}{2} \sin u \\ \cos \frac{u}{2} \cos u \\ \sin \frac{u}{2} \end{pmatrix}$$

$$\Phi_u = \frac{\partial \Phi}{\partial u} = \begin{pmatrix} -\frac{1}{2} v \sin \frac{u}{2} \sin u + (1 + v \cos \frac{u}{2}) \cos u \\ -\frac{1}{2} v \sin \frac{u}{2} \cos u - (1 + v \cos \frac{u}{2}) \sin u \\ \frac{1}{2} v \cos \frac{u}{2} \end{pmatrix} \tag{2.78}$$

$$\begin{aligned}
(\frac{\partial \Phi}{\partial v})^T \frac{\partial \Phi}{\partial u} &= \begin{pmatrix} \cos \frac{u}{2} \sin u \\ \cos \frac{u}{2} \cos u \\ \sin \frac{u}{2} \end{pmatrix}^T \begin{pmatrix} -\frac{1}{2} v \sin \frac{u}{2} \sin u + (1 + v \cos \frac{u}{2}) \cos u \\ -\frac{1}{2} v \sin \frac{u}{2} \cos u - (1 + v \cos \frac{u}{2}) \sin u \\ \frac{1}{2} v \cos \frac{u}{2} \end{pmatrix} \\
&= -\frac{1}{2} \left(\cos \frac{u}{2} \sin^2 u \right) v \sin \frac{u}{2} - \frac{1}{2} \left(\cos \frac{u}{2} \cos^2 u \right) v \sin \frac{u}{2} \\
&\qquad\qquad + \frac{1}{2} \sin \frac{u}{2} v \cos \frac{u}{2} = 0
\end{aligned} \tag{2.79}$$

$$\begin{aligned}
(\frac{\partial \Phi}{\partial v})^T \frac{\partial \Phi}{\partial v} &= \begin{pmatrix} \cos \frac{u}{2} \sin u \\ \cos \frac{u}{2} \cos u \\ \sin \frac{u}{2} \end{pmatrix}^T \begin{pmatrix} \cos \frac{u}{2} \sin u \\ \cos \frac{u}{2} \cos u \\ \sin \frac{u}{2} \end{pmatrix} \\
&= \cos^2 \frac{u}{2} \sin^2 u + \cos^2 \frac{u}{2} \cos^2 u + \sin^2 \frac{u}{2} = 1
\end{aligned} \tag{2.80}$$

$$\left(\frac{\partial \Phi}{\partial u}\right)^T \frac{\partial \Phi}{\partial u} =$$

$$\begin{pmatrix} -\frac{1}{2}v\sin\frac{u}{2}\sin u + \left(1+v\cos\frac{u}{2}\right)\cos u \\ -\frac{1}{2}v\sin\frac{u}{2}\cos u - \left(1+v\cos\frac{u}{2}\right)\sin u \\ \frac{1}{2}v\cos\frac{u}{2} \end{pmatrix}^T \cdot$$

$$\cdot \begin{pmatrix} -\frac{1}{2}v\sin\frac{u}{2}\sin u + \left(1+v\cos\frac{u}{2}\right)\cos u \\ -\frac{1}{2}v\sin\frac{u}{2}\cos u - \left(1+v\cos\frac{u}{2}\right)\sin u \\ \frac{1}{2}v\cos\frac{u}{2} \end{pmatrix} \quad (2.81)$$

$$= \frac{1}{4}v^2\sin^2\frac{u}{2}\sin^2 u + \cos^2 u + 2v\cos^2 u\cos\frac{u}{2} + v^2\cos^2 u\cos^2\frac{u}{2}$$

$$+ \frac{1}{4}v^2\sin^2\frac{u}{2}\cos^2 u + \sin^2 u + 2v\sin^2 u\cos\frac{u}{2} + v^2\sin^2 u\cos^2\frac{u}{2}$$

$$+ \frac{1}{4}v^2\cos^2\frac{1}{2}u = \frac{1}{4}v^2 + v^2\cos^2\frac{u}{2} + 1 + 2v\cos\frac{1}{2}u$$

$$= \left(1+v\cos\frac{u}{2}\right)^2 + \frac{1}{4}v^2$$

Thus the metric tensor is given by

$$g'_{ij} = \frac{\partial X^s}{\partial Y^i}\frac{\partial X^t}{\partial Y^j} g_{st} = \frac{\partial X^s}{\partial Y^i}\frac{\partial X^t}{\partial Y^j}\delta_{st}$$

$$\equiv \begin{pmatrix} \Phi_u \cdot \Phi_u & \Phi_v \cdot \Phi_u \\ \Phi_v \cdot \Phi_u & \Phi_v \cdot \Phi_v \end{pmatrix} = \mathrm{diag}\left(\left(1+v\cos\frac{u}{2}\right)^2 + \frac{1}{4}v^2, 1\right). \quad (2.82)$$

2.8   Decomposition of tensors

Although a tensor of type (or rank) n transforms like the tensor product of n tensors of type 1, not all type-n tensors can be decomposed into a single tensor product of n tensors of type (or rank) 1.

Nevertheless, by a generalized Schmidt decomposition (cf. page 63), any type-2 tensor can be decomposed into the sum of tensor products of two tensors of type 1.

2.9   Form invariance of tensors

A tensor (field) is form invariant with respect to some basis change if its representation in the new basis has the same form as in the old basis. For instance, if the "12122–component" $T_{12122}(x)$ of the tensor T with respect to the old basis and old coordinates x equals some function $f(x)$ (say, $f(x) = x^2$), then, a necessary condition for T to be form invariant is that, in terms of the new basis, that component $T'_{12122}(x')$ equals the same function $f(x')$ as before, but in the new coordinates x' [say, $f(x') = (x')^2$]. A sufficient condition for form invariance of T is that *all* coordinates or components of T are form invariant in that way.

Although form invariance is a gratifying feature for the reasons explained shortly, a tensor (field) needs not necessarily be form invariant with respect to all or even any (symmetry) transformation(s).

A physical motivation for the use of form invariant tensors can be given as follows. What makes some tuples (or matrix, or tensor components in general) of numbers or scalar functions a tensor? It is the

interpretation of the scalars as tensor components *with respect to a particular basis*. In another basis, if we were talking about the same tensor, the tensor components; that is, the numbers or scalar functions, would be different. Pointedly stated, the tensor coordinates represent some encoding of a multilinear function with respect to a particular basis.

Formally, the tensor coordinates are numbers; that is, scalars, which are grouped together in vector touples or matrices or whatever form we consider useful. As the tensor coordinates are scalars, they can be treated as scalars. For instance, due to commutativity and associativity, one can exchange their order. (Notice, though, that this is generally not the case for differential operators such as $\partial_i = \partial/\partial \mathbf{x}^i$.)

A *form invariant* tensor with respect to certain transformations is a tensor which retains the same functional form if the transformations are performed; that is, if the basis changes accordingly. That is, in this case, the functional form of mapping numbers or coordinates or other entities remains unchanged, regardless of the coordinate change. Functions remain the same but with the new parameter components as arguement. For instance; $4 \mapsto 4$ and $f(X_1, X_2, X_3) \mapsto f(Y_1, Y_2, Y_3)$.

Furthermore, if a tensor is invariant with respect to one transformation, it need not be invariant with respect to another transformation, or with respect to changes of the scalar product; that is, the metric.

Nevertheless, totally symmetric (antisymmetric) tensors remain totally symmetric (antisymmetric) in all cases:

$$A_{i_1 i_2 \ldots i_s i_t \ldots i_k} = \pm A_{i_1 i_2 \ldots i_t i_s \ldots i_k} \tag{2.83}$$

implies

$$\begin{aligned}
A'_{j_1 i_2 \ldots j_s j_t \ldots j_k} &= a^{i_1}{}_{j_1} a^{i_2}{}_{j_2} \cdots a_{j_s}{}^{i_s} a_{j_t}{}^{i_t} \cdots a^{i_k}{}_{j_k} A_{i_1 i_2 \ldots i_s i_t \ldots i_k} \\
&= \pm a^{i_1}{}_{j_1} a^{i_2}{}_{j_2} \cdots a_{j_s}{}^{i_s} a_{j_t}{}^{i_t} \cdots a^{i_k}{}_{j_k} A_{i_1 i_2 \ldots i_t i_s \ldots i_k} \\
&= \pm a^{i_1}{}_{j_1} a^{i_2}{}_{j_2} \cdots a_{j_t}{}^{i_t} a_{j_s}{}^{i_s} \cdots a^{i_k}{}_{j_k} A_{i_1 i_2 \ldots i_t i_s \ldots i_k} \\
&= \pm A'_{j_1 i_2 \ldots j_t j_s \ldots j_k}.
\end{aligned} \tag{2.84}$$

In physics, it would be nice if the natural laws could be written into a form which does not depend on the particular reference frame or basis used. Form invariance thus is a gratifying physical feature, reflecting the *symmetry* against changes of coordinates and bases.

After all, physicists want the formalization of their fundamental laws not to artificially depend on, say, spacial directions, or on some particular basis, if there is no physical reason why this should be so. Therefore, physicists tend to be crazy to write down everything in a form invariant manner.

One strategy to accomplish form invariance is to start out with form invariant tensors and compose – by tensor products and index reduction – everything from them. This method guarantees form invarince.

The "simplest" form invariant tensor under all transformations is the constant tensor of rank 0.

Another constant form invariant tensor under all transformations is represented by the Kronecker symbol δ^i_j, because

$$(\delta')^i_j = (a^{-1})_{i'}{}^i a_j{}^{j'} \delta^{i'}_{j'} = (a^{-1})_{i'}{}^i a_j{}^{i'} = a_j{}^{i'} (a^{-1})_{i'}{}^i = \delta^i_j. \tag{2.85}$$

A simple form invariant tensor field is a vector $\mathbf{x}$, because if $T(\mathbf{x}) = x^i t_i = x^i \mathbf{e}_i = \mathbf{x}$, then the "inner transformation" $\mathbf{x} \mapsto \mathbf{x}'$ and the "outer transformation" $T \mapsto T' = \mathbf{A}T$ just compensate each other; that is, in coordinate representation, Eqs.(2.11) and (2.39) yield

$$T'(\mathbf{x}') = x'^i t'_i = (a^{-1})_l{}^i x^l a_i{}^j t_j = (a^{-1})_l{}^i a_i{}^j \mathbf{e}_j x^l = \delta_l^j \mathbf{e}_j x^l = \mathbf{x} = T(\mathbf{x}). \tag{2.86}$$

For the sake of another demonstration of form invariance, consider the following two factorizable tensor fields: while

$$S(x) = \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} \otimes \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix}^T = (x_2, -x_1)^T \otimes (x_2, -x_1) \equiv \begin{pmatrix} x_2^2 & -x_1 x_2 \\ -x_1 x_2 & x_1^2 \end{pmatrix} \tag{2.87}$$

is a form invariant tensor field with respect to the basis $\{(0,1),(1,0)\}$ and orthogonal transformations (rotations around the origin)

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}, \tag{2.88}$$

$$T(x) = \begin{pmatrix} x_2 \\ x_1 \end{pmatrix} \otimes \begin{pmatrix} x_2 \\ x_1 \end{pmatrix}^T = (x_2, x_1)^T \otimes (x_2, x_1) \equiv \begin{pmatrix} x_2^2 & x_1 x_2 \\ x_1 x_2 & x_1^2 \end{pmatrix} \tag{2.89}$$

is not.

This can be proven by considering the single factors from which S and T are composed. Eqs. (2.38)-(2.39) and (2.43)-(2.44) show that the form invariance of the factors implies the form invariance of the tensor products.

For instance, in our example, the factors $(x_2, -x_1)^T$ of S are invariant, as they transform as

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x_2 \\ -x_1 \end{pmatrix} = \begin{pmatrix} x_2 \cos\varphi - x_1 \sin\varphi \\ -x_2 \sin\varphi - x_1 \cos\varphi \end{pmatrix} = \begin{pmatrix} x'_2 \\ -x'_1 \end{pmatrix},$$

where the transformation of the coordinates

$$\begin{pmatrix} x'_1 \\ x'_2 \end{pmatrix} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \cos\varphi + x_2 \sin\varphi \\ -x_1 \sin\varphi + x_2 \cos\varphi \end{pmatrix}$$

has been used.

Note that the notation identifying tensors of type (or rank) two with matrices, creates an "artifact" insofar as the transformation of the "second index" must then be represented by the exchanged multiplication order, together with the transposed transformation matrix; that is,

$$a_{ik} a_{jl} A_{kl} = a_{ik} A_{kl} a_{jl} = a_{ik} A_{kl} (a^T)_{lj} \equiv a \cdot A \cdot a^T.$$

Thus for a transformation of the transposed touple $(x_2, -x_1)$ we must consider the *transposed* transformation matrix arranged *after* the factor; that is,

$$(x_2, -x_1) \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} = (x_2 \cos\varphi - x_1 \sin\varphi, -x_2 \sin\varphi - x_1 \cos\varphi) = (x'_2, -x'_1).$$

In contrast, a similar calculation shows that the factors $(x_2, x_1)^T$ of T do not transform invariantly. However, noninvariance with respect to

certain transformations does not imply that T is not a valid, "respectable" tensor field; it is just not form invariant under rotations.

Nevertheless, note again that, while the tensor product of form invariant tensors is again a form invariant tensor, not every form invariant tensor might be decomposed into products of form invariant tensors.

Let $|+\rangle \equiv (0,1)$ and $|-\rangle \equiv (1,0)$. For a nondecomposable tensor, consider the sum of two-partite tensor products (associated with two "entangled" particles) Bell state (cf. Eq. (1.72) on page 25) in the standard basis

$$
\begin{aligned}
|\Psi^-\rangle &= \frac{1}{\sqrt{2}} \left(|+-\rangle - |-+\rangle \right) \\
&\equiv \left(0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}, 0 \right) \\
&\equiv \frac{1}{2} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
\end{aligned}
\tag{2.90}
$$

Why is $|\Psi^-\rangle$ not decomposable? In order to be able to answer this question (see also Section 1.10.3 on page 24), consider the most general two-partite state

$$
|\psi\rangle = \psi_{--}|--\rangle + \psi_{-+}|-+\rangle + \psi_{+-}|+-\rangle + \psi_{++}|++\rangle,
\tag{2.91}
$$

with $\psi_{ij} \in \mathbb{C}$, and compare it to the most general state obtainable through products of single-partite states $|\phi_1\rangle = \alpha_-|-\rangle + \alpha_+|+\rangle$, and $|\phi_2\rangle = \beta_-|-\rangle + \beta_+|+\rangle$ with $\alpha_i, \beta_i \in \mathbb{C}$; that is,

$$
\begin{aligned}
|\phi\rangle &= |\phi_1\rangle|\phi_2\rangle \\
&= (\alpha_-|-\rangle + \alpha_+|+\rangle)(\beta_-|-\rangle + \beta_+|+\rangle) \\
&= \alpha_-\beta_-|--\rangle + \alpha_-\beta_+|-+\rangle + \alpha_+\beta_-|+-\rangle + \alpha_+\beta_+|++\rangle.
\end{aligned}
\tag{2.92}
$$

Since the two-partite basis states

$$
\begin{aligned}
|--\rangle &\equiv (1,0,0,0), \\
|-+\rangle &\equiv (0,1,0,0), \\
|+-\rangle &\equiv (0,0,1,0), \\
|++\rangle &\equiv (0,0,0,1)
\end{aligned}
\tag{2.93}
$$

are linear independent (indeed, orthonormal), a comparison of $|\psi\rangle$ with $|\phi\rangle$ yields

$$
\begin{aligned}
\psi_{--} &= \alpha_-\beta_-, \\
\psi_{-+} &= \alpha_-\beta_+, \\
\psi_{+-} &= \alpha_+\beta_-, \\
\psi_{++} &= \alpha_+\beta_+.
\end{aligned}
\tag{2.94}
$$

Hence, $\psi_{--}/\psi_{-+} = \beta_-/\beta_+ = \psi_{+-}/\psi_{++}$, and thus a necessary and sufficient condition for a two-partite quantum state to be decomposable into a product of single-particle quantum states is that its amplitudes obey

$$
\psi_{--}\psi_{++} = \psi_{-+}\psi_{+-}.
\tag{2.95}
$$

$|\Psi^-\rangle$, together with the other three Bell states $|\Psi^+\rangle = \frac{1}{\sqrt{2}} \left(|+-\rangle + |-+\rangle \right)$, $|\Phi^+\rangle = \frac{1}{\sqrt{2}} \left(|--\rangle + |++\rangle \right)$, and $|\Phi^-\rangle = \frac{1}{\sqrt{2}} \left(|--\rangle - |++\rangle \right)$, forms an orthonormal basis of $\mathbb{C}^4$.

This is not satisfied for the Bell state $|\Psi^-\rangle$ in Eq. (2.90), because in this case $\psi_{--} = \psi_{++} = 0$ and $\psi_{-+} = -\psi_{+-} = 1/\sqrt{2}$. Such nondecomposability is in physics referred to as *entanglement* [3].

Note also that $|\Psi^-\rangle$ is a *singlet state*, as it is form invariant under the following generalized rotations in two-dimensional complex Hilbert subspace; that is, (if you do not believe this please check yourself)

$$|+\rangle = e^{i\frac{\varphi}{2}}\left(\cos\frac{\theta}{2}|+'\rangle - \sin\frac{\theta}{2}|-'\rangle\right),$$
$$|-\rangle = e^{-i\frac{\varphi}{2}}\left(\sin\frac{\theta}{2}|+'\rangle + \cos\frac{\theta}{2}|-'\rangle\right) \tag{2.96}$$

in the spherical coordinates θ, φ, but it cannot be composed or written as a product of a *single* (let alone form invariant) two-partite tensor product.

In order to prove form invariance of a constant tensor, one has to transform the tensor according to the standard transformation laws (2.39) and (2.42), and compare the result with the input; that is, with the untransformed, original, tensor. This is sometimes referred to as the "outer transformation."

In order to prove form invariance of a tensor field, one has to additionally transform the spatial coordinates on which the field depends; that is, the arguments of that field; and then compare. This is sometimes referred to as the "inner transformation." This will become clearer with the following example.

Consider again the tensor field defined earlier in Eq. (2.87), but let us not choose the "elegant" ways of proving form invariance by factoring; rather we explicitly consider the transformation of all the components

$$S_{ij}(x_1, x_2) = \begin{pmatrix} -x_1 x_2 & -x_2^2 \\ x_1^2 & x_1 x_2 \end{pmatrix}$$

with respect to the standard basis $\{(1,0),(0,1)\}$.

Is S form invariant with respect to rotations around the origin? That is, S should be form invariant with repect to transformations $x_i' = a_{ij}x_j$ with

$$a_{ij} = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}.$$

Consider the "outer" transformation first. As has been pointed out earlier, the term on the right hand side in $S_{ij}' = a_{ik}a_{jl}S_{kl}$ can be rewritten as a product of three matrices; that is,

$$a_{ik}a_{jl}S_{kl}(x_n) = a_{ik}S_{kl}a_{jl} = a_{ik}S_{kl}\left(a^T\right)_{lj} \equiv a \cdot S \cdot a^T.$$

a^T stands for the transposed matrix; that is, $(a^T)_{ij} = a_{ji}$.

$$\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} -x_1 x_2 & -x_2^2 \\ x_1^2 & x_1 x_2 \end{pmatrix}\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} =$$

$$= \begin{pmatrix} -x_1 x_2 \cos\varphi + x_1^2 \sin\varphi & -x_2^2 \cos\varphi + x_1 x_2 \sin\varphi \\ x_1 x_2 \sin\varphi + x_1^2 \cos\varphi & x_2^2 \sin\varphi + x_1 x_2 \cos\varphi \end{pmatrix}\begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} =$$

[3] Erwin Schrödinger. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(04):555–563, 1935a. DOI: 10.1017/S0305004100013554. URL http://doi.org/10.1017/S0305004100013554; Erwin Schrödinger. Probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32(03):446–452, 1936. DOI: 10.1017/S0305004100019137. URL http://doi.org/10.1017/S0305004100019137; and Erwin Schrödinger. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften*, 23:807–812, 823–828, 844–849, 1935b. DOI: 10.1007/BF01491891, 10.1007/BF01491914, 10.1007/BF01491987. URL http://doi.org/10.1007/BF01491891, http://doi.org/10.1007/BF01491914, http://doi.org/10.1007/BF01491987

$$
= \left(
\begin{array}{cc}
\cos\varphi\left(-x_1 x_2 \cos\varphi + x_1^2 \sin\varphi\right)+ & -\sin\varphi\left(-x_1 x_2 \cos\varphi + x_1^2 \sin\varphi\right)+ \\
+\sin\varphi\left(-x_2^2 \cos\varphi + x_1 x_2 \sin\varphi\right) & +\cos\varphi\left(-x_2^2 \cos\varphi + x_1 x_2 \sin\varphi\right) \\
\\
\cos\varphi\left(x_1 x_2 \sin\varphi + x_1^2 \cos\varphi\right)+ & -\sin\varphi\left(x_1 x_2 \sin\varphi + x_1^2 \cos\varphi\right)+ \\
+\sin\varphi\left(x_2^2 \sin\varphi + x_1 x_2 \cos\varphi\right) & +\cos\varphi\left(x_2^2 \sin\varphi + x_1 x_2 \cos\varphi\right)
\end{array}
\right) =
$$

$$
= \left(
\begin{array}{cc}
x_1 x_2\left(\sin^2\varphi - \cos^2\varphi\right)+ & 2x_1 x_2 \sin\varphi\cos\varphi \\
+\left(x_1^2 - x_2^2\right)\sin\varphi\cos\varphi & -x_1^2 \sin^2\varphi - x_2^2 \cos^2\varphi \\
\\
2x_1 x_2 \sin\varphi\cos\varphi+ & -x_1 x_2\left(\sin^2\varphi - \cos^2\varphi\right)- \\
+x_1^2 \cos^2\varphi + x_2^2 \sin^2\varphi & -\left(x_1^2 - x_2^2\right)\sin\varphi\cos\varphi
\end{array}
\right)
$$

Let us now perform the "inner" transform

$$
x_i' = a_{ij} x_j \Longrightarrow \quad
\begin{array}{rcl}
x_1' & = & x_1 \cos\varphi + x_2 \sin\varphi \\
x_2' & = & -x_1 \sin\varphi + x_2 \cos\varphi.
\end{array}
$$

Thereby we assume (to be corroborated) that the functional form in the new coordinates are identical to the functional form of the old coordinates. A comparison yields

$$
\begin{array}{rcl}
-x_1' x_2' & = & -\left(x_1 \cos\varphi + x_2 \sin\varphi\right)\left(-x_1 \sin\varphi + x_2 \cos\varphi\right) = \\
& = & -\left(-x_1^2 \sin\varphi\cos\varphi + x_2^2 \sin\varphi\cos\varphi - x_1 x_2 \sin^2\varphi + x_1 x_2 \cos^2\varphi\right) = \\
& = & x_1 x_2\left(\sin^2\varphi - \cos^2\varphi\right) + \left(x_1^2 - x_2^2\right)\sin\varphi\cos\varphi \\
(x_1')^2 & = & \left(x_1 \cos\varphi + x_2 \sin\varphi\right)\left(x_1 \cos\varphi + x_2 \sin\varphi\right) = \\
& = & x_1^2 \cos^2\varphi + x_2^2 \sin^2\varphi + 2x_1 x_2 \sin\varphi\cos\varphi \\
(x_2')^2 & = & \left(-x_1 \sin\varphi + x_2 \cos\varphi\right)\left(-x_1 \sin\varphi + x_2 \cos\varphi\right) = \\
& = & x_1^2 \sin^2\varphi + x_2^2 \cos^2\varphi - 2x_1 x_2 \sin\varphi\cos\varphi
\end{array}
$$

and hence

$$
S'(x_1', x_2') = \left(
\begin{array}{cc}
-x_1' x_2' & -(x_2')^2 \\
(x_1')^2 & x_1' x_2'
\end{array}
\right)
$$

is invariant with respect to rotations by angles φ, yielding the new basis $\{(\cos\varphi, -\sin\varphi), (\sin\varphi, \cos\varphi)\}$.

Incidentally, as has been stated earlier, $S(x)$ can be written as the product of two invariant tensors $b_i(x)$ and $c_j(x)$:

$$
S_{ij}(x) = b_i(x) c_j(x),
$$

with $b(x_1, x_2) = (-x_2, x_1)$, and $c(x_1, x_2) = (x_1, x_2)$. This can be easily checked by comparing the components:

$$
\begin{array}{rcl}
b_1 c_1 & = & -x_1 x_2 = S_{11}, \\
b_1 c_2 & = & -x_2^2 = S_{12}, \\
b_2 c_1 & = & x_1^2 = S_{21}, \\
b_2 c_2 & = & x_1 x_2 = S_{22}.
\end{array}
$$

Under rotations, b and c transform into

$$a_{ij}b_j = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} -x_2 \\ x_1 \end{pmatrix} = \begin{pmatrix} -x_2\cos\varphi + x_1\sin\varphi \\ x_2\sin\varphi + x_1\cos\varphi \end{pmatrix} = \begin{pmatrix} -x_2' \\ x_1' \end{pmatrix}$$

$$a_{ij}c_j = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1\cos\varphi + x_2\sin\varphi \\ -x_1\sin\varphi + x_2\cos\varphi \end{pmatrix} = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}.$$

This factorization of S is nonunique, since Eq. (2.87) uses a different factorization; also, S is decomposible into, for example,

$$S(x_1, x_2) = \begin{pmatrix} -x_1 x_2 & -x_2^2 \\ x_1^2 & x_1 x_2 \end{pmatrix} = \begin{pmatrix} -x_2^2 \\ x_1 x_2 \end{pmatrix} \otimes \left(\frac{x_1}{x_2}, 1 \right).$$

2.10   The Kronecker symbol δ

For vector spaces of dimension n the totally symmetric Kronecker symbol δ, sometimes referred to as the delta symbol δ–tensor, can be defined by

$$\delta_{i_1 i_2 \cdots i_k} = \begin{cases} +1 & \text{if } i_1 = i_2 = \cdots = i_k \\ 0 & \text{otherwise (that is, some indices are not identical).} \end{cases}$$
(2.97)

Note that, with the Einstein summation convention,

$$\delta_{ij}a_j = a_j\delta_{ij} = \delta_{i1}a_1 + \delta_{i2}a_2 + \cdots + \delta_{in}a_n = a_i,$$
$$\delta_{ji}a_j = a_j\delta_{ji} = \delta_{1i}a_1 + \delta_{2i}a_2 + \cdots + \delta_{ni}a_n = a_i.$$
(2.98)

2.11   The Levi-Civita symbol ε

For vector spaces of dimension n the totally antisymmetric Levi-Civita symbol ε, sometimes referred to as the Levi-Civita symbol ε–tensor, can be defined by the number of permutations of its indices; that is,

$$\varepsilon_{i_1 i_2 \cdots i_k} = \begin{cases} +1 & \text{if } (i_1 i_2 \ldots i_k) \text{ is an } even \text{ permutation of } (1, 2, \ldots k) \\ -1 & \text{if } (i_1 i_2 \ldots i_k) \text{ is an } odd \text{ permutation of } (1, 2, \ldots k) \\ 0 & \text{otherwise (that is, some indices are identical).} \end{cases}$$
(2.99)

Hence, $\varepsilon_{i_1 i_2 \cdots i_k}$ stands for the the sign of the permutation in the case of a permutation, and zero otherwise.

In two dimensions,

$$\varepsilon_{ij} \equiv \begin{pmatrix} \varepsilon_{11} & \varepsilon_{12} \\ \varepsilon_{21} & \varepsilon_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

In threedimensional Euclidean space, the cross product, or vector product of two vectors $\mathbf{x} \equiv x_i$ and $\mathbf{y} \equiv y_i$ can be written as $\mathbf{x} \times \mathbf{y} \equiv \varepsilon_{ijk} x_j y_k$.

For a direct proof, consider, for arbirtrary threedimensional vectors $\mathbf{x}$ and $\mathbf{y}$, and by enumerating all nonvanishing terms; that is, all permuta-

tions,

$$\mathbf{x} \times \mathbf{y} \equiv \varepsilon_{ijk} x_j y_k \equiv \begin{pmatrix} \varepsilon_{123} x_2 y_3 + \varepsilon_{132} x_3 y_2 \\ \varepsilon_{213} x_1 y_3 + \varepsilon_{231} x_3 y_1 \\ \varepsilon_{312} x_2 y_3 + \varepsilon_{321} x_3 y_2 \end{pmatrix}$$

$$= \begin{pmatrix} \varepsilon_{123} x_2 y_3 - \varepsilon_{123} x_3 y_2 \\ -\varepsilon_{123} x_1 y_3 + \varepsilon_{123} x_3 y_1 \\ \varepsilon_{123} x_2 y_3 - \varepsilon_{123} x_3 y_2 \end{pmatrix} = \begin{pmatrix} x_2 y_3 - x_3 y_2 \\ -x_1 y_3 + x_3 y_1 \\ x_2 y_3 - x_3 y_2 \end{pmatrix}. \tag{2.100}$$

2.12   The nabla, Laplace, and D'Alembert operators

The *nabla operator*

$$\nabla_i \equiv \left(\frac{\partial}{\partial X^1}, \frac{\partial}{\partial X^2}, \dots, \frac{\partial}{\partial X^n} \right). \tag{2.101}$$

is a vector differential operator in an n-dimensional vector space $\mathfrak{V}$. In index notation, ∇_i is also written as

$$\nabla_i = \partial_i = \partial_{X^i} = \frac{\partial}{\partial X^i}. \tag{2.102}$$

Why is the lower index indicating covariance used when differentiation with respect to upper indexed, contravariant coordinates? The nabla operator transforms in the following manners: $\nabla_i = \partial_i = \partial_{X^i}$ transforms like a *covariant* basis vector [compare with Eqs. (2.8) and (2.19)], since

$$\partial_i = \frac{\partial}{\partial X^i} = \frac{\partial Y^j}{\partial X^i} \frac{\partial}{\partial Y^j} = \frac{\partial Y^j}{\partial X^i} \partial'_i = (a^{-1})_i{}^j \partial'_i = J_{ij} \partial'_i, \tag{2.103}$$

where J_{ij} stands for the *Jacobian matrix* defined in Eq. (2.20).

As very similar calculation demonstrates that $\partial^i = \frac{\partial}{\partial X_i}$ transforms like a *contravariant* vector.

In three dimensions and in the standard Cartesian basis with the Euclidean metric, covariant and contravariant entities coincide, and

$$\nabla = \left(\frac{\partial}{\partial X^1}, \frac{\partial}{\partial X^2}, \frac{\partial}{\partial X^3} \right) = \mathbf{e}_1 \frac{\partial}{\partial X^1} + \mathbf{e}_2 \frac{\partial}{\partial X^2} + \mathbf{e}_3 \frac{\partial}{\partial X^3} =$$
$$= \left(\frac{\partial}{\partial X_1}, \frac{\partial}{\partial X_2}, \frac{\partial}{\partial X_3} \right) = \mathbf{e}^1 \frac{\partial}{\partial X_1} + \mathbf{e}^2 \frac{\partial}{\partial X_2} + \mathbf{e}^3 \frac{\partial}{\partial X_3}. \tag{2.104}$$

It is often used to define basic differential operations; in particular, (i) to denote the *gradient* of a scalar field $f(X_1, X_2, X_3)$ (rendering a vector field with respect to a particular basis), (ii) the *divergence* of a vector field $\mathbf{v}(X_1, X_2, X_3)$ (rendering a scalar field with respect to a particular basis), and (iii) the *curl* (rotation) of a vector field $\mathbf{v}(X_1, X_2, X_3)$ (rendering a vector field with respect to a particular basis) as follows:

$$\text{grad } f \quad = \quad \nabla f = \left(\frac{\partial f}{\partial X_1}, \frac{\partial f}{\partial X_2}, \frac{\partial f}{\partial X_3} \right), \tag{2.105}$$

$$\text{div } \mathbf{v} \quad = \quad \nabla \cdot \mathbf{v} = \frac{\partial v_1}{\partial X_1} + \frac{\partial v_2}{\partial X_2} + \frac{\partial v_3}{\partial X_3}, \tag{2.106}$$

$$\text{rot } \mathbf{v} \quad = \quad \nabla \times \mathbf{v} = \left(\frac{\partial v_3}{\partial X_2} - \frac{\partial v_2}{\partial X_3}, \frac{\partial v_1}{\partial X_3} - \frac{\partial v_3}{\partial X_1}, \frac{\partial v_2}{\partial X_1} - \frac{\partial v_1}{\partial X_2} \right) \tag{2.107}$$

$$\equiv \quad \varepsilon_{ijk} \partial_j v_k. \tag{2.108}$$

The *Laplace operator* is defined by

$$\Delta = \nabla^2 = \nabla \cdot \nabla = \frac{\partial^2}{\partial^2 X_1} + \frac{\partial^2}{\partial^2 X_2} + \frac{\partial^2}{\partial^2 X_3}. \tag{2.109}$$

In special relativity and electrodynamics, as well as in wave theory and quantized field theory, with the Minkowski space-time of dimension four (referring to the metric tensor with the signature "$\pm, \pm, \pm, \mp$"), the *D'Alembert operator* is defined by the Minkowski metric $\eta = \mathrm{diag}(1, 1, 1, -1)$

$$\Box = \partial_i \partial^i = \eta_{ij}\partial^i \partial^j = \nabla^2 - \frac{\partial^2}{\partial^2 t} = \nabla \cdot \nabla - \frac{\partial^2}{\partial^2 t} = \frac{\partial^2}{\partial^2 X_1} + \frac{\partial^2}{\partial^2 X_2} + \frac{\partial^2}{\partial^2 X_3} - \frac{\partial^2}{\partial^2 t}. \tag{2.110}$$

2.13   Index trickery and examples

We have already mentioned *Einstein's summation convention* requiring that, when an index variable appears twice in a single term, one has to sum over all of the possible index values. For instance, $a_{ij}b_{jk}$ stands for $\sum_j a_{ij}b_{jk}$.

There are other tricks which are commonly used. Here, some of them are enumerated:

(i)   Indices which appear as internal sums can be renamed arbitrarily (provided their name is not already taken by some other index). That is, $a_i b^i = a_j b^j$ for arbitrary a, b, i, j.

(ii)   With the Euclidean metric, $\delta_{ii} = n$.

(iii) $\frac{\partial X^i}{\partial X^j} = \delta^i_j = \delta^{ij}$ and $\frac{\partial X_i}{\partial X_j} = \delta^j_i = \delta^{ij}$.

(iv)   With the Euclidean metric, $\frac{\partial X^i}{\partial X^i} = n$.

(v) $\varepsilon_{ij}\delta_{ij} = -\varepsilon_{ji}\delta_{ij} = -\varepsilon_{ji}\delta_{ji} = (i \leftrightarrow j) = -\varepsilon_{ij}\delta_{ij} = 0$, since $a = -a$ implies $a = 0$; likewise, $\varepsilon_{ij}x_i x_j = 0$. In general, the Einstein summations $s_{ij...}a_{ij...}$ over objects $s_{ij...}$ which are *symmetric* with respect to index exchanges over objects $a_{ij...}$ which are *antisymmetric* with respect to index exchanges yields zero.

(vi)   For threedimensional vector spaces ($n = 3$) and the Euclidean metric, the *Grassmann identity* holds:

$$\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}. \tag{2.111}$$

For the sake of a proof, consider

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z}) \equiv$$

in index notation

$$x_j \varepsilon_{ijk} y_l z_m \varepsilon_{klm} = x_j y_l z_m \varepsilon_{ijk} \varepsilon_{klm} \equiv$$

in coordinate notation

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \left[\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} \times \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \right] = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \times \begin{pmatrix} y_2 z_3 - y_3 z_2 \\ y_3 z_1 - y_1 z_3 \\ y_1 z_2 - y_2 z_1 \end{pmatrix} =$$

$$\begin{pmatrix} x_2(y_1 z_2 - y_2 z_1) - x_3(y_3 z_1 - y_1 z_3) \\ x_3(y_2 z_3 - y_3 z_2) - x_1(y_1 z_2 - y_2 z_1) \\ x_1(y_3 z_1 - y_1 z_3) - x_2(y_2 z_3 - y_3 z_2) \end{pmatrix} = \tag{2.112}$$

$$\begin{pmatrix} x_2 y_1 z_2 - x_2 y_2 z_1 - x_3 y_3 z_1 + x_3 y_1 z_3 \\ x_3 y_2 z_3 - x_3 y_3 z_2 - x_1 y_1 z_2 + x_1 y_2 z_1 \\ x_1 y_3 z_1 - x_1 y_1 z_3 - x_2 y_2 z_3 + x_2 y_3 z_2 \end{pmatrix} =$$

$$\begin{pmatrix} y_1(x_2 z_2 + x_3 z_3) - z_1(x_2 y_2 + x_3 y_3) \\ y_2(x_3 z_3 + x_1 z_1) - z_2(x_1 y_1 + x_3 y_3) \\ y_3(x_1 z_1 + x_2 z_2) - z_3(x_1 y_1 + x_2 y_2) \end{pmatrix}$$

The "incomplete" dot products can be completed through addition and subtraction of the same term, respectively; that is,

$$\begin{pmatrix} y_1(x_1 z_1 + x_2 z_2 + x_3 z_3) - z_1(x_1 y_1 + x_2 y_2 + x_3 y_3) \\ y_2(x_1 z_1 + x_2 z_2 + x_3 z_3) - z_2(x_1 y_1 + x_2 y_2 + x_3 y_3) \\ y_3(x_1 z_1 + x_2 z_2 + x_3 z_3) - z_3(x_1 y_1 + x_2 y_2 + x_3 y_3) \end{pmatrix} \equiv$$

in vector notation $\qquad$ (2.113)

$$\mathbf{y}(\mathbf{x} \cdot \mathbf{z}) - \mathbf{z}(\mathbf{x} \cdot \mathbf{y}) \equiv$$

in index notation

$$x_j y_l z_m \left(\delta_{il} \delta_{jm} - \delta_{im} \delta_{jl} \right).$$

(vii) For threedimensional vector spaces ($n = 3$) and the Euclidean metric,

$$|a \times b| = \sqrt{\varepsilon_{ijk} \varepsilon_{ist} a_j a_s b_k b_t} = \sqrt{|a|^2 |b|^2 - (a \cdot b)^2} =$$

$$\sqrt{\det \begin{pmatrix} a \cdot a & a \cdot b \\ a \cdot b & b \cdot b \end{pmatrix}} = |a||b| \sin \theta_{ab}.$$

(viii) Let $u, v \equiv X_1', X_2'$ be two parameters associated with an orthonormal Cartesian basis $\{(0, 1), (1, 0)\}$, and let $\Phi : (u, v) \mapsto \mathbb{R}^3$ be a mapping from some area of $\mathbb{R}^2$ into a twodimensional surface of $\mathbb{R}^3$. Then the metric tensor is given by $g_{ij} = \frac{\partial \Phi^k}{\partial Y^i} \frac{\partial \Phi^m}{\partial Y^j} \delta_{km}$.

Consider the following examples in threedimensional vector space. Let $r^2 = \sum_{i=1}^{3} x_i^2$.

1.

$$\partial_j r = \partial_j \sqrt{\sum_i x_i^2} = \frac{1}{2} \frac{1}{\sqrt{\sum_i x_i^2}} 2 x_j = \frac{x_j}{r} \tag{2.114}$$

By using the chain rule one obtains

$$\partial_j r^\alpha = \alpha r^{\alpha-1}\left(\partial_j r\right) = \alpha r^{\alpha-1}\left(\frac{x_j}{r}\right) = \alpha r^{\alpha-2} x_j \tag{2.115}$$

and thus $\nabla r^\alpha = \alpha r^{\alpha-2}\mathbf{x}$.

2.

$$\partial_j \log r = \frac{1}{r}\left(\partial_j r\right) \tag{2.116}$$

With $\partial_j r = \frac{x_j}{r}$ derived earlier in Eq. (2.115) one obtains $\partial_j \log r = \frac{1}{r}\frac{x_j}{r} = \frac{x_j}{r^2}$, and thus $\nabla \log r = \frac{\mathbf{x}}{r^2}$.

3.

$$\partial_j\left[\left(\sum_i (x_i - a_i)^2\right)^{-\frac{1}{2}} + \left(\sum_i (x_i + a_i)^2\right)^{-\frac{1}{2}}\right] =$$

$$= -\frac{1}{2}\left[\frac{1}{\left(\sum_i (x_i - a_i)^2\right)^{\frac{3}{2}}}2(x_j - a_j) + \frac{1}{\left(\sum_i (x_i + a_i)^2\right)^{\frac{3}{2}}}2(x_j + a_j)\right] =$$

$$-\left(\sum_i (x_i - a_i)^2\right)^{-\frac{3}{2}}(x_j - a_j) - \left(\sum_i (x_i + a_i)^2\right)^{-\frac{3}{2}}(x_j + a_j). \tag{2.117}$$

4. For three dimensions and for $r \neq 0$,

$$\nabla\left(\frac{\mathbf{r}}{r^3}\right) \equiv \partial_i\left(\frac{r_i}{r^3}\right) = \frac{1}{r^3}\underbrace{\partial_i r_i}_{=3} + r_i\left(-3\frac{1}{r^4}\right)\left(\frac{1}{2r}\right)2r_i = 3\frac{1}{r^3} - 3\frac{1}{r^3} = 0. \tag{2.118}$$

5. With this solution (2.118) one obtains, for three dimensions and $r \neq 0$,

$$\Delta\left(\frac{1}{r}\right) \equiv \partial_i\partial_i\frac{1}{r} = \partial_i\left(-\frac{1}{r^2}\right)\left(\frac{1}{2r}\right)2r_i = -\partial_i\frac{r_i}{r^3} = 0. \tag{2.119}$$

6. With the earlier solution (2.118) one obtains

$$\Delta\left(\frac{\mathbf{rp}}{r^3}\right) \equiv$$

$$\partial_i\partial_i\frac{r_j p_j}{r^3} = \partial_i\left[\frac{p_i}{r^3} + r_j p_j\left(-3\frac{1}{r^5}\right)r_i\right] =$$

$$= p_i\left(-3\frac{1}{r^5}\right)r_i + p_i\left(-3\frac{1}{r^5}\right)r_i +$$

$$+r_j p_j\left[\left(15\frac{1}{r^6}\right)\left(\frac{1}{2r}\right)2r_i\right]r_i + r_j p_j\left(-3\frac{1}{r^5}\right)\underbrace{\partial_i r_i}_{=3} = \tag{2.120}$$

$$= r_i p_i\frac{1}{r^5}(-3 - 3 + 15 - 9) = 0$$

7. With $r \neq 0$ and constant $\mathbf{p}$ one obtains

Note that, in three dimensions, the Grassmann identity (2.111) $\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$ holds.

$$\nabla \times (\mathbf{p} \times \frac{\mathbf{r}}{r^3}) \equiv \varepsilon_{ijk}\partial_j \varepsilon_{klm}p_l\frac{r_m}{r^3} = p_l\varepsilon_{ijk}\varepsilon_{klm}\left[\partial_j \frac{r_m}{r^3}\right]$$

$$= p_l\varepsilon_{ijk}\varepsilon_{klm}\left[\frac{1}{r^3}\partial_j r_m + r_m\left(-3\frac{1}{r^4}\right)\left(\frac{1}{2r}\right)2r_j\right]$$

$$= p_l\varepsilon_{ijk}\varepsilon_{klm}\left[\frac{1}{r^3}\delta_{jm} - 3\frac{r_jr_m}{r^5}\right]$$

$$= p_l(\delta_{il}\delta_{jm} - \delta_{im}\delta_{jl})\left[\frac{1}{r^3}\delta_{jm} - 3\frac{r_jr_m}{r^5}\right] \qquad (2.121)$$

$$= p_i\underbrace{\left(3\frac{1}{r^3} - 3\frac{1}{r^3}\right)}_{=0} - p_j\left(\frac{1}{r^3}\underbrace{\partial_j r_i}_{=\delta_{ij}} - 3\frac{r_jr_i}{r^5}\right)$$

$$= -\frac{\mathbf{p}}{r^3} + 3\frac{(\mathbf{rp})\mathbf{r}}{r^5}.$$

8.

$$\nabla \times (\nabla\Phi)$$

$$\equiv \varepsilon_{ijk}\partial_j\partial_k\Phi$$

$$= \varepsilon_{ikj}\partial_k\partial_j\Phi \qquad (2.122)$$

$$= \varepsilon_{ikj}\partial_j\partial_k\Phi$$

$$= -\varepsilon_{ijk}\partial_j\partial_k\Phi = 0.$$

This is due to the fact that $\partial_j\partial_k$ is symmetric, whereas ε_{ijk} ist totally antisymmetric.

9. For a proof that $(\mathbf{x} \times \mathbf{y}) \times \mathbf{z} \neq \mathbf{x} \times (\mathbf{y} \times \mathbf{z})$ consider

$$(\mathbf{x} \times \mathbf{y}) \times \mathbf{z}$$

$$\equiv \underbrace{\varepsilon_{ijm}}_{\text{second} \times}\underbrace{\varepsilon_{jkl}}_{\text{first} \times}x_k y_l z_m$$

$$= -\varepsilon_{imj}\varepsilon_{jkl}x_k y_l z_m \qquad (2.123)$$

$$= -(\delta_{ik}\delta_{ml} - \delta_{im}\delta_{lk})x_k y_l z_m$$

$$= -x_i\mathbf{y}\cdot\mathbf{z} + y_i\mathbf{x}\cdot\mathbf{z}.$$

versus

$$\mathbf{x} \times (\mathbf{y} \times \mathbf{z})$$

$$\equiv \underbrace{\varepsilon_{ikj}}_{\text{first} \times}\underbrace{\varepsilon_{jlm}}_{\text{second} \times}x_k y_l z_m$$

$$= (\delta_{il}\delta_{km} - \delta_{im}\delta_{kl})x_k y_l z_m \qquad (2.124)$$

$$= y_i\mathbf{x}\cdot\mathbf{z} - z_i\mathbf{x}\cdot\mathbf{y}.$$

10. Let $\mathbf{w} = \frac{\mathbf{p}}{r}$ with $p_i = p_i\left(t - \frac{r}{c}\right)$, whereby t and c are constants. Then,

$$\text{div}\mathbf{w} \quad = \quad \nabla\cdot\mathbf{w}$$

$$\equiv \partial_i w_i \quad = \quad \partial_i\left[\frac{1}{r}p_i\left(t - \frac{r}{c}\right)\right] =$$

$$= \quad \left(-\frac{1}{r^2}\right)\left(\frac{1}{2r}\right)2r_i p_i + \frac{1}{r}p_i'\left(-\frac{1}{c}\right)\left(\frac{1}{2r}\right)2r_i =$$

$$= \quad -\frac{r_i p_i}{r^3} - \frac{1}{cr^2}p_i'r_i.$$

Hence, $\text{div}\mathbf{w} = \nabla\cdot\mathbf{w} = -\left(\frac{\mathbf{rp}}{r^3} + \frac{\mathbf{rp'}}{cr^2}\right).$

$$\begin{aligned}
\text{rot}\,\mathbf{w} &= \nabla \times \mathbf{w} \\
\varepsilon_{ijk}\partial_j w_k &= \equiv \varepsilon_{ijk}\left[\left(-\frac{1}{r^2}\right)\left(\frac{1}{2r}\right)2r_j p_k + \frac{1}{r}p'_k\left(-\frac{1}{c}\right)\left(\frac{1}{2r}\right)2r_j\right] = \\
&= -\frac{1}{r^3}\varepsilon_{ijk}r_j p_k - \frac{1}{cr^2}\varepsilon_{ijk}r_j p'_k = \\
&\equiv -\frac{1}{r^3}\left(\mathbf{r}\times\mathbf{p}\right) - \frac{1}{cr^2}\left(\mathbf{r}\times\mathbf{p}'\right).
\end{aligned}$$

11. Let us verify some specific examples of Gauss' (divergence) theorem, stating that the outward flux of a vector field through a closed surface is equal to the volume integral of the divergence of the region inside the surface. That is, the sum of all sources subtracted by the sum of all sinks represents the net flow out of a region or volume of threedimensional space:

$$\int_V \nabla\cdot\mathbf{w}\,dv = \int_{F_V}\mathbf{w}\cdot d\mathbf{f}. \tag{2.125}$$

Consider the vector field $\mathbf{w} = \left(4x, -2y^2, z^2\right)^T$ and the (cylindric) volume bounded by the planes $z = 0$ und $z = 3$, as well as by the surface $x^2 + y^2 = 4$.

Let us first look at the left hand side $\int_V \nabla\cdot\mathbf{w}\,dv$ of Eq. (2.125):

$$\nabla\mathbf{w} = \text{div}\,\mathbf{w} = 4 - 4y + 2z$$

$$\begin{aligned}
\Longrightarrow \int_V \text{div}\,\mathbf{w}\,dv &= \int_{z=0}^{3}dz\int_{x=-2}^{2}dx\int_{y=-\sqrt{4-x^2}}^{\sqrt{4-x^2}}dy\left(4-4y+2z\right) = \\
&\quad \text{cylindric coordinates:}\begin{bmatrix} x &=& r\cos\varphi \\ y &=& r\sin\varphi \\ z &=& z \end{bmatrix} \\
&= \int_{z=0}^{3}dz\int_0^2 r\,dr\int_0^{2\pi}d\varphi\left(4-4r\sin\varphi+2z\right) = \\
&= \int_{z=0}^{3}dz\int_0^2 r\,dr\left(4\varphi+4r\cos\varphi+2\varphi z\right)\Big|_{\varphi=0}^{2\pi} = \\
&= \int_{z=0}^{3}dz\int_0^2 r\,dr\left(8\pi+4r+4\pi z-4r\right) = \\
&= \int_{z=0}^{3}dz\int_0^2 r\,dr\left(8\pi+4\pi z\right) \\
&= 2\left(8\pi z+4\pi\frac{z^2}{2}\right)\Big|_{z=0}^{z=3} = 2(24+18)\pi = 84\pi
\end{aligned}$$

Now consider the right hand side $\int_F \mathbf{w}\cdot d\mathbf{f}$ of Eq. (2.125). The surface consists of three parts: the lower plane F_1 of the cylinder is characterized by $z = 0$; the upper plane F_2 of the cylinder is characterized

by $z = 3$; the surface on the side of the zylinder F_3 is characterized by $x^2 + y^2 = 4$. $d\mathbf{f}$ must be normal to these surfaces, pointing outwards; hence

$$F_1 : \int_{\mathscr{F}_1} \mathbf{w} \cdot d\mathbf{f}_1 \;\;=\;\; \int_{\mathscr{F}_1} \begin{pmatrix} 4x \\ -2y^2 \\ z^2 = 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ -1 \end{pmatrix} d\,x d\,y = 0$$

$$F_2 : \int_{\mathscr{F}_2} \mathbf{w} \cdot d\mathbf{f}_2 \;\;=\;\; \int_{\mathscr{F}_2} \begin{pmatrix} 4x \\ -2y^2 \\ z^2 = 9 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} d\,x d\,y =$$

$$\;\;=\;\; 9 \int_{K_{r=2}} d\,f = 9 \cdot 4\pi = 36\pi$$

$$F_3 : \int_{\mathscr{F}_3} \mathbf{w} \cdot d\mathbf{f}_3 \;\;=\;\; \int_{\mathscr{F}_3} \begin{pmatrix} 4x \\ -2y^2 \\ z^2 \end{pmatrix} \left(\frac{\partial \mathbf{x}}{\partial \varphi} \times \frac{\partial \mathbf{x}}{\partial z} \right) d\varphi\, dz \quad (r = 2 = \text{const.})$$

$$\frac{\partial \mathbf{x}}{\partial \varphi} = \begin{pmatrix} -r\sin\varphi \\ r\cos\varphi \\ 0 \end{pmatrix} = \begin{pmatrix} -2\sin\varphi \\ 2\cos\varphi \\ 0 \end{pmatrix} ; \;\; \frac{\partial \mathbf{x}}{\partial z} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

$$\Longrightarrow \left(\frac{\partial \mathbf{x}}{\partial \varphi} \times \frac{\partial \mathbf{x}}{\partial z} \right) = \begin{pmatrix} 2\cos\varphi \\ 2\sin\varphi \\ 0 \end{pmatrix}$$

$$\Longrightarrow F_3 \;\;=\;\; \int_{\varphi=0}^{2\pi} d\varphi \int_{z=0}^{3} dz \begin{pmatrix} 4\cdot 2\cos\varphi \\ -2(2\sin\varphi)^2 \\ z^2 \end{pmatrix} \begin{pmatrix} 2\cos\varphi \\ 2\sin\varphi \\ 0 \end{pmatrix} =$$

$$\;\;=\;\; \int_{\varphi=0}^{2\pi} d\varphi \int_{z=0}^{3} dz \left(16\cos^2\varphi - 16\sin^3\varphi \right) =$$

$$\;\;=\;\; 3\cdot 16 \int_{\varphi=0}^{2\pi} d\varphi \left(\cos^2\varphi - \sin^3\varphi \right) =$$

$$\;\;=\;\; \left[\begin{array}{lcl} \int \cos^2\varphi\, d\varphi & = & \frac{\varphi}{2} + \frac{1}{4}\sin 2\varphi \\ \int \sin^3\varphi\, d\varphi & = & -\cos\varphi + \frac{1}{3}\cos^3\varphi \end{array} \right] =$$

$$\;\;=\;\; 3\cdot 16 \left\{ \frac{2\pi}{2} - \underbrace{\left[\left(1 + \frac{1}{3} \right) - \left(1 + \frac{1}{3} \right) \right]}_{=0} \right\} = 48\pi$$

For the flux through the surfaces one thus obtains

$$\oint_F \mathbf{w} \cdot d\mathbf{f} = F_1 + F_2 + F_3 = 84\pi.$$

12. Let us verify some specific examples of Stokes' theorem in three dimensions, stating that

$$\int_{\mathscr{F}} \mathrm{rot}\, \mathbf{b} \cdot d\mathbf{f} = \oint_{\mathscr{C}_{\mathscr{F}}} \mathbf{b} \cdot d\mathbf{s}. \qquad (2.126)$$

Consider the vector field $\mathbf{b} = \left(yz, -xz, 0\right)^T$ and the volume bounded by spherical cap formed by the plane at $z = a/\sqrt{2}$ of a sphere of radius a centered around the origin.

Let us first look at the left hand side $\int_{\mathscr{F}} \operatorname{rot} \mathbf{b} \cdot d\mathbf{f}$ of Eq. (2.126):

$$\mathbf{b} = \begin{pmatrix} yz \\ -xz \\ 0 \end{pmatrix} \Longrightarrow \operatorname{rot} \mathbf{b} = \nabla \times \mathbf{b} = \begin{pmatrix} x \\ y \\ -2z \end{pmatrix}$$

Let us transform this into spherical coordinates:

$$\mathbf{x} = \begin{pmatrix} r\sin\theta\cos\varphi \\ r\sin\theta\sin\varphi \\ r\cos\theta \end{pmatrix}$$

$$\Rightarrow \frac{\partial \mathbf{x}}{\partial \theta} = r \begin{pmatrix} \cos\theta\cos\varphi \\ \cos\theta\sin\varphi \\ -\sin\theta \end{pmatrix}; \quad \frac{\partial \mathbf{x}}{\partial \varphi} = r \begin{pmatrix} -\sin\theta\sin\varphi \\ \sin\theta\cos\varphi \\ 0 \end{pmatrix}$$

$$d\mathbf{f} = \left(\frac{\partial \mathbf{x}}{\partial \theta} \times \frac{\partial \mathbf{x}}{\partial \varphi}\right) d\theta\, d\varphi = r^2 \begin{pmatrix} \sin^2\theta\cos\varphi \\ \sin^2\theta\sin\varphi \\ \sin\theta\cos\theta \end{pmatrix} d\theta\, d\varphi$$

$$\nabla \times \mathbf{b} = r \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ -2\cos\theta \end{pmatrix}$$

$$\begin{aligned}
\int_{\mathscr{F}} \operatorname{rot} \mathbf{b} \cdot d\mathbf{f} &= \int_{\theta=0}^{\pi/4} d\theta \int_{\varphi=0}^{2\pi} d\varphi\, a^3 \begin{pmatrix} \sin\theta\cos\varphi \\ \sin\theta\sin\varphi \\ -2\cos\theta \end{pmatrix} \begin{pmatrix} \sin^2\theta\cos\varphi \\ \sin^2\theta\sin\varphi \\ \sin\theta\cos\theta \end{pmatrix} = \\
&= a^3 \int_{\theta=0}^{\pi/4} d\theta \int_{\varphi=0}^{2\pi} d\varphi \left[\sin^3\theta \underbrace{\left(\cos^2\varphi + \sin^2\varphi\right)}_{=1} - 2\sin\theta\cos^2\theta\right] = \\
&= 2\pi a^3 \left[\int_{\theta=0}^{\pi/4} d\theta \left(1 - \cos^2\theta\right)\sin\theta - 2 \int_{\theta=0}^{\pi/4} d\theta \sin\theta\cos^2\theta\right] = \\
&= 2\pi a^3 \int_{\theta=0}^{\pi/4} d\theta \sin\theta \left(1 - 3\cos^2\theta\right) = \\
&\qquad \left[\begin{array}{l} \text{transformation of variables:} \\ \cos\theta = u \Rightarrow du = -\sin\theta\, d\theta \Rightarrow d\theta = -\frac{du}{\sin\theta} \end{array}\right] \\
&= 2\pi a^3 \int_{\theta=0}^{\pi/4} (-du)\left(1 - 3u^2\right) = 2\pi a^3 \left.\left(\frac{3u^3}{3} - u\right)\right|_{\theta=0}^{\pi/4} = \\
&= 2\pi a^3 \left.\left(\cos^3\theta - \cos\theta\right)\right|_{\theta=0}^{\pi/4} = 2\pi a^3 \left(\frac{2\sqrt{2}}{8} - \frac{\sqrt{2}}{2}\right) = \\
&= \frac{2\pi a^3}{8}\left(-2\sqrt{2}\right) = -\frac{\pi a^3 \sqrt{2}}{2}
\end{aligned}$$

Now consider the right hand side $\oint_{\mathscr{C}_{\mathscr{F}}} \mathbf{b} \cdot d\mathbf{s}$ of Eq. (2.126). The radius r' of the circle surface $\{(x, y, z) \mid x, y \in \mathbb{R}, z = a/\sqrt{2}\}$ bounded by the

sphere with radius a is determined by $a^2 = (r')^2 + (a/\sqrt{2})^2$; hence, $r' = a/\sqrt{2}$. The curve of integration $\mathscr{C}_{\mathscr{F}}$ can be parameterized by

$$\{(x, y, z) \mid x = \frac{a}{\sqrt{2}}\cos\varphi, y = \frac{a}{\sqrt{2}}\sin\varphi, z = \frac{a}{\sqrt{2}}\}.$$

Therefore,

$$\mathbf{x} = a \begin{pmatrix} \frac{1}{\sqrt{2}}\cos\varphi \\ \frac{1}{\sqrt{2}}\sin\varphi \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{a}{\sqrt{2}} \begin{pmatrix} \cos\varphi \\ \sin\varphi \\ 1 \end{pmatrix} \in \mathscr{C}_{\mathscr{F}}$$

Let us transform this into polar coordinates:

$$d\mathbf{s} = \frac{d\mathbf{x}}{d\varphi}\, d\varphi = \frac{a}{\sqrt{2}} \begin{pmatrix} -\sin\varphi \\ \cos\varphi \\ 0 \end{pmatrix} d\varphi$$

$$\mathbf{b} = \begin{pmatrix} \frac{a}{\sqrt{2}}\sin\varphi \cdot \frac{a}{\sqrt{2}} \\ -\frac{a}{\sqrt{2}}\cos\varphi \cdot \frac{a}{\sqrt{2}} \\ 0 \end{pmatrix} = \frac{a^2}{2} \begin{pmatrix} \sin\varphi \\ -\cos\varphi \\ 0 \end{pmatrix}$$

Hence the circular integral is given by

$$\oint_{\mathscr{C}_F} \mathbf{b} \cdot d\mathbf{s} = \frac{a^2}{2}\frac{a}{\sqrt{2}} \int_{\varphi=0}^{2\pi} \underbrace{\left(-\sin^2\varphi - \cos^2\varphi\right)}_{=-1} d\varphi = -\frac{a^3}{2\sqrt{2}}2\pi = -\frac{a^3\pi}{\sqrt{2}}.$$

13. In machine learning, a linear regression *Ansatz* [4] is to find a linear model for the prediction of some unknown observable, given some anecdotal instances of its performance. More formally, let y be an arbitrary observable which depends on n parameters $x_1, \ldots, x_n$ by linear means; that is, by

$$y = \sum_{i=1}^{n} x_i r_i = \langle \mathbf{x} | \mathbf{r} \rangle, \tag{2.127}$$

where $\langle \mathbf{x} | = (|\mathbf{x}\rangle)^T$ is the transpose of the vector $|\mathbf{x}\rangle$. The tuple

$$|\mathbf{r}\rangle = \left(r_1, \ldots, r_n\right)^T \tag{2.128}$$

contains the unknown weights of the approximation – the "theory," if you like – and $\langle \mathbf{a} | \mathbf{b} \rangle = \sum_i a_i b_i$ stands for the Euclidean scalar product of the tuples interpreted as (dual) vectors in n-dimensional (dual) vector space $\mathbb{R}^n$.

Given are m known instances of (2.127); that is, suppose m pairs $\left(z_j, |\mathbf{x}_j\rangle\right)$ are known. These data can be bundled into an m-tuple

$$|\mathbf{z}\rangle \equiv \left(z_{j_1}, \ldots, z_{j_m}\right)^T, \tag{2.129}$$

and an $(m \times n)$-matrix

$$\mathbf{X} \equiv \begin{pmatrix} x_{j_1 i_1} & \cdots & x_{j_1 i_n} \\ \vdots & \vdots & \vdots \\ x_{j_m i_1} & \cdots & x_{j_m i_n} \end{pmatrix} \tag{2.130}$$

[4] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. November 2016. ISBN 9780262035613, 9780262337434. URL `https://mitpress.mit.edu/books/deep-learning`

where $j_1, \ldots, j_m$ are arbitrary permutations of $1, \ldots, m$, and the matrix rows are just the vectors $|\mathbf{x}_{j_k}\rangle \equiv \left(x_{j_k i_1} \quad \ldots, x_{j_k i_n} \right)^T$.

The task is to compute a "good" estimate of $|\mathbf{r}\rangle$; that is, an estimate of $|\mathbf{r}\rangle$ which allows an "optimal" computation of the prediction y.

Suppose that a good way to measure the performance of the prediction from some particular definite but unknown $|\mathbf{r}\rangle$ with respect to the m given data $\left(z_j, |\mathbf{x}_j\rangle \right)$ is by the mean squared error (MSE)

$$
\begin{aligned}
\text{MSE} &= \frac{1}{m} \left\| |\mathbf{y}\rangle - |\mathbf{z}\rangle \right\|^2 = \frac{1}{m} \left\| \mathbf{X}|\mathbf{r}\rangle - |\mathbf{z}\rangle \right\|^2 \\
&= \frac{1}{m} \left(\mathbf{X}|\mathbf{r}\rangle - |\mathbf{z}\rangle \right)^T \left(\mathbf{X}|\mathbf{r}\rangle - |\mathbf{z}\rangle \right) \\
&= \frac{1}{m} \left(\langle \mathbf{r}|\mathbf{X}^T - \langle \mathbf{z}| \right) \left(\mathbf{X}|\mathbf{r}\rangle - |\mathbf{z}\rangle \right) \\
&= \frac{1}{m} \left(\langle \mathbf{r}|\mathbf{X}^T \mathbf{X}|\mathbf{r}\rangle - \langle \mathbf{z}|\mathbf{X}|\mathbf{r}\rangle - \langle \mathbf{r}|\mathbf{X}^T|\mathbf{z}\rangle + \langle \mathbf{z}|\mathbf{z}\rangle \right) \\
&= \frac{1}{m} \left[\langle \mathbf{r}|\mathbf{X}^T \mathbf{X}|\mathbf{r}\rangle - \langle \mathbf{z}| \left(\langle \mathbf{r}|\mathbf{X}^T \right)^T - \langle \mathbf{r}|\mathbf{X}^T|\mathbf{z}\rangle + \langle \mathbf{z}|\mathbf{z}\rangle \right] \\
&= \frac{1}{m} \left\{ \langle \mathbf{r}|\mathbf{X}^T \mathbf{X}|\mathbf{r}\rangle - \left[\left(\langle \mathbf{r}|\mathbf{X}^T \right)^T \right]^T |\mathbf{z}\rangle \right. \\
&\qquad\qquad \left. - \langle \mathbf{r}|\mathbf{X}^T|\mathbf{z}\rangle + \langle \mathbf{z}|\mathbf{z}\rangle \right\} \\
&= \frac{1}{m} \left(\langle \mathbf{r}|\mathbf{X}^T \mathbf{X}|\mathbf{r}\rangle - 2\langle \mathbf{r}|\mathbf{X}^T|\mathbf{z}\rangle + \langle \mathbf{z}|\mathbf{z}\rangle \right).
\end{aligned}
\tag{2.131}
$$

In order to minimize the mean squared error (2.131) with respect to variations of $|\mathbf{r}\rangle$ one obtains a condition for "the linear theory" $|\mathbf{y}\rangle$ by setting its derivatives (its gradient) to zero; that is

$$
\partial_{|\mathbf{r}\rangle} \text{MSE} = \mathbf{0}.
\tag{2.132}
$$

A lengthy but straightforward computation yields

$$
\begin{aligned}
&\frac{\partial}{\partial r_i} \left(r_j \mathbf{X}_{jk}^T \mathbf{X}_{kl} r_l - 2 r_j \mathbf{X}_{jk}^T z_k + z_j z_j \right) \\
&= \delta_{ij} \mathbf{X}_{jk}^T \mathbf{X}_{kl} r_l + r_j \mathbf{X}_{jk}^T \mathbf{X}_{kl} \delta_{il} - 2\delta_{ij} \mathbf{X}_{jk}^T z_k \\
&= \mathbf{X}_{ik}^T \mathbf{X}_{kl} r_l + r_j \mathbf{X}_{jk}^T \mathbf{X}_{ki} - 2\mathbf{X}_{ik}^T z_k \\
&= \mathbf{X}_{ik}^T \mathbf{X}_{kl} r_l + \mathbf{X}_{ik}^T \mathbf{X}_{kj} r_j - 2\mathbf{X}_{ik}^T z_k \\
&= 2\mathbf{X}_{ik}^T \mathbf{X}_{kj} r_j - 2\mathbf{X}_{ik}^T z_k \\
&\equiv 2 \left(\mathbf{X}^T \mathbf{X}|\mathbf{r}\rangle - \mathbf{X}^T|\mathbf{z}\rangle \right) = 0
\end{aligned}
\tag{2.133}
$$

and finally, upon multiplication with $\left(\mathbf{X}^T \mathbf{X} \right)^{-1}$ from the left,

$$
|\mathbf{r}\rangle = \left(\mathbf{X}^T \mathbf{X} \right)^{-1} \mathbf{X}^T |\mathbf{z}\rangle.
\tag{2.134}
$$

A short plausibility check for $n = m = 1$ yields the linear dependency $|\mathbf{z}\rangle = \mathbf{X}|\mathbf{r}\rangle$.

2.14   Some common misconceptions

2.14.1   Confusion between component representation and "the real thing"

Given a particular basis, a tensor is uniquely characterized by its components. However, without reference to a particular basis, any components

are just blurbs.

Example (wrong!): a type-1 tensor (i.e., a vector) is given by $(1, 2)$.

Correct: with respect to the basis $\{(0, 1), (1, 0)\}$, a rank-1 tensor (i.e., a vector) is given by $(1, 2)$.

2.14.2   A matrix is a tensor

See the above section. Example (wrong!): A matrix is a tensor of type (or rank) 2. Correct: with respect to the basis $\{(0, 1), (1, 0)\}$, a matrix represents a type-2 tensor. The matrix components are the tensor components.

Also, for non-orthogonal bases, covariant, contravariant, and mixed tensors correspond to different matrices.

3
Projective and incidence geometry

PROJECTIVE GEOMETRY is about the *geometric properties* that are invariant under *projective transformations. Incidence geometry* is about which points lie on which line.

3.1 Notation

In what follows, for the sake of being able to formally represent *geometric transformations as "quasi-linear" transformations and matrices,* the coordinates of n-dimensional Euclidean space will be augmented with one additional coordinate which is set to one. For instance, in the plane $\mathbb{R}^2$, we define new "three-componen" coordinates by

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \equiv \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} = \mathbf{X}. \tag{3.1}$$

In order to differentiate these new coordinates $\mathbf{X}$ from the usual ones $\mathbf{x}$, they will be written in capital letters.

3.2 Affine transformations

Affine transformations

$$f(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{t} \tag{3.2}$$

with the *translation* $\mathbf{t}$, and encoded by a touple $(t_1, t_2)^T$, and an arbitrary linear transformation $\mathbf{A}$ encoding *rotations,* as well as *dilatation and skewing transformations* and represented by an arbitrary matrix A, can be "wrapped together" to form the new transformation matrix ("$\mathbf{0}^T$" indicates a row matrix with entries zero)

$$\mathbf{f} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{t} \\ \hline \mathbf{0}^T & 1 \end{array} \right) \equiv \left(\begin{array}{cc|c} a_{11} & a_{12} & t_1 \\ a_{21} & a_{22} & t_2 \\ \hline 0 & 0 & 1 \end{array} \right). \tag{3.3}$$

As a result, the affine transformation f can be represented in the "quasi-linear" form

$$\mathbf{f}(\mathbf{X}) = \mathbf{f}\mathbf{X} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{t} \\ \hline \mathbf{0}^T & 1 \end{array} \right) \mathbf{X}. \tag{3.4}$$

3.2.1   One-dimensional case

In one dimension, that is, for $\mathbf{z} \in \mathbb{C}$, among the five basic operatios

(i)   scaling: $\mathbf{f}(\mathbf{z}) = r\mathbf{z}$ for $r \in \mathbb{R}$,

(ii)   translation: $\mathbf{f}(\mathbf{z}) = \mathbf{z} + \mathbf{w}$ for $w \in \mathbb{C}$,

(iii)   rotation: $\mathbf{f}(\mathbf{z}) = e^{i\varphi}\mathbf{z}$ for $\varphi \in \mathbb{R}$,

(iv)   complex conjugation: $\mathbf{f}(\mathbf{z}) = \overline{\mathbf{z}}$,

(v)   inversion: $\mathbf{f}(\mathbf{z}) = \mathbf{z}^{-1}$,

there are three types of affine transformations (i)–(iii) which can be combined.

3.3   Similarity transformations

Similarity transformations involve translations $\mathbf{t}$, rotations $\mathbf{R}$ and a dilatation r and can be represented by the matrix

$$\left(\begin{array}{c|c} r\mathbf{R} & \mathbf{t} \\ \hline \mathbf{0}^T & 1 \end{array} \right) \equiv \left(\begin{array}{cc|c} m\cos\varphi & -m\sin\varphi & t_1 \\ m\sin\varphi & m\cos\varphi & t_2 \\ \hline 0 & 0 & 1 \end{array} \right). \tag{3.5}$$

3.4   Fundamental theorem of affine geometry

Any bijection from $\mathbb{R}^n$, $n \geq 2$, onto itself which maps all lines onto lines is an affine transformation.

For a proof and further references, see
June A. Lester. Distance preserving transformations. In Francis Buekenhout, editor, *Handbook of Incidence Geometry*, pages 921–944. Elsevier, Amsterdam, 1995

3.5   Alexandrov's theorem

Consider the Minkowski space-time $\mathbb{M}^n$; that is, $\mathbb{R}^n$, $n \geq 3$, and the Minkowski metric [cf. (2.65) on page 87] $\eta \equiv \{\eta_{ij}\} = \mathrm{diag}(\underbrace{1, 1, \ldots, 1}_{n-1 \text{ times}}, -1)$.
Consider further bijections $\mathbf{f}$ from $\mathbb{M}^n$ onto itself preserving light cones; that is for all $\mathbf{x}, \mathbf{y} \in \mathbb{M}^n$,

$$\eta_{ij}(x^i - y^i)(x^j - y^j) = 0 \text{ if and only if } \eta_{ij}(\mathbf{f}^i(x) - \mathbf{f}^i(y))(\mathbf{f}^j(x) - \mathbf{f}^j(y)) = 0.$$

Then $\mathbf{f}(x)$ is the product of a Lorentz transformation and a positive scale factor.

For a proof and further references, see
June A. Lester. Distance preserving transformations. In Francis Buekenhout, editor, *Handbook of Incidence Geometry*, pages 921–944. Elsevier, Amsterdam, 1995

4

Group theory

GROUP THEORY is about *transformations* and *symmetries.*

4.1 Definition

A *group* is a set of objects $\mathfrak{G}$ which satify the following conditions (or, stated differently, axioms):

(i) closedness: There exists a *composition rule* "$\circ$" such that $\mathfrak{G}$ is *closed* under any composition of elements; that is, the combination of any two elements $a, b \in \mathfrak{G}$ results in an element of the group $\mathfrak{G}$.

(ii) associativity: for all a, b, and c in $\mathfrak{G}$, the following equality holds: $a \circ (b \circ c) = (a \circ b) \circ c$;

(iii) identity (element): there exists an element of $\mathfrak{G}$, called the *identity* (element) and denoted by I, such that for all a in $\mathfrak{G}$, $a \circ I = a$.

(iv) inverse (element): for every a in $\mathfrak{G}$, there exists an element a^{-1} in $\mathfrak{G}$, such that $a^{-1} \circ a = I$.

(v) (optional) commutativity: if, for all a and b in $\mathfrak{G}$, the following equalities hold: $a \circ b = b \circ a$, then the group $\mathfrak{G}$ is called *Abelian (group)*; otherwise it is called *non-Abelian (group).*

A *subgroup* of a group is a subset which also satisfies the above axioms.

The *order* of a group is the number of distinct elements of that group.

In discussing groups one should keep in mind that there are two abstract spaces involved:

(i) *Representation space* is the space of elements on which the group elements – that is, the group transformations – act.

(ii) *Group space* is the space of elements of the group transformations. Its *dimension* is the *number of independent transformations* which the group is composed of. These independent elements – also called the *generators* of the group – form a basis for all group elements. The

coordinates in this space are defined relative (in terms of) the (basis elements, also called) generators. A *continuous group* can geometrically be imagined as a linear space (e.g., a linear vector or matrix space) *continuous group linear space* in which every point in this linear space is an element of the group.

Suppose we can find a structure- and disticion-preserving mapping U – that is, an injective mapping preserving the group operation ∘ – between elements of a group $\mathfrak{G}$ and the groups of general either real or complex non-singular matrices $\mathrm{GL}(n, \mathbb{R})$ or $\mathrm{GL}(n, \mathbb{C})$, respectively. Then this mapping is called a *representation* of the group $\mathfrak{G}$. In particular, for this $U : \mathfrak{G} \mapsto \mathrm{GL}(n, \mathbb{R})$ or $U : \mathfrak{G} \mapsto \mathrm{GL}(n, \mathbb{C})$,

$$U(a \circ b) = U(a) \cdot U(b), \tag{4.1}$$

for all $a, b, a \circ b \in \mathfrak{G}$.

Consider, for the sake of an example, the *Pauli spin matrices* which are proportional to the angular momentum operators along the x, y, z-axis [1]:

[1] Leonard I. Schiff. *Quantum Mechanics*. McGraw-Hill, New York, 1955

$$\sigma_1 = \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

$$\sigma_2 = \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{4.2}$$

$$\sigma_3 = \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Suppose these matrices $\sigma_1, \sigma_2, \sigma_3$ serve as generators of a group. With respect to this basis system of matrices $\{\sigma_1, \sigma_2, \sigma_3\}$ a general point in group in group space might be labelled by a three-dimensional vector with the coordinates (x_1, x_2, x_3) (relative to the basis $\{\sigma_1, \sigma_2, \sigma_3\}$); that is,

$$\mathbf{x} = x_1 \sigma_1 + x_2 \sigma_2 + x_3 \sigma_3. \tag{4.3}$$

If we form the exponential $A(\mathbf{x}) = e^{\frac{i}{2}\mathbf{x}}$, we can show (no proof is given here) that $A(\mathbf{x})$ is a two-dimensional matrix representation of the group SU(2), the special unitary group of degree 2 of 2×2 unitary matrices with determinant 1.

4.2   Lie theory

4.2.1   Generators

We can generalize this examply by defining the *generators* of a continuous group as the first coefficient of a Taylor expansion around unity; that is, if the dimension of the group is n, and the Taylor expansion is

$$G(\mathbf{X}) = \sum_{i=1}^{n} X_i T_i + \dots, \tag{4.4}$$

then the matrix generator T_i is defined by

$$T_i = \left. \frac{\partial G(\mathbf{X})}{\partial X_i} \right|_{\mathbf{X}=0}. \tag{4.5}$$

4.2.2   Exponential map

There is an exponential connection $\exp : \mathfrak{X} \mapsto \mathfrak{G}$ between a matrix Lie group and the Lie algebra $\mathfrak{X}$ generated by the generators T_i.

4.2.3   Lie algebra

A Lie algebra is a vector space $\mathfrak{X}$, together with a binary *Lie bracket* operation $[\cdot,\cdot] : \mathfrak{X} \times \mathfrak{X} \mapsto \mathfrak{X}$ satisfying

(i)   bilinearity;

(ii)   antisymmetry: $[X, Y] = -[Y, X]$, in particular $[X, X] = 0$;

(iii)   the Jacobi identity: $[X, [Y, Z]] + [Z, [X, Y]] + [Y, [Z, X]] = 0$

for all $X, Y, Z \in \mathfrak{X}$.

4.3   Some important groups

4.3.1   General linear group $GL(n, \mathbb{C})$

The *general linear group* $GL(n, \mathbb{C})$ contains all non-singular (i.e., invertible; there exist an inverse) $n \times n$ matrices with complex entries. The composition rule "∘" is identified with matrix multiplication (which is associative); the neutral element is the unit matrix $\mathbb{I}_n = \mathrm{diag}(\underbrace{1,\dots,1}_{n \text{ times}})$.

4.3.2   Orthogonal group $O(n)$

The *orthogonal group* $O(n)$ [2] contains all orthogonal [i.e., $A^{-1} = A^T$] $n \times n$ matrices. The composition rule "∘" is identified with matrix multiplication (which is associative); the neutral element is the unit matrix $\mathbb{I}_n = \mathrm{diag}(\underbrace{1,\dots,1}_{n \text{ times}})$.

Because of orthogonality, only half of the off-diagonal entries are independent of one another; also the diagonal elements must be real; that leaves us with the liberty of dimension $n(n+1)/2$: $(n^2 - n)/2$ complex numbers from the off-diagonal elements, plus n reals from the diagonal.

[2] Francis D. Murnaghan. *The Unitary and Rotation Groups*, volume 3 of *Lectures on Applied Mathematics*. Spartan Books, Washington, D.C., 1962

4.3.3   Rotation group $SO(n)$

The *special orthogonal group* or, by another name, the *rotation group* $SO(n)$ contains all orthogonal $n \times n$ matrices with unit determinant. $SO(n)$ is a subgroup of $O(n)$

The rotation group in two-dimensional configuration space $SO(2)$ corresponds to planar rotations around the origin. It has dimension 1 corresponding to one parameter θ. Its elements can be written as

$$R(\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}. \tag{4.6}$$

4.3.4   Unitary group U(n)

The *unitary group* U(n) [3] contains all unitary [i.e., $A^{-1} = A^{\dagger} = (\overline{A})^{T}$] $n \times n$ matrices. The composition rule "∘" is identified with matrix multiplication (which is associative); the neutral element is the unit matrix $\mathbb{I}_n = \mathrm{diag}(\underbrace{1, \ldots, 1}_{n \text{ times}})$.

[3] Francis D. Murnaghan. *The Unitary and Rotation Groups*, volume 3 of *Lectures on Applied Mathematics*. Spartan Books, Washington, D.C., 1962

Because of unitarity, only half of the off-diagonal entries are independent of one another; also the diagonal elements must be real; that leaves us with the liberty of dimension n^2: $(n^2 - n)/2$ complex numbers from the off-diagonal elements, plus n reals from the diagonal yield n^2 real parameters.

Not that, for instance, U(1) is the set of complex numbers $z = e^{i\theta}$ of unit modulus $|z|^2 = 1$. It forms an Abelian group.

4.3.5   Special unitary group SU(n)

The *special unitary group* SU(n) contains all unitary $n \times n$ matrices with unit determinant. SU(n) is a subgroup of U(n).

4.3.6   Symmetric group S(n)

The *symmetric group* S(n) on a finite set of n elements (or symbols) is the group whose elements are all the permutations of the n elements, and whose group operation is the composition of such permutations. The identity is the identity permutation. The *permutations* are bijective functions from the set of elements onto itself. The order (number of elements) of S(n) is $n!$. Generalizing these groups to an infinite number of elements S_∞ is straightforward.

The symmetric group should not be confused with a symmetry group.

4.3.7   Poincaré group

The Poincaré group is the group of *isometries* – that is, bijective maps preserving distances – in space-time modelled by $\mathbb{R}^4$ endowed with a scalar product and thus of a norm induced by the *Minkowski metric* $\eta \equiv \{\eta_{ij}\} = \mathrm{diag}(1, 1, 1, -1)$ introduced in (2.65).

It has dimension ten $(4 + 3 + 3 = 10)$, associated with the ten fundamental (distance preserving) operations from which general isometries can be composed: (i) translation through time and any of the three dimensions of space $(1 + 3 = 4)$, (ii) rotation (by a fixed angle) around any of the three spatial axes (3), and a (Lorentz) boost, increasing the velocity in any of the three spatial directions of two uniformly moving bodies (3).

The rotations and Lorentz boosts form the *Lorentz group.*

4.4   Cayley's representation theorem

Cayley's theorem states that every group $\mathfrak{G}$ can be imbedded as – equivalently, is isomorphic to – a subgroup of the symmetric group; that is, it is a isomorphic with some permutation group. In particular, every finite group $\mathfrak{G}$ of order n can be imbedded as – equivalently, is isomorphic to – a subgroup of the symmetric group S(n).

Stated pointedly: permutations exhaust the possible structures of (finite) groups. The study of subgroups of the symmetric groups is no less general than the study of all groups. No proof is given here.

✦

For a proof, see

Joseph J. Rotman. *An Introduction to the Theory of Groups*, volume 148 of *Graduate texts in mathematics*. Springer, New York, fourth edition, 1995. ISBN 0387942858

Part II
Functional analysis

5

Brief review of complex analysis

Is it not amazing that complex numbers [1] can be used for physics? Robert Musil (a mathematician educated in Vienna), in *"Verwirrungen des Zögling Törleß"*, expressed the amazement of a youngster confronted with the applicability of imaginaries, states that, at the beginning of any computation involving imaginary numbers are "solid" numbers which could represent something measurable, like lengths or weights, or something else tangible; or are at least real numbers. At the end of the computation there are also such "solid" numbers. But the beginning and the end of the computation are connected by something seemingly nonexisting. Does this not appear, Musil's *Zögling Törleß* wonders, like a bridge crossing an abyss with only a bridge pier at the very beginning and one at the very end, which could nevertheless be crossed with certainty and securely, as if this bridge would exist entirely?

In what follows, a very brief review of *complex analysis*, or, by another term, *function theory*, will be presented. For much more detailed introductions to complex analysis, including proofs, take, for instance, the "classical" books [2], among a zillion of other very good ones [3]. We shall study complex analysis not only for its beauty, but also because it yields very important analytical methods and tools; for instance for the solution of (differential) equations and the computation of definite integrals. These methods will then be required for the computation of distributions and Green's functions, as well for the solution of differential equations of mathematical physics – such as the Schrödinger equation.

One motivation for introducing *imaginary numbers* is the (if you perceive it that way) "malady" that not every polynomial such as $P(x) = x^2 + 1$ has a root x – and thus not every (polynomial) equation $P(x) = x^2 + 1 = 0$ has a solution x – which is a real number. Indeed, you need the imaginary unit $i^2 = -1$ for a factorization $P(x) = (x + i)(x - i)$ yielding the two roots $\pm i$ to achieve this. In that way, the introduction of imaginary numbers is a further step towards omni-solvability. No wonder that the fundamental theorem of algebra, stating that every non-constant polynomial with complex coefficients has at least one complex root – and thus total factorizability of polynomials into linear factors, follows!

If not mentioned otherwise, it is assumed that the *Riemann surface*, representing a "deformed version" of the complex plane for functional

[1] Edmund Hlawka. Zum Zahlbegriff. *Philosophia Naturalis*, 19:413–470, 1982

German original (http://www.gutenberg.org/ebooks/34717): *"In solch einer Rechnung sind am Anfang ganz solide Zahlen, die Meter oder Gewichte, oder irgend etwas anderes Greifbares darstellen können und wenigstens wirkliche Zahlen sind. Am Ende der Rechnung stehen ebensolche. Aber diese beiden hängen miteinander durch etwas zusammen, das es gar nicht gibt. Ist das nicht wie eine Brücke, von der nur Anfangs- und Endpfeiler vorhanden sind und die man dennoch so sicher überschreitet, als ob sie ganz dastünde? Für mich hat so eine Rechnung etwas Schwindliges; als ob es ein Stück des Weges weiß Gott wohin ginge. Das eigentlich Unheimliche ist mir aber die Kraft, die in solch einer Rechnung steckt und einen so festhalt, daß man doch wieder richtig landet."*

[2] Eberhard Freitag and Rolf Busam. *Funktionentheorie 1*. Springer, Berlin, Heidelberg, fourth edition, 1993,1995,2000,2006. English translation in ; E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge University Press, Cambridge, fourth edition, 1927. URL http://archive.org/details/ACourseOfModernAnalysis. Reprinted in 1996. Table errata: Math. Comp. v. 36 (1981), no. 153, p. 319; Robert E. Greene and Stephen G. Krantz. *Function theory of one complex variable*, volume 40 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, third edition, 2006; Einar Hille. *Analytic Function Theory*. Ginn, New York, 1962. 2 Volumes; and Lars V. Ahlfors. *Complex Analysis: An Introduction of the Theory of Analytic Functions of One Complex Variable*. McGraw-Hill Book Co., New York, third edition, 1978

Eberhard Freitag and Rolf Busam. *Complex Analysis*. Springer, Berlin, Heidelberg, 2005

[3] Klaus Jänich. *Funktionentheorie. Eine Einführung*. Springer, Berlin, Heidelberg, sixth edition, 2008. DOI: 10.1007/978-3-540-35015-6. URL 10.1007/978-3-540-35015-6; and Dietmar A. Salamon. *Funktionentheorie*. Birkhäuser, Basel,

purposes, is simply connected. Simple connectedness means that the Riemann surface it is path-connected so that every path between two points can be continuously transformed, staying within the domain, into any other path while preserving the two endpoints between the paths. In particular, suppose that there are no "holes" in the Riemann surface; it is not "punctured."

Furthermore, let i be the *imaginary unit* with the property that $i^2 = -1$ is the solution of the equation $x^2 + 1 = 0$. With the introduction of imaginary numbers we can guarantee that all quadratic equations have two roots (i.e., solutions).

By combining imaginary and real numbers, any *complex number* can be defined to be some linear combination of the real unit number "1" with the imaginary unit number i that is, $z = 1 \times (\Re z) + i \times (\Im z)$, with the real valued factors $(\Re z)$ and $(\Im z)$, respectively. By this definition, a complex number z can be decomposed into real numbers x, y, r and φ such that

$$z \overset{\text{def}}{=} \Re z + i \Im z = x + i y = r e^{i\varphi}, \tag{5.1}$$

with $x = r\cos\varphi$ and $y = r\sin\varphi$, where Euler's formula

$$e^{i\varphi} = \cos\varphi + i\sin\varphi \tag{5.2}$$

has been used. If $z = \Re z$ we call z a real number. If $z = i\Im z$ we call z a purely imaginary number.

The *modulus* or *absolute value* of a complex number z is defined by

$$|z| \overset{\text{def}}{=} +\sqrt{(\Re z)^2 + (\Im z)^2}. \tag{5.3}$$

Many rules of classical arithmetic can be carried over to complex arithmetic [4]. Note, however, that, for instance, $\sqrt{a}\sqrt{b} = \sqrt{ab}$ is only valid if at least one factor a or b is positive; hence $-1 = i^2 = \sqrt{i}\sqrt{i} = \sqrt{-1}\sqrt{-1} \neq \sqrt{(-1)^2} = 1$. More generally, for two arbitrary numbers, u and v, $\sqrt{u}\sqrt{v}$ is not always equal to $\sqrt{uv}$.

For many mathematicians *Euler's identity*

$$e^{i\pi} = -1, \text{ or } e^{i\pi} + 1 = 0, \tag{5.4}$$

is the "most beautiful" theorem [5].

Euler's formula (5.2) can be used to derive *de Moivre's formula* for integer n (for non-integer n the formula is multi-valued for different arguments φ):

$$e^{in\varphi} = (\cos\varphi + i\sin\varphi)^n = \cos(n\varphi) + i\sin(n\varphi). \tag{5.5}$$

It is quite suggestive to consider the complex numbers z, which are linear combinations of the real and the imaginary unit, in the *complex plane* $\mathbb{C} = \mathbb{R} \times \mathbb{R}$ as a geometric representation of complex numbers. Thereby, the real and the imaginary unit are identified with the (orthonormal) basis vectors of the *standard (Cartesian) basis*; that is, with the tuples

$$\begin{aligned} 1 &\equiv (1,0), \\ i &\equiv (0,1). \end{aligned} \tag{5.6}$$

The addition and multiplication of two complex numbers represented by (x, y) and (u, v) with $x, y, u, v \in \mathbb{R}$ are then defined by

$$(x, y) + (u, v) = (x + u, y + v),$$
$$(x, y) \cdot (u, v) = (xu - yv, xv + yu),$$

(5.7)

and the neutral elements for addition and multiplication are $(0, 0)$ and $(1, 0)$, respectively.

We shall also consider the *extended plane* $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ consisting of the entire complex plane $\mathbb{C}$ *together* with the point "∞" representing infinity. Thereby, ∞ is introduced as an ideal element, completing the one-to-one (bijective) mapping $w = \frac{1}{z}$, which otherwise would have no image at $z = 0$, and no pre-image (argument) at $w = 0$.

5.1   Differentiable, holomorphic (analytic) function

Consider the function $f(z)$ on the domain $G \subset \text{Domain}(f)$.

f is called *differentiable* at the point z_0 if the differential quotient

$$\left. \frac{df}{dz} \right|_{z_0} = \left. f'(z) \right|_{z_0} = \left. \frac{\partial f}{\partial x} \right|_{z_0} = \left. \frac{1}{i} \frac{\partial f}{\partial y} \right|_{z_0}$$

(5.8)

exists.

If f is (arbitrarily often) differentiable in the domain G it is called *holomorphic*. We shall state without proof that, if a holomorphic function is differentiable, it is also arbitrarily often differentiable.

If f it can be expanded as a convergent power series in the domain G it is called *analytic* in the domain G. We state without proof that holomorphic functions are analytic, and *vice versa*; that is the terms holomorphic and analytic will be used synonymuously.

5.2   Cauchy-Riemann equations

The function $f(z) = u(z) + iv(z)$ (where u and v are real valued functions) is analytic or holomorph if and only if ($a_b = \partial a / \partial b$)

$$u_x = v_y, \qquad u_y = -v_x \quad .$$

(5.9)

For a proof, differentiate along the real, and then along the complex axis, taking

$$f'(z) = \lim_{x \to 0} \frac{f(z + x) - f(z)}{x} = \frac{\partial f}{\partial x} = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x},$$
$$\text{and } f'(z) = \lim_{y \to 0} \frac{f(z + iy) - f(z)}{iy} = \frac{\partial f}{\partial iy} = -i\frac{\partial f}{\partial y} = -i\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}.$$

(5.10)

For f to be analytic, both partial derivatives have to be identical, and thus $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial iy}$, or

$$\frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = -i\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}.$$

(5.11)

By comparing the real and imaginary parts of this equation, one obtains the two real Cauchy-Riemann equations

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y},$$
$$\frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}.$$

(5.12)

5.3   Definition analytical function

If f is analytic in G, all derivatives of f exist, and all mixed derivatives are independent on the order of differentiations. Then the Cauchy-Riemann equations imply that

$$
\frac{\partial}{\partial x}\left(\frac{\partial u}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial v}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial v}{\partial x}\right) = -\frac{\partial}{\partial y}\left(\frac{\partial u}{\partial y}\right),
$$
$$
\text{and } \frac{\partial}{\partial y}\left(\frac{\partial v}{\partial y}\right) = \frac{\partial}{\partial y}\left(\frac{\partial u}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{\partial u}{\partial y}\right) = -\frac{\partial}{\partial x}\left(\frac{\partial v}{\partial x}\right),
$$

(5.13)

and thus

$$
\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) u = 0, \text{ and } \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) v = 0 \quad .
$$

(5.14)

If $f = u + iv$ is analytic in G, then the lines of constant u and v are orthogonal.

The tangential vectors of the lines of constant u and v in the two-dimensional complex plane are defined by the two-dimensional nabla operator $\nabla u(x, y)$ and $\nabla v(x, y)$. Since, by the Cauchy-Riemann equations $u_x = v_y$ and $u_y = -v_x$

$$
\nabla u(x, y) \cdot \nabla v(x, y) = \begin{pmatrix} u_x \\ u_y \end{pmatrix} \cdot \begin{pmatrix} v_x \\ v_y \end{pmatrix} = u_x v_x + u_y v_y = u_x v_x + (-v_x) u_x = 0
$$

(5.15)

these tangential vectors are normal.

f is *angle (shape) preserving conformal* if and only if it is holomorphic and its derivative is everywhere non-zero.

Consider an analytic function f and an arbitrary path C in the complex plane of the arguments parameterized by $z(t)$, $t \in \mathbb{R}$. The image of C associated with f is $f(C) = C' : f(z(t))$, $t \in \mathbb{R}$.

The tangent vector of C' in $t = 0$ and $z_0 = z(0)$ is

$$
\frac{d}{dt} f(z(t))\Big|_{t=0} = \frac{d}{dz} f(z)\Big|_{z_0} \frac{d}{dt} z(t)\Big|_{t=0} = \lambda_0 e^{i\varphi_0} \frac{d}{dt} z(t)\Big|_{t=0} . \quad (5.16)
$$

Note that the first term $\frac{d}{dz} f(z)\Big|_{z_0}$ is independent of the curve C and only depends on z_0. Therefore, it can be written as a product of a squeeze (stretch) λ_0 and a rotation $e^{i\varphi_0}$. This is independent of the curve; hence two curves C_1 and C_2 passing through z_0 yield the same transformation of the image $\lambda_0 e^{i\varphi_0}$.

5.4   Cauchy's integral theorem

If f is analytic on G and on its borders ∂G, then any closed line integral of f vanishes

$$
\oint_{\partial G} f(z) dz = 0 \quad .
$$

(5.17)

No proof is given here.

In particular, $\oint_{C \subset \partial G} f(z) dz$ is independent of the particular curve, and only depends on the initial and the end points.

For a proof, subtract two line integral which follow arbitrary paths C_1 and C_2 to a common initial and end point, and which have the same integral kernel. Then reverse the integration direction of one of the line integrals. According to Cauchy's integral theorem the resulting integral over the closed loop has to vanish.

Often it is useful to parameterize a contour integral by some form of

$$\int_C f(z)\,dz = \int_a^b f(z(t)) \frac{dz(t)}{dt}\,dt. \tag{5.18}$$

Let $f(z) = 1/z$ and $C : z(\varphi) = Re^{i\varphi}$, with $R > 0$ and $-\pi < \varphi \leq \pi$. Then

$$
\begin{aligned}
\oint_{|z|=R} f(z)\,dz \quad &= \int_{-\pi}^{\pi} f(z(\varphi)) \frac{dz(\varphi)}{d\varphi}\,d\varphi \\
&= \int_{-\pi}^{\pi} \frac{1}{Re^{i\varphi}} R\,i\,e^{i\varphi}\,d\varphi \\
&= \int_{-\pi}^{\pi} i\varphi \\
&= 2\pi i
\end{aligned}
\tag{5.19}
$$

is independent of R.

5.5   Cauchy's integral formula

If f is analytic on G and on its borders ∂G, then

$$f(z_0) = \frac{1}{2\pi i} \oint_{\partial G} \frac{f(z)}{z - z_0}\,dz \quad . \tag{5.20}$$

No proof is given here.

Note that because of Cauchy's integral formula, analytic functions have an integral representation. This might appear as not very exciting; alas it has far-reaching consequences, because analytic functions have integral representation, they have higher derivatives, which also have integral representation. And, as a result, if a function has one complex derivative, then it has infnitely many complex derivatives. This statement can be expressed formally precisely by the generalized Cauchy's integral formula or, by another term, Cauchy's differentiation formula states that if f is analytic on G and on its borders ∂G, then

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_{\partial G} \frac{f(z)}{(z - z_0)^{n+1}}\,dz \quad . \tag{5.21}$$

No proof is given here.

Cauchy's integral formula presents a powerful method to compute integrals. Consider the following examples.

(i)   First, let us calculate

$$\oint_{|z|=3} \frac{3z+2}{z(z+1)^3}\,dz.$$

The kernel has two poles at $z = 0$ and $z = -1$ which are both inside the domain of the contour defined by $|z| = 3$. By using Cauchy's integral

formula we obtain for "small" ϵ

$$\oint_{|z|=3} \frac{3z+2}{z(z+1)^3} dz$$

$$= \oint_{|z|=\epsilon} \frac{3z+2}{z(z+1)^3} dz + \oint_{|z+1|=\epsilon} \frac{3z+2}{z(z+1)^3} dz$$

$$= \oint_{|z|=\epsilon} \frac{3z+2}{(z+1)^3} \frac{1}{z} dz + \oint_{|z+1|=\epsilon} \frac{3z+2}{z} \frac{1}{(z+1)^3} dz$$

$$= \frac{2\pi i}{0!} [[\frac{d^0}{dz^0}]] \frac{3z+2}{(z+1)^3} \Big|_{z=0} + \frac{2\pi i}{2!} \frac{d^2}{dz^2} \frac{3z+2}{z} \Big|_{z=-1} \qquad (5.22)$$

$$= \frac{2\pi i}{0!} \frac{3z+2}{(z+1)^3} \Big|_{z=0} + \frac{2\pi i}{2!} \frac{d^2}{dz^2} \frac{3z+2}{z} \Big|_{z=-1}$$

$$= 4\pi i - 4\pi i$$

$$= 0.$$

(ii) Consider

$$\oint_{|z|=3} \frac{e^{2z}}{(z+1)^4} dz$$

$$= \frac{2\pi i}{3!} \frac{3!}{2\pi i} \oint_{|z|=3} \frac{e^{2z}}{(z-(-1))^{3+1}} dz$$

$$= \frac{2\pi i}{3!} \frac{d^3}{dz^3} |e^{2z}|_{z=-1} \qquad (5.23)$$

$$= \frac{2\pi i}{3!} 2^3 |e^{2z}|_{z=-1}$$

$$= \frac{8\pi i e^{-2}}{3}.$$

Suppose $g(z)$ is a function with a pole of order n at the point z_0; that is

$$g(z) = \frac{f(z)}{(z-z_0)^n} \qquad (5.24)$$

where $f(z)$ is an analytic function. Then,

$$\oint_{\partial G} g(z) dz = \frac{2\pi i}{(n-1)!} f^{(n-1)}(z_0) \quad . \qquad (5.25)$$

5.6   Series representation of complex differentiable functions

As a consequence of Cauchy's (generalized) integral formula, analytic functions have power series representations.

For the sake of a proof, we shall recast the denominator $z - z_0$ in Cauchy's integral formula (5.20) as a geometric series as follows (we shall assume that $|z_0 - a| < |z - a|$)

$$\frac{1}{z-z_0} = \frac{1}{(z-a)-(z_0-a)}$$

$$= \frac{1}{(z-a)} \left[\frac{1}{1-\frac{z_0-a}{z-a}} \right]$$

$$= \frac{1}{(z-a)} \left[\sum_{n=0}^{\infty} \frac{(z_0-a)^n}{(z-a)^n} \right] \qquad (5.26)$$

$$= \sum_{n=0}^{\infty} \frac{(z_0-a)^n}{(z-a)^{n+1}}.$$

By substituting this in Cauchy's integral formula (5.20) and using
Cauchy's generalized integral formula (5.21) yields an expansion of the
analytical function f around z_0 by a power series

$$
\begin{aligned}
f(z_0) &= \frac{1}{2\pi i} \oint_{\partial G} \frac{f(z)}{z - z_0} dz \\
&= \frac{1}{2\pi i} \oint_{\partial G} f(z) \sum_{n=0}^{\infty} \frac{(z_0 - a)^n}{(z - a)^{n+1}} dz \\
&= \sum_{n=0}^{\infty} (z_0 - a)^n \frac{1}{2\pi i} \oint_{\partial G} \frac{f(z)}{(z - a)^{n+1}} dz \\
&= \sum_{n=0}^{\infty} \frac{f^n(z_0)}{n!} (z_0 - a)^n.
\end{aligned}
\tag{5.27}
$$

5.7   Laurent series

Every function f which is analytic in a concentric region $R_1 < |z - z_0| < R_2$
can in this region be uniquely written as a *Laurent series*

$$
f(z) = \sum_{k=-\infty}^{\infty} (z - z_0)^k a_k
\tag{5.28}
$$

The coefficients a_k are (the closed contour C must be in the concentric
region)

$$
a_k = \frac{1}{2\pi i} \oint_C (\chi - z_0)^{-k-1} f(\chi) d\chi \quad .
\tag{5.29}
$$

The coefficient

$$
\text{Res}(f(z_0)) = a_{-1} = \frac{1}{2\pi i} \oint_C f(\chi) d\chi
\tag{5.30}
$$

is called the *residue,* denoted by "Res."

For a proof, as in Eqs. (5.26) we shall recast $(a - b)^{-1}$ for $|a| > |b|$ as a
geometric series

$$
\begin{aligned}
\frac{1}{a - b} &= \frac{1}{a} \left(\frac{1}{1 - \frac{b}{a}} \right) == \frac{1}{a} \left(\sum_{n=0}^{\infty} \frac{b^n}{a^n} \right) = \sum_{n=0}^{\infty} \frac{b^n}{a^{n+1}} \\
[\text{substitution } n + 1 \to -k, \, n \to -k - 1 \, k \to -n - 1] &= \sum_{k=-1}^{-\infty} \frac{a^k}{b^{k+1}},
\end{aligned}
\tag{5.31}
$$

and, for $|a| < |b|$,

$$
\begin{aligned}
\frac{1}{a - b} &= -\frac{1}{b - a} = -\sum_{n=0}^{\infty} \frac{a^n}{b^{n+1}} \\
[\text{substitution } n + 1 \to -k, \, n \to -k - 1 \, k \to -n - 1] &= -\sum_{k=-1}^{-\infty} \frac{b^k}{a^{k+1}}.
\end{aligned}
\tag{5.32}
$$

Furthermore since $a + b = a - (-b)$, we obtain, for $|a| > |b|$,

$$
\frac{1}{a + b} = \sum_{n=0}^{\infty} (-1)^n \frac{b^n}{a^{n+1}} = \sum_{k=-1}^{-\infty} (-1)^{-k-1} \frac{a^k}{b^{k+1}} = -\sum_{k=-1}^{-\infty} (-1)^k \frac{a^k}{b^{k+1}},
\tag{5.33}
$$

and, for $|a| < |b|$,

$$
\begin{aligned}
\frac{1}{a + b} &= -\sum_{n=0}^{\infty} (-1)^{n+1} \frac{a^n}{b^{n+1}} = \sum_{n=0}^{\infty} (-1)^n \frac{a^n}{b^{n+1}} \\
&= \sum_{k=-1}^{-\infty} (-1)^{-k-1} \frac{b^k}{a^{k+1}} = -\sum_{k=-1}^{-\infty} (-1)^k \frac{b^k}{a^{k+1}}.
\end{aligned}
\tag{5.34}
$$

Suppose that some function $f(z)$ is analytic in an annulus bounded by the radius r_1 and $r_2 > r_1$. By substituting this in Cauchy's integral formula (5.20) for an annulus bounded by the radius r_1 and $r_2 > r_1$ (note that the orientations of the boundaries with respect to the annulus are opposite, rendering a relative factor "-1") and using Cauchy's generalized integral formula (5.21) yields an expansion of the analytical function f around z_0 by the Laurent series for a point a on the annulus; that is, for a path containing the point z around a circle with radius r_1, $|z - a| < |z_0 - a|$; likewise, for a path containing the point z around a circle with radius $r_2 > a > r_1$, $|z - a| > |z_0 - a|$,

$$f(z_0) = \frac{1}{2\pi i} \oint_{r_1} \frac{f(z)}{z - z_0} dz - \frac{1}{2\pi i} \oint_{r_2} \frac{f(z)}{z - z_0} dz$$

$$= \frac{1}{2\pi i} \left[\oint_{r_1} f(z) \sum_{n=0}^{\infty} \frac{(z_0 - a)^n}{(z - a)^{n+1}} dz + \oint_{r_2} f(z) \sum_{n=-1}^{-\infty} \frac{(z_0 - a)^n}{(z - a)^{n+1}} dz \right]$$

$$= \frac{1}{2\pi i} \left[\sum_{n=0}^{\infty} (z_0 - a)^n \oint_{r_1} \frac{f(z)}{(z - a)^{n+1}} dz + \sum_{n=-1}^{-\infty} (z_0 - a)^n \oint_{r_2} \frac{f(z)}{(z - a)^{n+1}} dz \right]$$

$$= \sum_{-\infty}^{\infty} (z_0 - a)^n \left[\frac{1}{2\pi i} \oint_{r_1 \le r \le r_2} \frac{f(z)}{(z - a)^{n+1}} dz \right].$$

$$(5.35)$$

Suppose that $g(z)$ is a function with a pole of order n at the point z_0; that is $g(z) = h(z)/(z - z_0)^n$, where $h(z)$ is an analytic function. Then the terms $k \le -(n + 1)$ vanish in the Laurent series. This follows from Cauchy's integral formula

$$a_k = \frac{1}{2\pi i} \oint_c (\chi - z_0)^{-k-n-1} h(\chi) d\chi = 0 \qquad (5.36)$$

for $-k - n - 1 \ge 0$.

Note that, if f has a simple pole (pole of order 1) at z_0, then it can be rewritten into $f(z) = g(z)/(z - z_0)$ for some analytic function $g(z) = (z - z_0)f(z)$ that remains after the singularity has been "split" from f. Cauchy's integral formula (5.20), and the residue can be rewritten as

$$a_{-1} = \frac{1}{2\pi i} \oint_{\partial G} \frac{g(z)}{z - z_0} dz = g(z_0). \qquad (5.37)$$

For poles of higher order, the generalized Cauchy integral formula (5.21) can be used.

5.8   Residue theorem

Suppose f is analytic on a simply connected open subset G with the exception of finitely many (or denumerably many) points z_i. Then,

$$\oint_{\partial G} f(z) dz = 2\pi i \sum_{z_i} \operatorname{Res} f(z_i) \quad . \qquad (5.38)$$

No proof is given here.

The residue theorem presents a powerful tool for calculating integrals, both real and complex. Let us first mention a rather general case of a situation often used. Suppose we are interested in the integral

$$I = \int_{-\infty}^{\infty} R(x) dx$$

with rational kernel R; that is, $R(x) = P(x)/Q(x)$, where $P(x)$ and $Q(x)$ are polynomials (or can at least be bounded by a polynomial) with no common root (and therefore factor). Suppose further that the degrees of the polynomial is

$$\deg P(x) \le \deg Q(x) - 2.$$

This condition is needed to assure that the additional upper or lower path we want to add when completing the contour does not contribute; that is, vanishes.

Now first let us analytically continue $R(x)$ to the complex plane $R(z)$; that is,

$$I = \int_{-\infty}^{\infty} R(x)\,dx = \int_{-\infty}^{\infty} R(z)\,dz.$$

Next let us close the contour by adding a (vanishing) path integral

$$\int_{\cap} R(z)\,dz = 0$$

in the upper (lower) complex plane

$$I = \int_{-\infty}^{\infty} R(z)\,dz + \int_{\cap} R(z)\,dz = \oint_{\to \&\cap} R(z)\,dz.$$

The added integral vanishes because it can be approximated by

$$\left| \int_{\cap} R(z)\,dz \right| \le \lim_{r \to \infty} \left(\frac{\text{const.}}{r^2} \pi r \right) = 0.$$

With the contour closed the residue theorem can be applied for an evaluation of I; that is,

$$I = 2\pi i \sum_{z_i} \operatorname{Res} R(z_i)$$

for all singularities z_i in the region enclosed by "$\to \& \cap$."

Let us consider some examples.

(i) Consider

$$I = \int_{-\infty}^{\infty} \frac{dx}{x^2 + 1}.$$

The analytic continuation of the kernel and the addition with vanishing a semicircle "far away" closing the integration path in the *upper* complex half-plane of z yields

$$
\begin{aligned}
I &= \int_{-\infty}^{\infty} \frac{dx}{x^2 + 1} \\
&= \int_{-\infty}^{\infty} \frac{dz}{z^2 + 1} \\
&= \int_{-\infty}^{\infty} \frac{dz}{z^2 + 1} + \int_{\cap} \frac{dz}{z^2 + 1} \\
&= \int_{-\infty}^{\infty} \frac{dz}{(z+i)(z-i)} + \int_{\cap} \frac{dz}{(z+i)(z-i)} \\
&= \oint \frac{1}{(z-i)} f(z)\,dz \text{ with } f(z) = \frac{1}{(z+i)} \\
&= 2\pi i \operatorname{Res}\left(\frac{1}{(z+i)(z-i)} \right)\Bigg|_{z=+i} \\
&= 2\pi i f(+i) \\
&= 2\pi i \frac{1}{(2i)} \\
&= \pi.
\end{aligned}
\tag{5.39}
$$

Here, Eq. (5.37) has been used. Closing the integration path in the *lower* complex half-plane of z yields (note that in this case the contour integral is negative because of the path orientation)

$$
\begin{aligned}
I &= \int_{-\infty}^{\infty} \frac{dx}{x^2 + 1} \\
&= \int_{-\infty}^{\infty} \frac{dz}{z^2 + 1} \\
&= \int_{-\infty}^{\infty} \frac{dz}{z^2 + 1} + \int_{\text{lower path}} \frac{dz}{z^2 + 1} \\
&= \int_{-\infty}^{\infty} \frac{dz}{(z+i)(z-i)} + \int_{\text{lower path}} \frac{dz}{(z+i)(z-i)} \\
&= \oint \frac{1}{(z+i)} f(z) dz \text{ with } f(z) = \frac{1}{(z-i)} \\
&= -2\pi i \operatorname{Res}\left(\frac{1}{(z+i)(z-i)} \right)\Big|_{z=-i} \\
&= -2\pi i f(-i) \\
&= 2\pi i \frac{1}{(2i)} \\
&= \pi.
\end{aligned}
$$ (5.40)

(ii) Consider

$$
F(p) = \int_{-\infty}^{\infty} \frac{e^{ipx}}{x^2 + a^2} dx
$$

with $a \neq 0$.

The analytic continuation of the kernel yields

$$
F(p) = \int_{-\infty}^{\infty} \frac{e^{ipz}}{z^2 + a^2} dz = \int_{-\infty}^{\infty} \frac{e^{ipz}}{(z - ia)(z + ia)} dz.
$$

Suppose first that $p > 0$. Then, if $z = x + iy$, $e^{ipz} = e^{ipx} e^{-py} \to 0$ for $z \to \infty$ in the *upper* half plane. Hence, we can close the contour in the upper half plane and obtain $F(p)$ with the help of the residue theorem.

If $a > 0$ only the pole at $z = +ia$ is enclosed in the contour; thus we obtain

$$
\begin{aligned}
F(p) &= 2\pi i \operatorname{Res} \frac{e^{ipz}}{(z + ia)}\Big|_{z = +ia} \\
&= 2\pi i \frac{e^{i^2 pa}}{2ia} \\
&= \frac{\pi}{a} e^{-pa}.
\end{aligned}
$$ (5.41)

If $a < 0$ only the pole at $z = -ia$ is enclosed in the contour; thus we obtain

$$
\begin{aligned}
F(p) &= 2\pi i \operatorname{Res} \frac{e^{ipz}}{(z - ia)}\Big|_{z = -ia} \\
&= 2\pi i \frac{e^{-i^2 pa}}{-2ia} \\
&= \frac{\pi}{-a} e^{-i^2 pa} \\
&= \frac{\pi}{-a} e^{pa}.
\end{aligned}
$$ (5.42)

Hence, for $a \neq 0$,

$$F(p) = \frac{\pi}{|a|} e^{-|pa|}.\qquad(5.43)$$

For $p < 0$ a very similar consideration, taking the *lower* path for continuation – and thus aquiring a minus sign because of the "clockwork" orientation of the path as compared to its interior – yields

$$F(p) = \frac{\pi}{|a|} e^{-|pa|}.\qquad(5.44)$$

(iii) If some function $f(z)$ can be expanded into a Taylor series or Laurent series, the residue can be directly obtained by the coefficient of the $\frac{1}{z}$ term. For instance, let $f(z) = e^{\frac{1}{z}}$ and $C : z(\varphi) = Re^{i\varphi}$, with $R = 1$ and $-\pi < \varphi \leq \pi$. This function is singular only in the origin $z = 0$, but this is an *essential singularity* near which the function exhibits extreme behavior. Nevertheless, $f(z) = e^{\frac{1}{z}}$ can be expanded into a Laurent series

$$f(z) = e^{\frac{1}{z}} = \sum_{l=0}^{\infty} \frac{1}{l!} \left(\frac{1}{z}\right)^l$$

around this singularity. The residue can be found by using the series expansion of $f(z)$; that is, by *comparing* its coefficient of the $1/z$ term. Hence, $\mathrm{Res}\left(e^{\frac{1}{z}}\right)\Big|_{z=0}$ is the coefficient 1 of the $1/z$ term. Thus,

$$\oint_{|z|=1} e^{\frac{1}{z}} dz = 2\pi i \, \mathrm{Res}\left(e^{\frac{1}{z}}\right)\Big|_{z=0} = 2\pi i.\qquad(5.45)$$

For $f(z) = e^{-\frac{1}{z}}$, a similar argument yields $\mathrm{Res}\left(e^{-\frac{1}{z}}\right)\Big|_{z=0} = -1$ and thus $\oint_{|z|=1} e^{-\frac{1}{z}} dz = -2\pi i$.

An alternative attempt to compute the residue, with $z = e^{i\varphi}$, yields

$$\begin{aligned}
a_{-1} = \mathrm{Res}\left(e^{\pm\frac{1}{z}}\right)\Big|_{z=0} &= \frac{1}{2\pi i} \oint_C e^{\pm\frac{1}{z}} dz \\
&= \frac{1}{2\pi i} \int_{-\pi}^{\pi} e^{\pm\frac{1}{e^{i\varphi}}} \frac{dz(\varphi)}{d\varphi} d\varphi \\
&= \pm\frac{1}{2\pi i} \int_{-\pi}^{\pi} e^{\pm\frac{1}{e^{i\varphi}}} i\, e^{\pm i\varphi} d\varphi \qquad(5.46) \\
&= \pm\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\pm e^{\mp i\varphi}} e^{\pm i\varphi} d\varphi \\
&= \pm\frac{1}{2\pi} \int_{-\pi}^{\pi} e^{\pm e^{\mp i\varphi} \pm i\varphi} d\varphi.
\end{aligned}$$

5.9   Multi-valued relationships, branch points and and branch cuts

Suppose that the Riemann surface of is *not* simply connected.

Suppose further that f is a multi-valued function (or multifunction). An argument z of the function f is called *branch point* if there is a closed curve C_z around z whose image $f(C_z)$ is an open curve. That is, the multifunction f is discontinuous in z. Intuitively speaking, branch points are the points where the various sheets of a multifunction come together.

A *branch cut* is a curve (with ends possibly open, closed, or half-open) in the complex plane across which an analytic multifunction is discontinuous. Branch cuts are often taken as lines.

5.10   Riemann surface

Suppose $f(z)$ is a multifunction. Then the various z-surfaces on which $f(z)$ is uniquely defined, together with their connections through branch points and branch cuts, constitute the Riemann surface of f. The required leafs are called *Riemann sheet*.

A point z of the function $f(z)$ is called a *branch point of order n* if through it and through the associated cut(s) $n+1$ Riemann sheets are connected.

5.11   Some special functional classes

The requirement that a function is holomorphic (analytic, differentiable) puts some stringent conditions on its type, form, and on its behaviour. For instance, let $z_0 \in G$ the limit of a sequence $\{z_n\} \in G$, $z_n \neq z_0$. Then it can be shown that, if two analytic functions f und g on the domain G coincide in the points z_n, then they coincide on the entire domain G.

5.11.1   Entire function

An function is said to be an *entire function* if it is defined and differentiable (holomorphic, analytic) in the entire *finite complex plane* $\mathbb{C}$.

An entire function may be either a *rational function* $f(z) = P(z)/Q(z)$ which can be written as the ratio of two polynomial functions $P(z)$ and $Q(z)$, or it may be a *transcendental function* such as e^z or $\sin z$.

The *Weierstrass factorization theorem* states that an entire function can be represented by a (possibly infinite [6]) product involving its zeroes [i.e., the points z_k at which the function vanishes $f(z_k) = 0$]. For example (for a proof, see Eq. (6.2) of [7]),

$$\sin z = z \prod_{k=1}^{\infty} \left[1 - \left(\frac{z}{\pi k} \right)^2 \right]. \tag{5.47}$$

[6] Theodore W. Gamelin. *Complex Analysis.* Springer, New York, NY, 2001

[7] J. B. Conway. *Functions of Complex Variables. Volume I.* Springer, New York, NY, 1973

5.11.2   Liouville's theorem for bounded entire function

Liouville's theorem states that a bounded (i.e., its absolute square is finite everywhere in $\mathbb{C}$) entire function which is defined at infinity is a constant. Conversely, a nonconstant entire function cannot be bounded.

For a proof, consider the integral representation of the derivative $f'(z)$ of some bounded entire function $f(z) < C$ (suppose the bound is C) obtained through Cauchy's integral formula (5.21), taken along a circular path with arbitrarily large radius r of length $2\pi r$ in the limit of infinite radius; that is,

$$\left| f'(z_0) \right| = \left| \frac{1}{2\pi i} \oint_{\partial G} \frac{f(z)}{(z-z_0)^2} dz \right|$$
$$< \frac{1}{2\pi i} \oint_{\partial G} \frac{|f(z)|}{(z-z_0)^2} dz < \frac{1}{2\pi i} 2\pi r \frac{C}{r^2} = \frac{C}{r} \overset{r \to \infty}{\longrightarrow} 0. \tag{5.48}$$

As a result, $f(z_0) = 0$ and thus $f = A \in \mathbb{C}$.

A *generalized Liouville theorem* states that if $f : \mathbb{C} \to \mathbb{C}$ is an entire function and if for some real number c and some positive integer k it

It may (wrongly) appear that $\sin z$ is nonconstant and bounded. However it is only bounded on the real axis; indeeed, $\sin i y = (1/2)(e^y + e^{-y}) \to \infty$ for $y \to \infty$.

holds that $|f(z)| \le c|z|^k$ for all z with $|z| > 1$, then f is a polynomial in z of degree at most k.

No proof [8] is presented here. This theorem is needed for an investigation into the general form of the Fuchsian differential equation on page 203.

[8] Robert E. Greene and Stephen G. Krantz. *Function theory of one complex variable*, volume 40 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, third edition, 2006

5.11.3   Picard's theorem

Picard's theorem states that any entire function that misses two or more points $f : \mathbb{C} \mapsto \mathbb{C} - \{z_1, z_2, \ldots\}$ is constant. Conversely, any nonconstant entire function covers the entire complex plane $\mathbb{C}$ except a single point.

An example for a nonconstant entire function is e^z which never reaches the point 0.

5.11.4   Meromorphic function

If f has no singularities other than poles in the domain G it is called *meromorphic* in the domain G.

We state without proof (e.g., Theorem 8.5.1 of [9]) that a function f which is meromorphic in the extended plane is a rational function $f(z) = P(z)/Q(z)$ which can be written as the ratio of two polynomial functions $P(z)$ and $Q(z)$.

[9] Einar Hille. *Analytic Function Theory*. Ginn, New York, 1962. 2 Volumes

5.12   Fundamental theorem of algebra

The *factor theorem* states that a polynomial $P(z)$ in z of degree k has a factor $z - z_0$ if and only if $P(z_0) = 0$, and can thus be written as $P(z) = (z - z_0)Q(z)$, where $Q(z)$ is a polynomial in z of degree $k - 1$. Hence, by iteration,

$$P(z) = \alpha \prod_{i=1}^{k} (z - z_i), \qquad (5.49)$$

where $\alpha \in \mathbb{C}$.

No proof is presented here.

The *fundamental theorem of algebra* states that every polynomial (with arbitrary complex coefficients) has a root [i.e. solution of $f(z) = 0$] in the complex plane. Therefore, by the factor theorem, the number of roots of a polynomial, up to multiplicity, equals its degree.

Again, no proof is presented here.

https://www.dpmms.cam.ac.uk/ wtg10/ftalg.html

6

Brief review of Fourier transforms

6.0.1 Functional spaces

That complex continuous waveforms or functions are comprised of a number of harmonics seems to be an idea at least as old as the Pythagoreans. In physical terms, Fourier analysis [1] attempts to decompose a function into its constituent frequencies, known as a frequency spectrum. Thereby the goal is the expansion of periodic and aperiodic functions into sine and cosine functions. Fourier's observation or conjecture is, informally speaking, that any "suitable" function $f(x)$ can be expressed as a possibly infinite sum (i.e. linear combination), of sines and cosines of the form

$$f(x) = \sum_{k=-\infty}^{\infty} [A_k \cos(Ckx) + B_k \sin(Ckx)]. \tag{6.1}$$

Moreover, it is conjectured that any "suitable" function $f(x)$ can be expressed as a possibly infinite sum (i.e. linear combination), of exponentials; that is,

$$f(x) = \sum_{k=-\infty}^{\infty} D_k e^{ikx}. \tag{6.2}$$

More generally, it is conjectured that any "suitable" function $f(x)$ can be expressed as a possibly infinite sum (i.e. linear combination), of other (possibly orthonormal) functions $g_k(x)$; that is,

$$f(x) = \sum_{k=-\infty}^{\infty} \gamma_k g_k(x). \tag{6.3}$$

The bigger picture can then be viewed in terms of *functional (vector) spaces*: these are spanned by the elementary functions g_k, which serve as elements of a *functional basis* of a possibly infinite-dimensional vector space. Suppose, in further analogy to the set of all such functions $\mathfrak{G} = \bigcup_k g_k(x)$ to the (Cartesian) standard basis, we can consider these elementary functions g_k to be *orthonormal* in the sense of a *generalized functional scalar product* [cf. also Section 11.5 on page 222; in particular Eq. (11.115)]

$$\langle g_k \mid g_l \rangle = \int_a^b g_k(x) g_l(x) \, dx = \delta_{kl}. \tag{6.4}$$

[1] T. W. Körner. *Fourier Analysis*. Cambridge University Press, Cambridge, UK, 1988; Kenneth B. Howell. *Principles of Fourier analysis*. Chapman & Hall/CRC, Boca Raton, London, New York, Washington, D.C., 2001; and Russell Herman. *Introduction to Fourier and Complex Analysis with Applications to the Spectral Analysis of Signals*. University of North Carolina Wilmington, Wilmington, NC, 2010. URL http://people.uncw.edu/hermanr/mat367/FCABook/Book2010/FTCA-book.pdf. Creative Commons Attribution-NoncommercialShare Alike 3.0 United States License

One could arrange the coefficients γ_k into a tuple (an ordered list of elements) $(\gamma_1, \gamma_2, \ldots)$ and consider them as components or coordinates of a vector with respect to the linear orthonormal functional basis $\mathfrak{G}$.

6.0.2   Fourier series

Suppose that a function $f(x)$ is periodic in the interval $[-\frac{L}{2}, \frac{L}{2}]$ with period L. (Alternatively, the function may be only defined in this interval.) A function $f(x)$ is *periodic* if there exist a period $L \in \mathbb{R}$ such that, for all x in the domain of f,

$$f(L + x) = f(x). \tag{6.5}$$

Then, under certain "mild" conditions – that is, f must be piecewise continuous and have only a finite number of maxima and minima – f can be decomposed into a *Fourier series*

$$
\begin{aligned}
f(x) &= \tfrac{a_0}{2} + \textstyle\sum_{k=1}^{\infty} \left[a_k \cos\left(\tfrac{2\pi}{L} kx\right) + b_k \sin\left(\tfrac{2\pi}{L} kx\right) \right], \text{ with} \\
a_k &= \tfrac{2}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x) \cos\left(\tfrac{2\pi}{L} kx\right) dx \text{ for } k \geq 0 \\
b_k &= \tfrac{2}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x) \sin\left(\tfrac{2\pi}{L} kx\right) dx \text{ for } k > 0.
\end{aligned}
\tag{6.6}
$$

For a (heuristic) proof, consider the Fourier conjecture (6.1), and compute the coefficients A_k, B_k, and C.

First, observe that we have assumed that f is periodic in the interval $[-\frac{L}{2}, \frac{L}{2}]$ with period L. This should be reflected in the sine and cosine terms of (6.1), which themselves are periodic functions in the interval $[-\pi, \pi]$ with period 2π. Thus in order to map the functional period of f into the sines and cosines, we can "stretch/shrink" L into 2π; that is, C in Eq. (6.1) is identified with

$$C = \frac{2\pi}{L}. \tag{6.7}$$

Thus we obtain

$$f(x) = \sum_{k=-\infty}^{\infty} \left[A_k \cos(\frac{2\pi}{L} kx) + B_k \sin(\frac{2\pi}{L} kx) \right]. \tag{6.8}$$

Now use the following properties: (i) for $k = 0$, $\cos(0) = 1$ and $\sin(0) = 0$. Thus, by comparing the coefficient a_0 in (6.6) with A_0 in (6.1) we obtain $A_0 = \frac{a_0}{2}$.

(ii) Since $\cos(x) = \cos(-x)$ is an *even function* of x, we can rearrange the summation by combining identical functions $\cos(-\frac{2\pi}{L} kx) = \cos(\frac{2\pi}{L} kx)$, thus obtaining $a_k = A_{-k} + A_k$ for $k > 0$.

(iii) Since $\sin(x) = -\sin(-x)$ is an *odd function* of x, we can rearrange the summation by combining identical functions $\sin(-\frac{2\pi}{L} kx) = -\sin(\frac{2\pi}{L} kx)$, thus obtaining $b_k = -B_{-k} + B_k$ for $k > 0$.

Having obtained the same form of the Fourier series of $f(x)$ as exposed in (6.6), we now turn to the derivation of the coefficients a_k and b_k. a_0 can be derived by just considering the functional scalar product exposedin Eq. (6.4) of $f(x)$ with the constant identity function $g(x) = 1$;

For proofs and additional information see §8.1 in

Kenneth B. Howell. *Principles of Fourier analysis*. Chapman & Hall/CRC, Boca Raton, London, New York, Washington, D.C., 2001

that is,

$$\langle g \mid f \rangle = \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x)\,dx$$

$$= \int_{-\frac{L}{2}}^{\frac{L}{2}} \left\{ \frac{a_0}{2} + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right) \right] \right\} dx \qquad (6.9)$$

$$= a_0 \frac{L}{2},$$

and hence

$$a_0 = \frac{2}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x)\,dx \qquad (6.10)$$

In a very similar manner, the other coefficients can be computed by considering $\left\langle \cos\left(\frac{2\pi}{L} kx\right) \mid f(x) \right\rangle \left\langle \sin\left(\frac{2\pi}{L} kx\right) \mid f(x) \right\rangle$ and exploiting the *orthogonality relations for sines and cosines*

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} \sin\left(\frac{2\pi}{L} kx\right) \cos\left(\frac{2\pi}{L} lx\right) dx = 0,$$

$$\int_{-\frac{L}{2}}^{\frac{L}{2}} \cos\left(\frac{2\pi}{L} kx\right) \cos\left(\frac{2\pi}{L} lx\right) dx = \int_{-\frac{L}{2}}^{\frac{L}{2}} \sin\left(\frac{2\pi}{L} kx\right) \sin\left(\frac{2\pi}{L} lx\right) dx = \frac{L}{2}\delta_{kl}.$$

$$(6.11)$$

For the sake of an example, let us compute the Fourier series of

$$f(x) = |x| = \begin{cases} -x, & \text{for } -\pi \le x < 0, \\ +x, & \text{for } 0 \le x \le \pi. \end{cases}$$

First observe that $L = 2\pi$, and that $f(x) = f(-x)$; that is, f is an *even* function of x; hence $b_n = 0$, and the coefficients a_n can be obtained by considering only the integration between 0 and π.

For $n = 0$,

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{\pi} dx f(x) = \frac{2}{\pi} \int_{0}^{\pi} x\,dx = \pi.$$

For $n > 0$,

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx)\,dx = \frac{2}{\pi} \int_{0}^{\pi} x \cos(nx)\,dx =$$

$$= \frac{2}{\pi} \left[\frac{\sin(nx)}{n} x \Big|_{0}^{\pi} - \int_{0}^{\pi} \frac{\sin(nx)}{n}\,dx \right] = \frac{2}{\pi} \frac{\cos(nx)}{n^2} \Big|_{0}^{\pi} =$$

$$= \frac{2}{\pi} \frac{\cos(n\pi) - 1}{n^2} = -\frac{4}{\pi n^2} \sin^2 \frac{n\pi}{2} = \begin{cases} 0 & \text{for even } n \\ -\dfrac{4}{\pi n^2} & \text{for odd } n \end{cases}$$

Thus,

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi} \left(\cos x + \frac{\cos 3x}{9} + \frac{\cos 5x}{25} + \cdots \right) =$$

$$= \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\cos[(2n+1)x]}{(2n+1)^2}.$$

One could arrange the coefficients $(a_0, a_1, b_1, a_2, b_2, \ldots)$ into a tuple (an ordered list of elements) and consider them as components or coordinats of a vector spanned by the linear independent sine and cosine functions which serve as a basis of an infinite dimensional vector space.

6.0.3 Exponential Fourier series

Suppose again that a function is periodic in the interval $[-\frac{L}{2}, \frac{L}{2}]$ with period L. Then, under certain "mild" conditions – that is, f must be piecewise continuous and have only a finite number of maxima and minima – f can be decomposed into an *exponential Fourier series*

$$f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \text{ with}$$
$$c_k = \frac{1}{L} \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x') e^{-ikx'} dx'. \tag{6.12}$$

The expontial form of the Fourier series can be derived from the Fourier series (6.6) by Euler's formula (5.2), in particular, $e^{ik\varphi} = \cos(k\varphi) + i\sin(k\varphi)$, and thus

$$\cos(k\varphi) = \frac{1}{2}\left(e^{ik\varphi} + e^{-ik\varphi}\right), \text{ as well as } \sin(k\varphi) = \frac{1}{2i}\left(e^{ik\varphi} - e^{-ik\varphi}\right).$$

By comparing the coefficients of (6.6) with the coefficients of (6.12), we obtain

$$a_k = c_k + c_{-k} \text{ for } k \geq 0,$$
$$b_k = i(c_k - c_{-k}) \text{ for } k > 0, \tag{6.13}$$

or

$$c_k = \begin{cases} \frac{1}{2}(a_k - ib_k) \text{ for } k > 0, \\ \frac{a_0}{2} \text{ for } k = 0, \\ \frac{1}{2}(a_{-k} + ib_{-k}) \text{ for } k < 0. \end{cases} \tag{6.14}$$

Eqs. (6.12) can be combined into

$$f(x) = \frac{1}{L} \sum_{\check{k}=-\infty}^{\infty} \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x') e^{-i\check{k}(x'-x)} dx'. \tag{6.15}$$

6.0.4 Fourier transformation

Suppose we define $\Delta k = 2\pi/L$, or $1/L = \Delta k/2\pi$. Then Eq. (6.15) can be rewritten as

$$f(x) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \int_{-\frac{L}{2}}^{\frac{L}{2}} f(x') e^{-ik(x'-x)} dx' \Delta k. \tag{6.16}$$

Now, in the "aperiodic" limit $L \to \infty$ we obtain the *Fourier transformation* and the *Fourier inversion* $\mathscr{F}^{-1}[\mathscr{F}[f(x)]] = \mathscr{F}[\mathscr{F}^{-1}[f(x)]] = f(x)$ by

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x') e^{-ik(x'-x)} dx' dk, \text{ whereby}$$
$$\mathscr{F}^{-1}[\tilde{f}(k)] = f(x) = \alpha \int_{-\infty}^{\infty} \tilde{f}(k) e^{\pm ikx} dk, \text{ and}$$
$$\mathscr{F}[f(x)] = \tilde{f}(k) = \beta \int_{-\infty}^{\infty} f(x') e^{\mp ikx'} dx'. \tag{6.17}$$

$\mathscr{F}[f(x)] = \tilde{f}(k)$ is called the *Fourier transform* of $f(x)$. *Per* convention, either one of the two sign pairs $+-$ or $-+$ must be chosen. The factors α and β must be chosen such that

$$\alpha\beta = \frac{1}{2\pi}; \tag{6.18}$$

that is, the factorization can be "spread evenly among α and β," such that $\alpha = \beta = 1/\sqrt{2\pi}$, or "unevenly," such as, for instance, $\alpha = 1$ and $\beta = 1/2\pi$, or $\alpha = 1/2\pi$ and $\beta = 1$.

Most generally, the Fourier transformations can be rewritten (change of integration constant), with arbitrary $A, B \in \mathbb{R}$, as

$$
\begin{aligned}
\mathscr{F}^{-1}[\tilde{f}(k)](x) = f(x) &= B \int_{-\infty}^{\infty} \tilde{f}(k) e^{iAkx} dk, \text{ and} \\
\mathscr{F}[f(x)](k) = \tilde{f}(k) &= \frac{A}{2\pi B} \int_{-\infty}^{\infty} f(x') e^{-iAkx'} dx'.
\end{aligned}
\tag{6.19}
$$

The coice $A = 2\pi$ and $B = 1$ renders a very symmetric form of (6.19); more precisely,

$$
\begin{aligned}
\mathscr{F}^{-1}[\tilde{f}(k)](x) = f(x) &= \int_{-\infty}^{\infty} \tilde{f}(k) e^{2\pi ikx} dk, \text{ and} \\
\mathscr{F}[f(x)](k) = \tilde{f}(k) &= \int_{-\infty}^{\infty} f(x') e^{-2\pi ikx'} dx'.
\end{aligned}
\tag{6.20}
$$

For the sake of an example, assume $A = 2\pi$ and $B = 1$ in Eq. (6.19), therefore starting with (6.20), and consider the Fourier transform of the *Gaussian function*

$$
\varphi(x) = e^{-\pi x^2} \quad . \tag{6.21}
$$

As a hint, notice that e^{-t^2} is analytic in the region $0 \le \operatorname{Im} t \le \sqrt{\pi}k$; also, as will be shown in Eqs. (7.18), the *Gaussian integral* is

$$
\int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi} \quad . \tag{6.22}
$$

With $A = 2\pi$ and $B = 1$ in Eq. (6.19), the Fourier transform of the Gaussian function is

$$
\begin{aligned}
\mathscr{F}[\varphi(x)](k) = \widetilde{\varphi}(k) &= \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi ikx} dx \\
&\quad \text{[completing the exponent]} \\
&= \int_{-\infty}^{\infty} e^{-\pi k^2} e^{-\pi(x+ik)^2} dx
\end{aligned}
\tag{6.23}
$$

The variable transformation $t = \sqrt{\pi}(x + ik)$ yields $dt/dx = \sqrt{\pi}$; thus $dx = dt/\sqrt{\pi}$, and

$$
\mathscr{F}[\varphi(x)](k) = \widetilde{\varphi}(k) = \frac{e^{-\pi k^2}}{\sqrt{\pi}} \int_{-\infty+i\sqrt{\pi}k}^{+\infty+i\sqrt{\pi}k} e^{-t^2} dt \tag{6.24}
$$

Let us rewrite the integration (6.24) into the Gaussian integral by considering the closed path whose "left and right pieces vanish;" moreover,

$$
\oint_{\mathscr{C}} dt\, e^{-t^2} = \int_{+\infty}^{-\infty} e^{-t^2} dt + \int_{-\infty+i\sqrt{\pi}k}^{+\infty+i\sqrt{\pi}k} e^{-t^2} dt = 0, \tag{6.25}
$$

because e^{-t^2} is analytic in the region $0 \le \operatorname{Im} t \le \sqrt{\pi}k$. Thus, by substituting

$$
\int_{-\infty+i\sqrt{\pi}k}^{+\infty+i\sqrt{\pi}k} e^{-t^2} dt = \int_{-\infty}^{+\infty} e^{-t^2} dt, \tag{6.26}
$$

in (6.24) and by insertion of the value $\sqrt{\pi}$ for the *Gaussian integral*, as shown in Eq. (7.18), we finally obtain

$$
\mathscr{F}[\varphi(x)](k) = \widetilde{\varphi}(k) = \frac{e^{-\pi k^2}}{\sqrt{\pi}} \underbrace{\int_{-\infty}^{+\infty} e^{-t^2} dt}_{\sqrt{\pi}} = e^{-\pi k^2}. \tag{6.27}
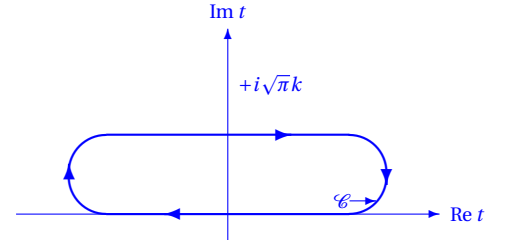$$

Figure 6.1: Integration path to compute the Fourier transform of the Gaussian.

A very similar calculation yields

$$\mathscr{F}^{-1}[\varphi(k)](x) = \varphi(x) = e^{-\pi x^2}.$$ (6.28)

Eqs. (6.27) and (6.28) establish the fact that the Gaussian function $\varphi(x) = e^{-\pi x^2}$ defined in (6.21) is an *eigenfunction* of the Fourier transformations $\mathscr{F}$ and $\mathscr{F}^{-1}$ with associated eigenvalue 1.

With a slightly different definition the Gaussian function $f(x) = e^{-x^2/2}$ is also an eigenfunction of the operator

$$\mathscr{H} = -\frac{d^2}{dx^2} + x^2$$ (6.29)

corresponding to a harmonic oscillator. The resulting eigenvalue equation is

$$\mathscr{H}f(x) = \left[-\frac{d^2}{dx^2} + x^2 \right] f(x) = \left[-\frac{d}{dx}(-x) + x^2 \right] f(x) = f(x);$$ (6.30)

with eigenvalue 1.

Instead of going too much into the details here, it may suffice to say that the *Hermite functions*

$$h_n(x) = \pi^{-1/4}(2^n n!)^{-1/2} \left(\frac{d}{dx} - x \right)^n e^{-x^2/2} = \pi^{-1/4}(2^n n!)^{-1/2} H_n(x) e^{-x^2/2}$$ (6.31)

are all eigenfunctions of the Fourier transform with the eigenvalue $i^n\sqrt{2\pi}$. The polynomial $H_n(x)$ of degree n is called *Hermite polynomial.* Hermite functions form a complete system, so that any function g (with $\int |g(x)|^2 dx < \infty$) has a *Hermite expansion*

$$g(x) = \sum_{k=0}^{\infty} \langle g, h_n \rangle h_n(x).$$ (6.32)

This is an example of an *eigenfunction expansion.*

7

Distributions as generalized functions

7.1 *Heuristically coping with discontinuities*

What follows are "recipes" and a "cooking course" for some "dishes"
Heaviside, Dirac and others have enjoyed "eating," alas without being
able to "explain their digestion" (cf. the citation by Heaviside on page).

Insofar theoretical physics is natural philosophy, the question arises
if physical entities need to be smooth and continuous – in particular, if
physical functions need to be *smooth* (i.e., in C^∞), having derivatives of
all orders [1] (such as polynomials, trigonometric and exponential func-
tions) – as "nature abhors sudden discontinuities," or if we are willing to
allow and conceptualize singularities of different sorts. Other, entirely
different, scenarios are discrete [2] computer-generated universes [3]. This
little course is no place for preference and judgments regarding these
matters. Let me just point out that contemporary mathematical physics
is not only leaning toward, but appears to be deeply committed to dis-
continuities; both in classical and quantized field theories dealing with
"point charges," as well as in general relativity, the (nonquantized field
theoretical) geometrodynamics of graviation, dealing with singularities
such as "black holes" or "initial singularities" of various sorts.

Discontinuities were introduced quite naturally as electromagnetic
pulses, which can, for instance be described with the *Heaviside function*
$H(t)$ representing vanishing, zero field strength until time $t = 0$, when
suddenly a constant electrical field is "switched on eternally." It is quite
natural to ask what the derivative of the (infinite pulse) function $H(t)$
might be. — At this point the reader is kindly ask to stop reading for a
moment and contemplate on what kind of function that might be.

Heuristically, if we call this derivative the *(Dirac) delta function* δ
defined by $\delta(t) = \frac{dH(t)}{dt}$, we can assure ourselves of two of its properties
(i) "$\delta(t) = 0$ for $t \neq 0$," as well as as the antiderivative of the Heaviside
function, yielding (ii) "$\int_{-\infty}^{\infty} \delta(t) dt = \int_{-\infty}^{\infty} \frac{dH(t)}{dt} dt = H(\infty) - H(-\infty) = 1 - 0 = 1$."

Indeed, we could follow a pattern of "growing discontinuity," reach-
able by ever higher and higher derivatives of the absolute value (or mod-

[1] William F. Trench. Introduction to real
analysis. Free Hyperlinked Edition 2.01,
2012. URL http://ramanujan.math.
trinity.edu/wtrench/texts/TRENCH_
REAL_ANALYSIS.PDF

[2] Konrad Zuse. Discrete mathematics
and Rechnender Raum. 1994. URL
http://www.zib.de/PaperWeb/
abstracts/TR-94-10/; and Konrad
Zuse. *Rechnender Raum*. Friedrich
Vieweg & Sohn, Braunschweig, 1969

[3] Edward Fredkin. Digital mechanics. an
informational process based on reversible
universal cellular automata. *Physica*,
D45:254–270, 1990. DOI: 10.1016/0167-
2789(90)90186-S. URL http://doi.
org/10.1016/0167-2789(90)90186-S;
Tommaso Toffoli. The role of the observer
in uniform systems. In George J. Klir, edi-
tor, *Applied General Systems Research: Re-
cent Developments and Trends*, pages 395–
400. Plenum Press, Springer US, New York,
London, and Boston, MA, 1978. ISBN
978-1-4757-0555-3. DOI: 10.1007/978-1-
4757-0555-3_29. URL http://doi.org/
10.1007/978-1-4757-0555-3_29; and
Karl Svozil. Computational universes.
Chaos, Solitons & Fractals, 25(4):845–859,
2006. DOI: 10.1016/j.chaos.2004.11.055.
URL http://doi.org/10.1016/j.
chaos.2004.11.055
This heuristic definition of the Dirac
delta function $\delta_y(x) = \delta(x, y) = \delta(x - y)$
with a discontinuity at y is not unlike
the discrete Kronecker symbol δ_{ij},
as (i) $\delta_{ij} = 0$ for $i \neq j$, as well as (ii)
"$\sum_{i=-\infty}^{\infty} \delta_{ij} = \sum_{i=-\infty}^{\infty} \delta_{ij} = 1$," We may
even define the Kronecker symbol δ_{ij} as
the difference quotient of some "discrete
Heaviside function" $H_{ij} = 1$ for $i \geq j$, and
$H_{i,j} = 0$ else: $\delta_{ij} = H_{ij} - H_{(i-1)j} = 1$ only
for $i = j$; else it vanishes.

ulus); that is, we shall pursue the path sketched by

$$|x| \xrightarrow{\frac{d}{dx}} \text{sgn}(x),\ H(x) \xrightarrow{\frac{d}{dx}} \delta(x) \xrightarrow{\frac{d^n}{dx^n}} \delta^{(n)}(x).$$

Objects like $|x|$, $H(t)$ or $\delta(t)$ may be heuristically understandable as "functions" not unlike the regular analytic functions; alas their nth derivatives cannot be straightforwardly defined. In order to cope with a formally precised definition and derivation of (infinite) pulse functions and to achieve this goal, a theory of *generalized functions,* or, used synonymously, *distributions* has been developed. In what follows we shall develop the theory of distributions; always keeping in mind the assumptions regarding (dis)continuities that make necessary this part of calculus.

The *Ansatz* pursued will be to "pair" (that is, to multiply) these generalized functions F with suitable "good" test functions φ, and integrate over these functional pairs $F\varphi$. Thereby we obtain a linear continuous functional $F[\varphi]$, also denoted by $\langle F, \varphi \rangle$. This strategy allows for the "transference" or "shift" of operations on, and transformations of, F – such as differentiations or Fourier transformations, but also multiplications with polynomials or other smooth functions – to the test function φ according to *adjoint identities*

$$\langle \mathbf{T}F, \varphi \rangle = \langle F, \mathbf{S}\varphi \rangle. \tag{7.1}$$

For example, for n-fold differention,

$$\mathbf{S} = (-1)^n \mathbf{T} = (-1)^n \frac{d^{(n)}}{dx^{(n)}}, \tag{7.2}$$

and for the Fourier transformation,

$$\mathbf{S} = \mathbf{T} = \mathscr{F}. \tag{7.3}$$

For some (smooth) functional multiplier $g(x) \in C^\infty$,

$$\mathbf{S} = \mathbf{T} = g(x). \tag{7.4}$$

One more issue is the problem of the meaning and existence of *weak solutions* (also called generalized solutions) of differential equations for which, if interpreted in terms of regular functions, the derivatives may not all exist.

Take, for example, the wave equation in one spatial dimension $\frac{\partial^2}{\partial t^2} u(x,t) = c^2 \frac{\partial^2}{\partial x^2} u(x,t)$. It has a solution of the form [4] $u(x,t) = f(x - ct) + g(x + ct)$, where f and g characterize a travelling "shape" of inert, unchanged form. There is no obvious physical reason why the pulse shape function f or g should be differentiable, alas if it is not, then u is not differentiable either. What if we, for instance, set $g = 0$, and identify $f(x - ct)$ with the Heaviside infinite pulse function $H(x - ct)$?

7.2   General distribution

Suppose we have some "function" $F(x)$; that is, $F(x)$ could be either a regular analytical function, such as $F(x) = x$, or some other, "weirder

function," such as the Dirac delta function, or the derivative of the Heaviside (unit step) function, which might be "highly discontinuous." As an *Ansatz,* we may associate with this "function" $F(x)$ a *distribution,* or, used synonymously, a *generalized function* $F[\varphi]$ or $\langle F, \varphi \rangle$ which in the "weak sense" is defined as a *continuous linear functional* by integrating $F(x)$ together with some "good" *test function* φ as follows [5]:

$$F(x) \longleftrightarrow \langle F, \varphi \rangle \equiv F[\varphi] = \int_{-\infty}^{\infty} F(x)\varphi(x)\,dx. \tag{7.5}$$

We say that $F[\varphi]$ or $\langle F, \varphi \rangle$ is the distribution *associated with* or *induced by* $F(x)$.

One interpretation of $F[\varphi] \equiv \langle F, \varphi \rangle$ is that F stands for a sort of "measurement device," and φ represents some "system to be measured;" then $F[\varphi] \equiv \langle F, \varphi \rangle$ is the "outcome" or "measurement result."

Thereby, it completely suffices to say what F "does to" some test function φ; there is nothing more to it.

For example, the Dirac Delta function $\delta(x)$ is completely characterised by

$$\delta(x) \longleftrightarrow \delta[\varphi] \equiv \langle \delta, \varphi \rangle = \varphi(0);$$

likewise, the shifted Dirac Delta function $\delta_y(x) \equiv \delta(x - y)$ is completely characterised by

$$\delta_y(x) \equiv \delta(x - y) \longleftrightarrow \delta_y[\varphi] \equiv \langle \delta_y, \varphi \rangle = \varphi(y).$$

Many other (regular) functions which are usually not integrable in the interval $(-\infty, +\infty)$ will, through the pairing with a "suitable" or "good" test function φ, induce a distribution.

For example, take

$$1 \longleftrightarrow 1[\varphi] \equiv \langle 1, \varphi \rangle = \int_{-\infty}^{\infty} \varphi(x)\,dx,$$

or

$$x \longleftrightarrow x[\varphi] \equiv \langle x, \varphi \rangle = \int_{-\infty}^{\infty} x\varphi(x)\,dx,$$

or

$$e^{2\pi i a x} \longleftrightarrow e^{2\pi i a x}[\varphi] \equiv \langle e^{2\pi i a x}, \varphi \rangle = \int_{-\infty}^{\infty} e^{2\pi i a x}\varphi(x)\,dx.$$

7.2.1 Duality

Sometimes, $F[\varphi] \equiv \langle F, \varphi \rangle$ is also written in a scalar product notation; that is, $F[\varphi] = \langle F \mid \varphi \rangle$. This emphasizes the pairing aspect of $F[\varphi] \equiv \langle F, \varphi \rangle$. In this view the set of all distributions F is the *dual space* of the set of test functions φ.

7.2.2 Linearity

Recall that a *linear* functional is some mathematical entity which maps a function or another mathematical object into scalars in a linear manner; that is, as the integral is linear, we obtain

$$F[c_1 \varphi_1 + c_2 \varphi_2] = c_1 F[\varphi_1] + c_2 F[\varphi_2]; \tag{7.6}$$

or, in the bracket notation,

$$\langle F, c_1\varphi_1 + c_2\varphi_2 \rangle = c_1\langle F, \varphi_1 \rangle + c_2\langle F, \varphi_2 \rangle. \tag{7.7}$$

This linearity is guaranteed by integration.

7.2.3   Continuity

One way of expressing *continuity* is the following:

$$\text{if } \varphi_n \overset{n\to\infty}{\longrightarrow} \varphi, \text{ then } F[\varphi_n] \overset{n\to\infty}{\longrightarrow} F[\varphi], \tag{7.8}$$

or, in the bracket notation,

$$\text{if } \varphi_n \overset{n\to\infty}{\longrightarrow} \varphi, \text{ then } \langle F, \varphi_n \rangle \overset{n\to\infty}{\longrightarrow} \langle F, \varphi \rangle. \tag{7.9}$$

7.3   Test functions

7.3.1   Desiderata on test functions

By invoking test functions, we would like to be able to differentiate distributions very much like ordinary functions. We would also like to transfer differentiations to the functional context. How can this be implemented in terms of possible "good" properties we require from the behaviour of test functions, in accord with our wishes?

Consider the partial integration obtained from $(uv)' = u'v + uv'$; thus $\int(uv)' = \int u'v + \int uv'$, and finally $\int u'v = \int(uv)' - \int uv'$, thereby effectively allowing us to "shift" or "transfer" the differentiation of the original function to the test function. By identifying u with the generalized function g (such as, for instance δ), and v with the test function φ, respectively, we obtain

$$\begin{aligned}
\langle g', \varphi \rangle \equiv g'[\varphi] &= \int_{-\infty}^{\infty} g'(x)\varphi(x)\,dx \\
&= g(x)\varphi(x)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} g(x)\varphi'(x)\,dx \\
&= \underbrace{g(\infty)\varphi(\infty)}_{\text{should vanish}} - \underbrace{g(-\infty)\varphi(-\infty)}_{\text{should vanish}} - \int_{-\infty}^{\infty} g(x)\varphi'(x)\,dx \\
&= -g[\varphi'] \equiv -\langle g, \varphi' \rangle.
\end{aligned} \tag{7.10}$$

We can justify the two main requirements of "good" test functions, at least for a wide variety of purposes:

1. that they "sufficiently" vanish at infinity – this can, for instance, be achieved by requiring that their support (the set of arguments x where $g(x) \neq 0$) is finite; and

2. that they are continuosly differentiable – indeed, by induction, that they are arbitrarily often differentiable.

In what follows we shall enumerate three types of suitable test functions satisfying these *desiderata*. One should, however, bear in mind that the class of "good" test functions depends on the distribution. Take, for example, the Dirac delta function $\delta(x)$. It is so "concentrated" that any

(infinitely often) differentiable – even constant – function $f(x)$ defind "around $x = 0$" can serve as a "good" test function (with respect to δ), as $f(x)$ is only evaluated at $x = 0$; that is, $\delta[f] = f(0)$. This is again an indication of the *duality* between distributions on the one hand, and their test functions on the other hand.

7.3.2 Test function class I

Recall that we require [6] our tests functions φ to be infinitely often differentiable. Furthermore, in order to get rid of terms at infinity "in a straightforward, simple way," suppose that their support is compact. Compact support means that $\varphi(x)$ does not vanish only at a finite, bounded region of x. Such a "good" test function is, for instance,

$$\varphi_{\sigma,a}(x) = \begin{cases} e^{-\frac{1}{1-((x-a)/\sigma)^2}} & \text{for } |\frac{x-a}{\sigma}| < 1, \\ 0 & \text{else.} \end{cases} \tag{7.11}$$

In order to show that $\varphi_{\sigma,a}$ is a suitable test function, we have to prove its infinite differetiability, as well as the compactness of its support $M_{\varphi_{\sigma,a}}$. Let

$$\varphi_{\sigma,a}(x) := \varphi\left(\frac{x-a}{\sigma}\right)$$

and thus

$$\varphi(x) = \begin{cases} e^{-\frac{1}{1-x^2}} & \text{for } |x| < 1 \\ 0 & \text{for } |x| \geq 1 \end{cases}$$

This function is drawn in Fig. 7.1.

First, note, by definition, the support $M_\varphi = (-1, 1)$, because $\varphi(x)$ vanishes outside $(-1, 1)$).

Second, consider the differentiability of $\varphi(x)$; that is $\varphi \in C^\infty(\mathbb{R})$? Note that $\varphi^{(0)} = \varphi$ is continuous; and that $\varphi^{(n)}$ is of the form

Figure 7.1: Plot of a test function $\varphi(x)$.

$$\varphi^{(n)}(x) = \begin{cases} \frac{P_n(x)}{(x^2-1)^{2n}} e^{\frac{1}{x^2-1}} & \text{for } |x| < 1 \\ 0 & \text{for } |x| \geq 1, \end{cases}$$

where $P_n(x)$ is a finite polynomial in x ($\varphi(u) = e^u \implies \varphi'(u) = \frac{d\varphi}{du}\frac{du}{dx^2}\frac{dx^2}{dx} = \varphi(u)\left(-\frac{1}{(x^2-1)^2}\right)2x$ etc.) and $[x = 1-\varepsilon] \implies x^2 = 1-2\varepsilon+\varepsilon^2 \implies x^2-1 = \varepsilon(\varepsilon-2)$

$$\begin{aligned} \lim_{x\uparrow 1} \varphi^{(n)}(x) &= \lim_{\varepsilon\downarrow 0} \frac{P_n(1-\varepsilon)}{\varepsilon^{2n}(\varepsilon-2)^{2n}} e^{\frac{1}{\varepsilon(\varepsilon-2)}} = \\ &= \lim_{\varepsilon\downarrow 0} \frac{P_n(1)}{\varepsilon^{2n}2^{2n}} e^{-\frac{1}{2\varepsilon}} = \left[\varepsilon = \frac{1}{R}\right] = \lim_{R\to\infty} \frac{P_n(1)}{2^{2n}} R^{2n} e^{-\frac{R}{2}} = 0, \end{aligned}$$

because the power e^{-x} of e decreases stronger than any polynomial x^n.

Note that the complex continuation $\varphi(z)$ is not an analytic function and cannot be expanded as a Taylor series on the entire complex plane $\mathbb{C}$ although it is infinitely often differentiable on the real axis; that is, although $\varphi \in C^\infty(\mathbb{R})$. This can be seen from a uniqueness theorem of complex analysis. Let $B \subseteq \mathbb{C}$ be a domain, and let $z_0 \in B$ the limit of a sequence $\{z_n\} \in B$, $z_n \neq z_0$. Then it can be shown that, if two analytic functions f und g on B coincide in the points z_n, then they coincide on the entire domain B.

Now, take $B = \mathbb{R}$ and the vanishing analytic function f; that is, $f(x) = 0$. $f(x)$ coincides with $\varphi(x)$ only in $\mathbb{R} - M_\varphi$. As a result, φ cannot be analytic. Indeed, $\varphi_{\sigma,\bar{a}}(x)$ diverges at $x = a \pm \sigma$. Hence $\varphi(x)$ cannot be Taylor expanded, and somoothness (that is, being in C^∞) not necessarily implies that its continuation into the complex plain results in an analytic function. (The converse is true though: analyticity implies smoothness.)

7.3.3   Test function class II

Other "good" test functions are [7]

[7] Laurent Schwartz. *Introduction to the Theory of Distributions.* University of Toronto Press, Toronto, 1952. collected and written by Israel Halperin

$$\left\{ \phi_{c,d}(x) \right\}^{\frac{1}{n}} \tag{7.12}$$

obtained by choosing $n \in \mathbb{N} - 0$ and $-\infty \le c < d \le \infty$ and by defining

$$\phi_{c,d}(x) = \begin{cases} e^{-\left(\frac{1}{x-c} + \frac{1}{d-x}\right)} & \text{for } c < x < d, \\ 0 & \text{else.} \end{cases} \tag{7.13}$$

If $\varphi(x)$ is a "good" test function, then

$$x^\alpha P_n(x)\varphi(x) \tag{7.14}$$

with any Polynomial $P_n(x)$, and, in particular, $x^n \varphi(x)$ also is a "good" test function.

7.3.4   Test function class III: Tempered distributions and Fourier transforms

A particular class of "good" test functions – having the property that they vanish "sufficiently fast" for large arguments, but are nonzero at any finite argument – are capable of rendering Fourier transforms of generalized functions. Such generalized functions are called *tempered distributions.*

One example of a test function yielding tempered distribution is the *Gaussian function*

$$\varphi(x) = e^{-\pi x^2}. \tag{7.15}$$

We can multiply the Gaussian function with polynomials (or take its derivatives) and thereby obtain a particular class of test functions inducing tempered distributions.

The Gaussian function is normalized such that

$$\int_{-\infty}^{\infty} \varphi(x)\, dx = \int_{-\infty}^{\infty} e^{-\pi x^2}\, dx$$

$$[\text{variable substitution } x = \frac{t}{\sqrt{\pi}}, dx = \frac{dt}{\sqrt{\pi}}]$$

$$= \int_{-\infty}^{\infty} e^{-\pi \left(\frac{t}{\sqrt{\pi}}\right)^2} d\left(\frac{t}{\sqrt{\pi}}\right) \tag{7.16}$$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-t^2}\, dt$$

$$= \frac{1}{\sqrt{\pi}} \sqrt{\pi} = 1.$$

In this evaluation, we have used the *Gaussian integral*

$$I = \int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}, \tag{7.17}$$

which can be obtained by considering its square and transforming into polar coordinates r, θ; that is,

$$
\begin{aligned}
I^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2} dx \right) \left(\int_{-\infty}^{\infty} e^{-y^2} dy \right) \\
&= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx \, dy \\
&= \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r \, d\theta \, dr \\
&= \int_0^{2\pi} d\theta \int_0^{\infty} e^{-r^2} r \, dr \\
&= 2\pi \int_0^{\infty} e^{-r^2} r \, dr \\
&\quad \left[u = r^2, \frac{du}{dr} = 2r, dr = \frac{du}{2r} \right] \\
&= \pi \int_0^{\infty} e^{-u} du \\
&= \pi \left(-e^{-u} \big|_0^{\infty} \right) \\
&= \pi \left(-e^{-\infty} + e^0 \right) \\
&= \pi.
\end{aligned}
\tag{7.18}
$$

The Gaussian test function (7.15) has the advantage that, as has been shown in (6.27), with a certain kind of definition for the Fourier transform, namely $A = 2\pi$ and $B = 1$ in Eq. (6.19), its functional form does not change under Fourier transforms. More explicitly, as derived in Eqs. (6.27) and (6.28),

$$
\mathscr{F}[\varphi(x)](k) = \widetilde{\varphi}(k) = \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i k x} dx = e^{-\pi k^2}.
\tag{7.19}
$$

Just as for differentiation discussed later it is possible to "shift" or "transfer" the Fourier transformation from the distribution to the test function as follows. Suppose we are interested in the Fourier transform $\mathscr{F}[F]$ of some distribution F. Then, with the convention $A = 2\pi$ and $B = 1$ adopted in Eq. (6.19), we must consider

$$
\begin{aligned}
\langle \mathscr{F}[F], \varphi \rangle &\equiv \mathscr{F}[F][\varphi] = \int_{-\infty}^{\infty} \mathscr{F}[F](x) \varphi(x) dx \\
&= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} F(y) e^{-2\pi i x y} dy \right] \varphi(x) dx \\
&= \int_{-\infty}^{\infty} F(y) \left[\int_{-\infty}^{\infty} \varphi(x) e^{-2\pi i x y} dx \right] dy \\
&= \int_{-\infty}^{\infty} F(y) \mathscr{F}[\varphi](y) dy \\
&= \langle F, \mathscr{F}[\varphi] \rangle \equiv F[\mathscr{F}[\varphi]].
\end{aligned}
\tag{7.20}
$$

in the same way we obtain the *Fourier inversion* for distributions

$$
\langle \mathscr{F}^{-1}[\mathscr{F}[F]], \varphi \rangle = \langle \mathscr{F}[\mathscr{F}^{-1}[F]], \varphi \rangle = \langle F, \varphi \rangle.
\tag{7.21}
$$

Note that, in the case of test functions with compact support – say, $\widehat{\varphi}(x) = 0$ for $|x| > a > 0$ and finite a – if the order of integrals is exchanged, the "new test function"

$$
\mathscr{F}[\widehat{\varphi}](y) = \int_{-\infty}^{\infty} \widehat{\varphi}(x) e^{-2\pi i x y} dx = \int_{-a}^{a} \widehat{\varphi}(x) e^{-2\pi i x y} dx
\tag{7.22}
$$

obtained through a Fourier transform of $\widehat{\varphi}(x)$, does not necessarily inherit a compact support from $\widehat{\varphi}(x)$; in particular, $\mathscr{F}[\widehat{\varphi}](y)$ may not necessarily vanish [i.e. $\mathscr{F}[\widehat{\varphi}](y) = 0$] for $|y| > a > 0$.

Let us, with these conventions, compute the Fourier transform of the tempered Dirac delta distribution. Note that, by the very definition of the Dirac delta distribution,

$$\langle \mathscr{F}[\delta], \varphi \rangle = \langle \delta, \mathscr{F}[\varphi] \rangle = \mathscr{F}[\varphi](0) = \int_{-\infty}^{\infty} e^{-2\pi i x 0} \varphi(x) dx = \int_{-\infty}^{\infty} 1\varphi(x) dx = \langle 1, \varphi \rangle.$$
$$(7.23)$$

Thus we may identify $\mathscr{F}[\delta]$ with 1; that is,

$$\mathscr{F}[\delta] = 1. \tag{7.24}$$

This is an extreme example of an *infinitely concentrated* object whose Fourier transform is *infinitely spread out.*

A very similar calculation renders the tempered distribution associated with the Fourier transform of the shifted Dirac delta distribution

$$\mathscr{F}[\delta_y] = e^{-2\pi i x y}. \tag{7.25}$$

Alas we shall pursue a different, more conventional, approach, sketched in Section 7.5.

7.3.5 Test function class IV: C^∞

If the generalized functions are "sufficiently concentrated" so that they themselves guarantee that the terms $g(\infty)\varphi(\infty)$ as well as $g(-\infty)\varphi(-\infty)$ in Eq. (7.10) to vanish, we may just require the test functions to be infinitely differentiable – and thus in C^∞ – for the sake of making possible a transfer of differentiation. (Indeed, if we are willing to sacrifice even infinite differentiability, we can widen this class of test functions even more.) We may, for instance, employ constant functions such as $\varphi(x) = 1$ as test functions, thus giving meaning to, for instance, $\langle \delta, 1 \rangle = \int_{-\infty}^{\infty} \delta(x) dx$, or $\langle f(x)\delta, 1 \rangle = \langle f(0)\delta, 1 \rangle = f(0) \int_{-\infty}^{\infty} \delta(x) dx$.

7.4 Derivative of distributions

Equipped with "good" test functions which have a finite support and are infinitely often (or at least sufficiently often) differentiable, we can now give meaning to the transferral of differential quotients from the objects entering the integral towards the test function by *partial integration*. First note again that $(uv)' = u'v + uv'$ and thus $\int (uv)' = \int u'v + \int uv'$ and finally $\int u'v = \int (uv)' - \int uv'$. Hence, by identifying u with g, and v with the test function φ, we obtain

$$\begin{aligned} \langle F', \varphi \rangle \equiv F'[\varphi] &= \int_{-\infty}^{\infty} \left(\frac{d}{dx} F(x) \right) \varphi(x) dx \\ &= F(x)\varphi(x)\big|_{x=-\infty}^{\infty} - \int_{-\infty}^{\infty} F(x) \left(\frac{d}{dx} \varphi(x) \right) dx \\ &= -\int_{-\infty}^{\infty} F(x) \left(\frac{d}{dx} \varphi(x) \right) dx \\ &= -F[\varphi'] \equiv -\langle F, \varphi' \rangle. \end{aligned} \tag{7.26}$$

By induction

$$\left\langle \frac{d^n}{dx^n}F, \varphi \right\rangle \equiv \langle F^{(n)}, \varphi \rangle \equiv F^{(n)}\left[\varphi\right] = (-1)^n F\left[\varphi^{(n)}\right] = (-1)^n \langle F, \varphi^{(n)} \rangle. \quad (7.27)$$

For the sake of a further example using *adjoint identities* , swapping products and differentiations forth and back in the F–φ pairing, let us compute $g(x)\delta'(x)$ where $g \in C^\infty$; that is

$$\begin{aligned}
\langle g\delta', \varphi \rangle &= \langle \delta', g\varphi \rangle \\
&= -\langle \delta, (g\varphi)' \rangle \\
&= -\langle \delta, g\varphi' + g'\varphi \rangle \quad (7.28) \\
&= -g(0)\varphi'(0) - g'(0)\varphi(0) \\
&= \langle g(0)\delta' - g'(0)\delta, \varphi \rangle.
\end{aligned}$$

Therefore,

$$g(x)\delta'(x) = g(0)\delta'(x) - g'(0)\delta(x). \quad (7.29)$$

7.5 Fourier transform of distributions

We mention without proof that, if $\{f_x(x)\}$ is a sequence of functions converging, for $n \to \infty$ toward a function f in the functional sense (i.e. *via* integration of f_n and f with "good" test functions), then the Fourier transform $\widetilde{f}$ of f can be defined by [8]

$$\mathscr{F}[f(x)] = \widetilde{f}(k) = \lim_{n\to\infty} \int_{-\infty}^{\infty} f_n(x) e^{-ikx} dx. \quad (7.30)$$

While this represents a method to calculate Fourier transforms of distributions, there are other, more direct ways of obtaining them. These were mentioned earlier. In what follows, we shall enumerate the Fourier transform of some species, mostly by complex analysis.

7.6 Dirac delta function

Historically, the Heaviside step function, which will be discussed later – was first used for the description of electromagnetic pulses. In the days when Dirac developed quantum mechanics there was a need to define "singular scalar products" such as "$\langle x \mid y \rangle = \delta(x - y)$," with some generalization of the Kronecker delta function δ_{ij}, depicted in Fig. 7.2, which is zero whenever $x \neq y$ and "large enough" needle shaped (see Fig. 7.2) to yield unity when integrated over the entire reals; that is, "$\int_{-\infty}^{\infty} \langle x \mid y \rangle dy = \int_{-\infty}^{\infty} \delta(x - y) dy = 1$."

Naturally, such "needle shaped functions" were viewed suspiciouly by many mathematicians at first, but later they embraced these types of functions [9] by developing a theory of *functional analysis* , *generalized functions* or, by another naming, *distributions.*

7.6.1 Delta sequence

One of the first attempts to formalize these objects with "large discontinuities" was in terms of functional limits. Take, for instance, the *delta*

Figure 7.2: Dirac's δ-function as a "needle shaped" generalized function.

sequence which is a sequence of strongly peaked functions for which in some limit the sequences $\{f_n(x-y)\}$ with, for instance,

$$\delta_n(x-y) = \begin{cases} n & \text{for } y - \frac{1}{2n} < x < y + \frac{1}{2n} \\ 0 & \text{else} \end{cases} \tag{7.31}$$

become the delta function $\delta(x-y)$. That is, in the functional sense

$$\lim_{n\to\infty} \delta_n(x-y) = \delta(x-y). \tag{7.32}$$

Note that the area of this particular $\delta_n(x-y)$ above the x-axes is independent of n, since its width is $1/n$ and the height is n.

In an *ad hoc* sense, other delta sequences can be enumerated as follows. They all converge towards the delta function in the sense of linear functionals (i.e. when integrated over a test function).

$$\delta_n(x) = \frac{n}{\sqrt{\pi}} e^{-n^2 x^2}, \tag{7.33}$$

$$= \frac{1}{\pi} \frac{\sin(nx)}{x}, \tag{7.34}$$

$$= = (1 \mp i) \left(\frac{n}{2\pi}\right)^{\frac{1}{2}} e^{\pm i n x^2} \tag{7.35}$$

$$= \frac{1}{\pi x} \frac{e^{inx} - e^{-inx}}{2i}, \tag{7.36}$$

$$= \frac{1}{\pi} \frac{n e^{-x^2}}{1 + n^2 x^2}, \tag{7.37}$$

$$= \frac{1}{2\pi} \int_{-n}^{n} e^{ixt} dt = \frac{1}{2\pi i x} e^{ixt} \Big|_{-n}^{n}, \tag{7.38}$$

$$= \frac{1}{2\pi} \frac{\sin\left[\left(n + \frac{1}{2}\right) x\right]}{\sin\left(\frac{1}{2} x\right)}, \tag{7.39}$$

$$= \frac{1}{\pi} \frac{n}{1 + n^2 x^2}, \tag{7.40}$$

$$= \frac{n}{\pi} \left(\frac{\sin(nx)}{nx}\right)^2. \tag{7.41}$$

Other commonly used limit forms of the δ-function are the Gaussian, Lorentzian, and Dirichlet forms

$$\delta_\epsilon(x) = \frac{1}{\sqrt{\pi}\epsilon} e^{-\frac{x^2}{\epsilon^2}}, \tag{7.42}$$

$$= \frac{1}{\pi} \frac{\epsilon}{x^2 + \epsilon^2} = \frac{1}{2\pi i} \left(\frac{1}{x - i\epsilon} - \frac{1}{x + i\epsilon}\right), \tag{7.43}$$

$$= \frac{1}{\pi} \frac{\sin\left(\frac{x}{\epsilon}\right)}{x}, \tag{7.44}$$

respectively. Note that (7.42) corresponds to (7.33), (7.43) corresponds to (7.40) with $\epsilon = n^{-1}$, and (7.44) corresponds to (7.34). Again, the limit

$$\delta(x) = \lim_{\epsilon \to 0} \delta_\epsilon(x) \tag{7.45}$$

has to be understood in the functional sense; that is, by integration over a test function.

Let us proof that the sequence $\{\delta_n\}$ with

$$\delta_n(x-y) = \begin{cases} n & \text{for } y - \frac{1}{2n} < x < y + \frac{1}{2n} \\ 0 & \text{else} \end{cases}$$
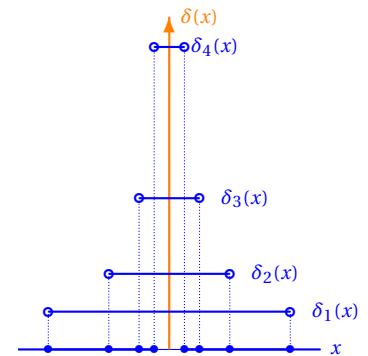

Figure 7.3: Delta sequence approximating Dirac's δ-function as a more and more "needle shaped" generalized function.

defined in Eq. (7.31) and depicted in Fig. 7.3 is a delta sequence; that is, if, for large n, it converges to δ in a functional sense. In order to verify this claim, we have to integrate $\delta_n(x)$ with "good" test functions $\varphi(x)$ and take the limit $n \to \infty$; if the result is $\varphi(0)$, then we can identify $\delta_n(x)$ in this limit with $\delta(x)$ (in the functional sense). Since $\delta_n(x)$ is uniform convergent, we can exchange the limit with the integration; thus

$$\lim_{n\to\infty} \int_{-\infty}^{\infty} \delta_n(x-y)\varphi(x)\,dx$$

[variable transformation:

$$x' = x - y, x = x' + y$$

$$dx' = dx, -\infty \le x' \le \infty]$$

$$= \lim_{n\to\infty} \int_{-\infty}^{\infty} \delta_n(x')\varphi(x'+y)\,dx'$$

$$= \lim_{n\to\infty} \int_{-\frac{1}{2n}}^{\frac{1}{2n}} n\varphi(x'+y)\,dx'$$

[variable transformation:

$$u = 2nx', x' = \frac{u}{2n},$$

$$du = 2n\,dx', -1 \le u \le 1]$$

$$= \lim_{n\to\infty} \int_{-1}^{1} n\varphi(\frac{u}{2n}+y)\frac{du}{2n}$$

$$= \lim_{n\to\infty} \frac{1}{2} \int_{-1}^{1} \varphi(\frac{u}{2n}+y)\,du$$

$$= \frac{1}{2} \int_{-1}^{1} \lim_{n\to\infty} \varphi(\frac{u}{2n}+y)\,du$$

$$= \frac{1}{2}\varphi(y) \int_{-1}^{1} du$$

$$= \varphi(y).$$

(7.46)

Hence, in the functional sense, this limit yields the shifted δ-function δ_y. Thus we obtain

$$\lim_{n\to\infty} \delta_n[\varphi] = \delta_y[\varphi] = \varphi(y).$$

7.6.2 $\delta[\varphi]$ distribution

The distribution (linear functional) associated with the δ function can be defined by mapping any test function into a scalar as follows:

$$\int_{-\infty}^{\infty} \delta(x-y)\varphi(x)\,dx = \varphi(y). \tag{7.47}$$

A common way of expressing this *delta function distribution* is by writing

$$\delta(x-y) \longleftrightarrow \langle \delta_y, \varphi \rangle \equiv \langle \delta_y | \varphi \rangle \equiv \delta_y[\varphi] \overset{\text{def}}{=} \varphi(y). \tag{7.48}$$

For $y = 0$, we just obtain

$$\delta(x) \longleftrightarrow \langle \delta, \varphi \rangle \equiv \langle \delta | \varphi \rangle \equiv \delta[\varphi] \overset{\text{def}}{=} \delta_0[\varphi] = \varphi(0). \tag{7.49}$$

7.6.3   Useful formulæ involving δ

The following formulae are sometimes enumerated without proofs.

$$\delta(x) = \delta(-x) \tag{7.50}$$

For a proof, note that $\varphi(x)\delta(-x) = \varphi(0)\delta(-x)$, and that, in particular, with the substitution $x \to -x$,

$$\int_{-\infty}^{\infty} \delta(-x)\varphi(x)dx = \varphi(0)\int_{\infty}^{-\infty} \delta(-(-x))d(-x)$$
$$= -\varphi(0)\int_{\infty}^{-\infty} \delta(x)dx = \varphi(0)\int_{-\infty}^{\infty} \delta(x)dx. \tag{7.51}$$

$$\delta(x) = \lim_{\epsilon \to 0} \frac{H(x+\epsilon) - H(x)}{\epsilon} = \frac{d}{dx}H(x) \tag{7.52}$$

$$f(x)\delta(x - x_0) = f(x_0)\delta(x - x_0) \tag{7.53}$$

This results from a direct application of Eq. (7.4); that is,

$$f(x)\delta[\varphi] = \delta\left[f\varphi\right] = f(0)\varphi(0) = f(0)\delta[\varphi], \tag{7.54}$$

and

$$f(x)\delta_{x_0}[\varphi] = \delta_{x_0}\left[f\varphi\right] = f(x_0)\varphi(x_0) = f(x_0)\delta_{x_0}[\varphi]. \tag{7.55}$$

For a more explicit direct proof, note that

$$\int_{-\infty}^{\infty} f(x)\delta(x - x_0)\varphi(x)dx = \int_{-\infty}^{\infty} \delta(x - x_0)(f(x)\varphi(x))dx = f(x_0)\varphi(x_0), \tag{7.56}$$

and hence $f\delta_{x_0}[\varphi] = f(x_0)\delta_{x_0}[\varphi]$.

For the δ distribution with its "extreme concentration" at the origin, a "nonconcentrated test function" suffices; in particular, a constant test function such as $\varphi(x) = 1$ is fine. This is the reason why test functions need not show up explicitly in expressions, and, in particular, integrals, containing δ. Because, say, for suitable functions $g(x)$ "well behaved" at the origin,

$$g(x)\delta(x - y)[1] = g(x)\delta(x - y)[\varphi = 1]$$
$$= \int_{-\infty}^{\infty} g(x)\delta(x - y)[\varphi = 1]\,dx = \int_{-\infty}^{\infty} g(x)\delta(x - y)\,1\,dx = g(y). \tag{7.57}$$

$$x\delta(x) = 0 \tag{7.58}$$

For a proof consider

$$x\delta[\varphi] = \delta\left[x\varphi\right] = 0\varphi(0) = 0. \tag{7.59}$$

For $a \neq 0$,

$$\delta(ax) = \frac{1}{|a|}\delta(x), \tag{7.60}$$

and, more generally,

$$\delta(a(x - x_0)) = \frac{1}{|a|}\delta(x - x_0) \tag{7.61}$$

For the sake of a proof, consider the case $a > 0$ as well as $x_0 = 0$ first:

$$\int_{-\infty}^{\infty} \delta(ax)\varphi(x)dx$$

[variable substitution $y = ax, x = \dfrac{y}{a}, dx = \dfrac{1}{a}dy$]

$$= \frac{1}{a}\int_{-\infty}^{\infty} \delta(y)\varphi\left(\frac{y}{a}\right)dy \qquad (7.62)$$

$$= \frac{1}{a}\varphi(0);$$

and, second, the case $a < 0$:

$$\int_{-\infty}^{\infty} \delta(ax)\varphi(x)dx$$

[variable substitution $y = ax, x = \dfrac{y}{a}, dx = \dfrac{1}{a}dy$]

$$= \frac{1}{a}\int_{\infty}^{-\infty} \delta(y)\varphi\left(\frac{y}{a}\right)dy \qquad (7.63)$$

$$= -\frac{1}{a}\int_{-\infty}^{\infty} \delta(y)\varphi\left(\frac{y}{a}\right)dy$$

$$= -\frac{1}{a}\varphi(0).$$

In the case of $x_0 \neq 0$ and $\pm a > 0$, we obtain

$$\int_{-\infty}^{\infty} \delta(a(x - x_0))\varphi(x)dx$$

[variable substitution $y = a(x - x_0), x = \dfrac{y}{a} + x_0, dx = \dfrac{1}{a}dy$]

$$= \pm\frac{1}{a}\int_{-\infty}^{\infty} \delta(y)\varphi\left(\frac{y}{a} + x_0\right)dy \qquad (7.64)$$

$$= \frac{1}{|a|}\varphi(x_0).$$

If there exists a simple singularity x_0 of $f(x)$ in the integration interval, then

$$\delta(f(x)) = \frac{1}{|f'(x_0)|}\delta(x - x_0). \qquad (7.65)$$

More generally, if f has only simple roots and f' is nonzero there,

$$\delta(f(x)) = \sum_{x_i} \frac{\delta(x - x_i)}{|f'(x_i)|} \qquad (7.66)$$

where the sum extends over all simple roots x_i in the integration interval. In particular,

$$\delta(x^2 - x_0^2) = \frac{1}{2|x_0|}[\delta(x - x_0) + \delta(x + x_0)] \qquad (7.67)$$

For a proof, note that, since f has only simple roots , it can be expanded around these roots by

$$f(x) \approx \underbrace{f(x_0)}_{=0} + (x - x_0)f'(x_0) = (x - x_0)f'(x_0)$$

An example is a polynomial of degree k of the form $f = A\prod_{i=1}^{k}(x - x_i)$; with mutually distinct x_i, $1 \leq i \leq k$.

with nonzero $f'(x_0) \in \mathbb{R}$. By identifying $f'(x_0)$ with a in Eq. (7.60) we obtain Eq. (7.66).

The simplest nontrivial case is $f(x) = a + bx = b\left(\frac{a}{b} + x\right)$, for which $x_0 = -\frac{a}{b}$ and $f'\left(x_0 = \frac{a}{b}\right) = b$.

$$\delta'(f(x)) = \sum_{i=0}^{N} \frac{f''(x_i)}{|f'(x_i)|^3}\delta(x - x_i) + \sum_{i=0}^{N} \frac{f'(x_i)}{|f'(x_i)|^3}\delta'(x - x_i) \qquad (7.68)$$

$$|x|\delta(x^2) = \delta(x) \tag{7.69}$$

$$-x\delta'(x) = \delta(x), \tag{7.70}$$

which is a direct consequence of Eq. (7.29). More explicitly, we can use partial integration and obtain

$$
\begin{aligned}
&\int_{-\infty}^{\infty} -x\delta'(x)\varphi(x)dx \\
=& -x\delta(x)|_{-\infty}^{\infty} + \int_{-\infty}^{\infty} \delta(x)\frac{d}{dx}\left(x\varphi(x)\right)dx \\
=& \int_{-\infty}^{\infty} \delta(x)x\varphi'(x)dx + \int_{-\infty}^{\infty} \delta(x)\varphi(x)dx \\
&= 0\varphi'(0) + \varphi(0) = \varphi(0).
\end{aligned}
\tag{7.71}
$$

$$\delta^{(n)}(-x) = (-1)^n \delta^{(n)}(x), \tag{7.72}$$

where the index $^{(n)}$ denotes n-fold differentiation, can be proven by [recall that, by the chain rule of differentiation, $\frac{d}{dx}\varphi(-x) = -\varphi'(-x)$]

$$
\begin{aligned}
&\int_{-\infty}^{\infty} \delta^{(n)}(-x)\varphi(x)dx \\
&[\text{variable substitution } x \to -x] \\
&= -\int_{\infty}^{-\infty} \delta^{(n)}(x)\varphi(-x)dx \\
&= \int_{-\infty}^{\infty} \delta^{(n)}(x)\varphi(-x)dx \\
&= (-1)^n \int_{-\infty}^{\infty} \delta(x)\left[\frac{d^n}{dx^n}\varphi(-x)\right]dx \\
&= (-1)^n \int_{-\infty}^{\infty} \delta(x)\left[(-1)^n\varphi^{(n)}(-x)\right]dx \\
&= \int_{\infty}^{-\infty} \delta(x)\varphi^{(n)}(-x)dx \\
&[\text{variable substitution } x \to -x] \\
&= -\int_{\infty}^{-\infty} \delta(-x)\varphi^{(n)}(x)dx \\
&= \int_{-\infty}^{\infty} \delta(x)\varphi^{(n)}(x)dx \\
&= (-1)^n \int_{\infty}^{-\infty} \delta^{(n)}(x)\varphi(x)dx.
\end{aligned}
\tag{7.73}
$$

Because of an additional factor $(-1)^n$ from the chain rule, in particular, from the n-fold "inner" differentiation of $-x$, follows that

$$\frac{d^n}{dx^n}\delta(-x) = (-1)^n \delta^{(n)}(-x) = \delta^{(n)}(x). \tag{7.74}$$

$$x^{m+1}\delta^{(m)}(x) = 0, \tag{7.75}$$

where the index $^{(m)}$ denotes m-fold differentiation;

$$x^2\delta'(x) = 0, \tag{7.76}$$

which is a consequence of Eq. (7.29).

More generally,

$$x^n \delta^{(m)}(x)[\varphi = 1] = \int_{-\infty}^{\infty} x^n \delta^{(m)}(x) dx = (-1)^n n! \delta_{nm}, \qquad (7.77)$$

which can be demonstrated by considering

$$
\begin{aligned}
\langle x^n \delta^{(m)} | 1 \rangle &= \langle \delta^{(m)} | x^n \rangle \\
&= (-1)^n \langle \delta | \frac{d^m}{dx^m} x^n \rangle \\
&= (-1)^n n! \delta_{nm} \underbrace{\langle \delta | 1 \rangle}_{1} \\
&= (-1)^n n! \delta_{nm}.
\end{aligned}
\qquad (7.78)
$$

$$\frac{d^2}{dx^2}[xH(x)] = \frac{d}{dx}[H(x) + \underbrace{x\delta(x)}_{0}] = \frac{d}{dx} H(x) = \delta(x) \qquad (7.79)$$

If $\delta^3(\vec{r}) = \delta(x)\delta(y)\delta(r)$ with $\vec{r} = (x, y, z)$, then

$$\delta^3(\vec{r}) = \delta(x)\delta(y)\delta(z) = -\frac{1}{4\pi} \Delta \frac{1}{r} \qquad (7.80)$$

$$\delta^3(\vec{r}) = -\frac{1}{4\pi}(\Delta + k^2) \frac{e^{ikr}}{r} \qquad (7.81)$$

$$\delta^3(\vec{r}) = -\frac{1}{4\pi}(\Delta + k^2) \frac{\cos kr}{r} \qquad (7.82)$$

In quantum field theory, phase space integrals of the form

$$\frac{1}{2E} = \int dp^0 \, H(p^0)\delta(p^2 - m^2) \qquad (7.83)$$

if $E = (\vec{p}^2 + m^2)^{(1/2)}$ are exploited.

7.6.4   Fourier transform of δ

The Fourier transform of the δ-function can be obtained straightforwardly by insertion into Eq. (6.19); that is, with $A = B = 1$

$$
\begin{aligned}
\mathscr{F}[\delta(x)] = \widetilde{\delta}(k) &= \int_{-\infty}^{\infty} \delta(x) e^{-ikx} dx \\
&= e^{-i0k} \int_{-\infty}^{\infty} \delta(x) dx \\
&= 1, \text{ and thus} \\
\mathscr{F}^{-1}[\widetilde{\delta}(k)] = \mathscr{F}^{-1}[1] &= \delta(x) \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} [\cos(kx) + i\sin(kx)] dk \\
&= \frac{1}{\pi} \int_{0}^{\infty} \cos(kx) dk + \frac{i}{2\pi} \int_{-\infty}^{\infty} \sin(kx) dk \\
&= \frac{1}{\pi} \int_{0}^{\infty} \cos(kx) dk.
\end{aligned}
\qquad (7.84)
$$

That is, the Fourier transform of the δ-function is just a constant. δ-spiked signals carry all frequencies in them. Note also that $\mathscr{F}[\delta(x-y)] = e^{iky}\mathscr{F}[\delta(x)]$.

From Eq. (7.84) we can compute

$$
\begin{aligned}
\mathscr{F}[1] = \widetilde{1}(k) &= \int_{-\infty}^{\infty} e^{-ikx} dx \\
&[\text{variable substitution } x \to -x] \\
&= \int_{+\infty}^{-\infty} e^{-ik(-x)} d(-x) \\
&= -\int_{+\infty}^{-\infty} e^{ikx} dx \\
&= \int_{-\infty}^{+\infty} e^{ikx} dx \\
&= 2\pi\delta(k).
\end{aligned}
\tag{7.85}
$$

7.6.5 Eigenfunction expansion of δ

The δ-function can be expressed in terms of, or "decomposed" into, various *eigenfunction expansions*. We mention without proof [10] that, for $0 < x, x_0 < L$, two such expansions in terms of trigonometric functions are

[10] Dean G. Duffy. *Green's Functions with Applications.* Chapman and Hall/CRC, Boca Raton, 2001

$$
\begin{aligned}
\delta(x - x_0) &= \frac{2}{L} \sum_{k=1}^{\infty} \sin\left(\frac{\pi k x_0}{L}\right) \sin\left(\frac{\pi k x}{L}\right) \\
&= \frac{1}{L} + \frac{2}{L} \sum_{k=1}^{\infty} \cos\left(\frac{\pi k x_0}{L}\right) \cos\left(\frac{\pi k x}{L}\right).
\end{aligned}
\tag{7.86}
$$

This "decomposition of unity" is analoguous to the expansion of the identity in terms of orthogonal projectors $\mathbf{E}_i$ (for one-dimensional projectors, $\mathbf{E}_i = |i\rangle\langle i|$) encountered in the spectral theorem 1.28.1.

Other decomposions are in terms of orthonormal (Legendre) polynomials (cf. Sect. 11.6 on page 223), or other functions of mathematical physics discussed later.

7.6.6 Delta function expansion

Just like "slowly varying" functions can be expanded into a Taylor series in terms of the power functions x^n, highly localized functions can be expanded in terms of derivatives of the δ-function in the form [11]

[11] Ismo V. Lindell. Delta function expansions, complex delta functions and the steepest descent method. *American Journal of Physics*, 61(5):438–442, 1993. DOI: 10.1119/1.17238. URL http://doi.org/10.1119/1.17238

$$
f(x) \sim f_0\delta(x) + f_1\delta'(x) + f_2\delta''(x) + \cdots + f_n\delta^{(n)}(x) + \cdots = \sum_{k=1}^{\infty} f_k\delta^{(k)}(x),
$$
$$
\text{with } f_k = \frac{(-1)^k}{k!} \int_{-\infty}^{\infty} f(y) y^k \, dy.
\tag{7.87}
$$

The sign "$\sim$" denotes the functional character of this "equation" (7.87).

The delta expansion (7.87) can be proven by considering a smooth

function $g(x)$, and integrating over its expansion; that is,

$$\int_{-\infty}^{\infty} f(x)\varphi(x)dx$$

$$= \int_{-\infty}^{\infty} \left[f_0\delta(x) + f_1\delta'(x) + f_2\delta''(x) + \cdots + (-1)^n f_n \delta^{(n)}(x) + \cdots \right] \varphi(x)dx$$

$$= f_0\varphi(0) - f_1\varphi'(0) + f_2\varphi''(0) + \cdots + (-1)^n f_n \varphi^{(n)}(0) + \cdots,$$

$$(7.88)$$

and comparing the coefficients in (7.88) with the coefficients of the
Taylor series expansion of φ at $x = 0$

$$\int_{-\infty}^{\infty} \varphi(x)f(x) = \int_{-\infty}^{\infty} \left[\varphi(0) + x\varphi'(0) + \cdots + \frac{x^n}{n!}\varphi^{(n)}(0) + \cdots \right] f(x)dx$$

$$= \varphi(0)\int_{-\infty}^{\infty} f(x)dx + \varphi'(0)\int_{-\infty}^{\infty} xf(x)dx + \cdots + \varphi^{(n)}(0)\int_{-\infty}^{\infty} \frac{x^n}{n!}f(x)dx + \cdots.$$

$$(7.89)$$

7.7   Cauchy principal value

7.7.1   Definition

The *(Cauchy) principal value* $\mathscr{P}$ (sometimes also denoted by p.v.) is a
value associated with a (divergent) integral as follows:

$$\mathscr{P}\int_a^b f(x)dx = \lim_{\varepsilon\to 0^+} \left[\int_a^{c-\varepsilon} f(x)dx + \int_{c+\varepsilon}^b f(x)dx \right]$$

$$= \lim_{\varepsilon\to 0^+} \int_{[a,c-\varepsilon]\cup[c+\varepsilon,b]} f(x)dx,$$

$$(7.90)$$

if c is the "location" of a singularity of $f(x)$.

For example, the integral $\int_{-1}^{1} \frac{dx}{x}$ diverges, but

$$\mathscr{P}\int_{-1}^{1} \frac{dx}{x} = \lim_{\varepsilon\to 0^+} \left[\int_{-1}^{-\varepsilon} \frac{dx}{x} + \int_{+\varepsilon}^{1} \frac{dx}{x} \right]$$

[variable substitution $x \to -x$ in the first integral]

$$= \lim_{\varepsilon\to 0^+} \left[\int_{+1}^{+\varepsilon} \frac{dx}{x} + \int_{+\varepsilon}^{1} \frac{dx}{x} \right]$$

$$(7.91)$$

$$= \lim_{\varepsilon\to 0^+} \left[\log\varepsilon - \log 1 + \log 1 - \log\varepsilon \right] = 0.$$

7.7.2   Principle value and pole function $\frac{1}{x}$ distribution

The "standalone function" $\frac{1}{x}$ does not define a distribution since it is
not integrable in the vicinity of $x = 0$. This issue can be "alleviated"
or "circumvented" by considering the principle value $\mathscr{P}\frac{1}{x}$. In this way
the principle value can be transferred to the context of distributions by

defining a *principal value distribution* in a functional sense by

$$\mathscr{P}\left(\frac{1}{x}\right)[\varphi] = \lim_{\varepsilon \to 0^+} \int_{|x|>\varepsilon} \frac{1}{x}\varphi(x)\,dx$$

$$= \lim_{\varepsilon \to 0^+}\left[\int_{-\infty}^{-\varepsilon} \frac{1}{x}\varphi(x)\,dx + \int_{+\varepsilon}^{\infty} \frac{1}{x}\varphi(x)\,dx\right]$$

[variable substitution $x \to -x$ in the first integral]

$$= \lim_{\varepsilon \to 0^+}\left[\int_{+\infty}^{+\varepsilon} \frac{1}{x}\varphi(-x)\,dx + \int_{+\varepsilon}^{\infty} \frac{1}{x}\varphi(x)\,dx\right] \tag{7.92}$$

$$= \lim_{\varepsilon \to 0^+}\left[-\int_{+\varepsilon}^{\infty} \frac{1}{x}\varphi(-x)\,dx + \int_{+\varepsilon}^{\infty} \frac{1}{x}\varphi(x)\,dx\right]$$

$$= \lim_{\varepsilon \to 0^+}\int_{\varepsilon}^{+\infty} \frac{\varphi(x) - \varphi(-x)}{x}\,dx$$

$$= \int_{0}^{+\infty} \frac{\varphi(x) - \varphi(-x)}{x}\,dx.$$

In the functional sense, $\frac{1}{x}[\varphi]$ can be interpreted as a principal value. That is,

$$\frac{1}{x}[\varphi] = \int_{-\infty}^{\infty} \frac{1}{x}\varphi(x)\,dx$$

$$= \int_{-\infty}^{0} \frac{1}{x}\varphi(x)\,dx + \int_{0}^{\infty} \frac{1}{x}\varphi(x)\,dx$$

[variable substitution $x \to -x, dx \to -dx$ in the first integral]

$$= \int_{+\infty}^{0} \frac{1}{(-x)}\varphi(-x)\,d(-x) + \int_{0}^{\infty} \frac{1}{x}\varphi(x)\,dx$$

$$= \int_{+\infty}^{0} \frac{1}{x}\varphi(-x)\,dx + \int_{0}^{\infty} \frac{1}{x}\varphi(x)\,dx \tag{7.93}$$

$$= -\int_{0}^{\infty} \frac{1}{x}\varphi(-x)\,dx + \int_{0}^{\infty} \frac{1}{x}\varphi(x)\,dx$$

$$= \int_{0}^{\infty} \frac{\varphi(x) - \varphi(-x)}{x}\,dx$$

$$= \mathscr{P}\left(\frac{1}{x}\right)[\varphi],$$

where in the last step the principle value distribution (7.92) has been used.

7.8   *Absolute value distribution*

The distribution associated with the absolute value $|x|$ is defined by

$$|x|[\varphi] = \int_{-\infty}^{\infty} |x|\varphi(x)\,dx. \tag{7.94}$$

$|x| [\varphi]$ can be evaluated and represented as follows:

$$|x| [\varphi] = \int_{-\infty}^{\infty} |x| \varphi(x) dx$$

$$= \int_{-\infty}^{0} (-x)\varphi(x) dx + \int_{0}^{\infty} x\varphi(x) dx$$

$$= -\int_{-\infty}^{0} x\varphi(x) dx + \int_{0}^{\infty} x\varphi(x) dx$$

[variable substitution $x \to -x, dx \to -dx$ in the first integral]   (7.95)

$$= -\int_{+\infty}^{0} x\varphi(-x) dx + \int_{0}^{\infty} x\varphi(x) dx$$

$$= \int_{0}^{\infty} x\varphi(-x) dx + \int_{0}^{\infty} x\varphi(x) dx$$

$$= \int_{0}^{\infty} x \left[\varphi(x) + \varphi(-x)\right] dx.$$

7.9   Logarithm distribution

7.9.1   Definition

Let, for $x \neq 0$,

$$\log|x| [\varphi] = \int_{-\infty}^{\infty} \log|x| \varphi(x) dx$$

$$= \int_{-\infty}^{0} \log(-x)\varphi(x) dx + \int_{0}^{\infty} \log x\varphi(x) dx$$

[variable substitution $x \to -x, dx \to -dx$ in the first integral]

$$= \int_{+\infty}^{0} \log(-(-x))\varphi(-x) d(-x) + \int_{0}^{\infty} \log x\varphi(x) dx \quad (7.96)$$

$$= -\int_{+\infty}^{0} \log x\varphi(-x) dx + \int_{0}^{\infty} \log x\varphi(x) dx$$

$$= \int_{0}^{\infty} \log x\varphi(-x) dx + \int_{0}^{\infty} \log x\varphi(x) dx$$

$$= \int_{0}^{\infty} \log x \left[\varphi(x) + \varphi(-x)\right] dx.$$

7.9.2   Connection with pole function

Note that

$$\mathscr{P}\left(\frac{1}{x}\right)[\varphi] = \frac{d}{dx} \log|x| [\varphi], \qquad (7.97)$$

and thus for the principal value of a pole of degree n

$$\mathscr{P}\left(\frac{1}{x^n}\right)[\varphi] = \frac{(-1)^{n-1}}{(n-1)!} \frac{d^n}{dx^n} \log|x| [\varphi]. \qquad (7.98)$$

For a proof of Eq. (7.97) consider the functional derivative $\log'|x|[\varphi]$ of

$\log|x|[\varphi]$ by insertion into Eq. (7.96); that is

$$
\begin{aligned}
\log'|x|[\varphi] &= \int_{-\infty}^{0} \frac{d\log(-x)}{dx}\varphi(x)\,dx + \int_{0}^{\infty} \frac{d\log x}{dx}\varphi(x)\,dx \\
&= \int_{-\infty}^{0}\left(-\frac{1}{(-x)}\right)\varphi(x)\,dx + \int_{0}^{\infty}\frac{1}{x}\varphi(x)\,dx \\
&= \int_{-\infty}^{0}\frac{1}{x}\varphi(x)\,dx + \int_{0}^{\infty}\frac{1}{x}\varphi(x)\,dx
\end{aligned}
$$

[variable substitution $x \to -x, dx \to -dx$ in the first integral]

$$
\begin{aligned}
&= \int_{+\infty}^{0}\frac{1}{(-x)}\varphi(-x)\,d(-x) + \int_{0}^{\infty}\frac{1}{x}\varphi(x)\,dx \quad (7.99) \\
&= \int_{+\infty}^{0}\frac{1}{x}\varphi(-x)\,dx + \int_{0}^{\infty}\frac{1}{x}\varphi(x)\,dx \\
&= -\int_{0}^{\infty}\frac{1}{x}\varphi(-x)\,dx + \int_{0}^{\infty}\frac{1}{x}\varphi(x)\,dx \\
&= \int_{0}^{\infty}\frac{\varphi(x)-\varphi(-x)}{x}\,dx \\
&= \mathscr{P}\left(\frac{1}{x}\right)[\varphi].
\end{aligned}
$$

The more general Eq. (7.98) follows by direct differentiation.

7.10   Pole function $\frac{1}{x^n}$ distribution

For $n \geq 2$, the integral over $\frac{1}{x^n}$ is undefined even if we take the principal value. Hence the direct route to an evaluation is blocked, and we have to take an indirect approch *via* derivatives of $\frac{1}{x}$ [12]. Thus, let

[12] Thomas Sommer. Verallgemeinerte Funktionen. unpublished manuscript, 2012

$$
\begin{aligned}
\frac{1}{x^2}[\varphi] &= -\frac{d}{dx}\frac{1}{x}[\varphi] \\
= \frac{1}{x}[\varphi'] &= \int_{0}^{\infty}\frac{1}{x}\left[\varphi'(x)-\varphi'(-x)\right]dx \quad (7.100) \\
&= \mathscr{P}\left(\frac{1}{x}\right)[\varphi'].
\end{aligned}
$$

Also,

$$
\begin{aligned}
\frac{1}{x^3}[\varphi] &= -\frac{1}{2}\frac{d}{dx}\frac{1}{x^2}[\varphi] = \frac{1}{2}\frac{1}{x^2}[\varphi'] = \frac{1}{2x}[\varphi''] \\
&= \frac{1}{2}\int_{0}^{\infty}\frac{1}{x}\left[\varphi''(x)-\varphi''(-x)\right]dx \quad (7.101) \\
&= \frac{1}{2}\mathscr{P}\left(\frac{1}{x}\right)[\varphi''].
\end{aligned}
$$

More generally, for $n > 1$, by induction, using (7.100) as induction

basis,

$$\frac{1}{x^n}[\varphi]$$

$$= -\frac{1}{n-1}\frac{d}{dx}\frac{1}{x^{n-1}}[\varphi] = \frac{1}{n-1}\frac{1}{x^{n-1}}[\varphi']$$

$$= -\left(\frac{1}{n-1}\right)\left(\frac{1}{n-2}\right)\frac{d}{dx}\frac{1}{x^{n-2}}[\varphi'] = \frac{1}{(n-1)(n-2)}\frac{1}{x^{n-2}}[\varphi'']$$

$$= \cdots = \frac{1}{(n-1)!}\frac{1}{x}[\varphi^{(n-1)}] \tag{7.102}$$

$$= \frac{1}{(n-1)!}\int_0^\infty \frac{1}{x}\left[\varphi^{(n-1)}(x) - \varphi^{(n-1)}(-x)\right]dx$$

$$= \frac{1}{(n-1)!}\mathscr{P}\left(\frac{1}{x}\right)[\varphi^{(n-1)}].$$

7.11   Pole function $\frac{1}{x\pm i\alpha}$ distribution

We are interested in the limit $\alpha \to 0$ of $\frac{1}{x+i\alpha}$. Let $\alpha > 0$. Then,

$$\frac{1}{x+i\alpha}[\varphi] = \int_{-\infty}^\infty \frac{1}{x+i\alpha}\varphi(x)dx$$

$$= \int_{-\infty}^\infty \frac{x-i\alpha}{(x+i\alpha)(x-i\alpha)}\varphi(x)dx$$

$$= \int_{-\infty}^\infty \frac{x-i\alpha}{x^2+\alpha^2}\varphi(x)dx \tag{7.103}$$

$$= \int_{-\infty}^\infty \frac{x}{x^2+\alpha^2}\varphi(x)dx - i\alpha\int_{-\infty}^\infty \frac{1}{x^2+\alpha^2}\varphi(x)dx.$$

Let us treat the two summands of (7.103) separately. (i) Upon variable substitution $x = \alpha y$, $dx = \alpha dy$ in the second integral in (7.103) we obtain

$$\alpha\int_{-\infty}^\infty \frac{1}{x^2+\alpha^2}\varphi(x)dx = \alpha\int_{-\infty}^\infty \frac{1}{\alpha^2 y^2+\alpha^2}\varphi(\alpha y)\alpha dy$$

$$= \alpha^2\int_{-\infty}^\infty \frac{1}{\alpha^2(y^2+1)}\varphi(\alpha y)dy \tag{7.104}$$

$$= \int_{-\infty}^\infty \frac{1}{y^2+1}\varphi(\alpha y)dy$$

In the limit $\alpha \to 0$, this is

$$\lim_{\alpha\to 0}\int_{-\infty}^\infty \frac{1}{y^2+1}\varphi(\alpha y)dy = \varphi(0)\int_{-\infty}^\infty \frac{1}{y^2+1}dy$$

$$= \varphi(0)\left(\arctan y\right)\big|_{y=-\infty}^{y=\infty} \tag{7.105}$$

$$= \pi\varphi(0) = \pi\delta[\varphi].$$

(ii) The first integral in (7.103) is

$$\int_{-\infty}^\infty \frac{x}{x^2+\alpha^2}\varphi(x)dx$$

$$= \int_{-\infty}^0 \frac{x}{x^2+\alpha^2}\varphi(x)dx + \int_0^\infty \frac{x}{x^2+\alpha^2}\varphi(x)dx$$

$$= \int_{+\infty}^0 \frac{-x}{(-x)^2+\alpha^2}\varphi(-x)d(-x) + \int_0^\infty \frac{x}{x^2+\alpha^2}\varphi(x)dx \tag{7.106}$$

$$= -\int_0^\infty \frac{x}{x^2+\alpha^2}\varphi(-x)dx + \int_0^\infty \frac{x}{x^2+\alpha^2}\varphi(x)dx$$

$$= \int_0^\infty \frac{x}{x^2+\alpha^2}\left[\varphi(x) - \varphi(-x)\right]dx.$$

In the limit $\alpha \to 0$, this becomes

$$\lim_{\alpha \to 0} \int_0^\infty \frac{x}{x^2 + \alpha^2} \left[\varphi(x) - \varphi(-x) \right] dx = \int_0^\infty \frac{\varphi(x) - \varphi(-x)}{x} dx$$
$$= \mathscr{P}\left(\frac{1}{x}\right)[\varphi], \tag{7.107}$$

where in the last step the principle value distribution (7.92) has been used.

Putting all parts together, we obtain

$$\frac{1}{x + i0^+}[\varphi] = \lim_{\alpha \to 0} \frac{1}{x + i\alpha}[\varphi] = \mathscr{P}\left(\frac{1}{x}\right)[\varphi] - i\pi\delta[\varphi] = \left\{ \mathscr{P}\left(\frac{1}{x}\right) - i\pi\delta \right\}[\varphi]. \tag{7.108}$$

A very similar calculation yields

$$\frac{1}{x - i0^+}[\varphi] = \lim_{\alpha \to 0} \frac{1}{x - i\alpha}[\varphi] = \mathscr{P}\left(\frac{1}{x}\right)[\varphi] + i\pi\delta[\varphi] = \left\{ \mathscr{P}\left(\frac{1}{x}\right) + i\pi\delta \right\}[\varphi]. \tag{7.109}$$

These equations (7.108) and (7.109) are often called the *Sokhotsky formula*, also known as the *Plemelj formula*, or the *Plemelj-Sokhotsky formula*.

7.12   Heaviside step function

7.12.1   Ambiguities in definition

Let us now turn to some very common generalized functions; in particular to Heaviside's electromagnetic *infinite pulse function*. One of the possible definitions of the Heaviside step function $H(x)$, and maybe the most common one – they differ by the difference of the value(s) of $H(0)$ at the origin $x = 0$, a difference which is irrelevant measure theoretically for "good" functions since it is only about an isolated point – is

$$H(x - x_0) = \begin{cases} 1 & \text{for } x \geq x_0 \\ 0 & \text{for } x < x_0 \end{cases} \tag{7.110}$$

The function is plotted in Fig. 7.4.

In the spirit of the above definition, it might have been more appropriate to define $H(0) = \frac{1}{2}$; that is,

$$H(x - x_0) = \begin{cases} 1 & \text{for } x > x_0 \\ \frac{1}{2} & \text{for } x = x_0 \\ 0 & \text{for } x < x_0 \end{cases} \tag{7.111}$$

and, since this affects only an isolated point at $x = 0$, we may happily do so if we prefer.

It is also very common to define the Heaviside step function as the *antiderivative of the δ function*; likewise the delta function is the derivative of the Heaviside step function; that is,

$$H(x - x_0) = \int_{-\infty}^{x - x_0} \delta(t)\, dt,$$
$$\frac{d}{dx} H(x - x_0) = \delta(x - x_0). \tag{7.112}$$

$H(x)$
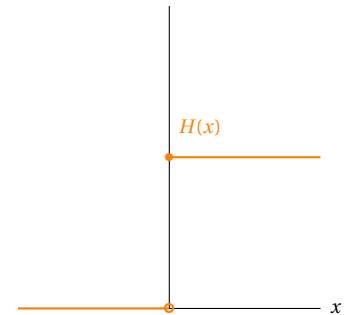
Figure 7.4: Plot of the Heaviside step function $H(x)$.

The latter equation can, in the functional sense; that is, by integration over a test function, be proven by

$$
\begin{aligned}
\langle H', \varphi \rangle &= -\langle H, \varphi' \rangle \\
&= -\int_{-\infty}^{\infty} H(x)\varphi'(x)\,dx \\
&= -\int_{0}^{\infty} \varphi'(x)\,dx \\
&= -\varphi(x)\Big|_{x=0}^{x=\infty} \\
&= \underbrace{-\varphi(\infty)}_{=0} + \varphi(0) \\
&= \langle \delta, \varphi \rangle
\end{aligned}
\tag{7.113}
$$

for all test functions $\varphi(x)$. Hence we can – in the functional sense – identify δ with H'. More explicitly, through integration by parts, we obtain

$$
\begin{aligned}
&\int_{-\infty}^{\infty} \left[\frac{d}{dx} H(x-x_0) \right] \varphi(x)\,dx \\
&= H(x-x_0)\varphi(x)\Big|_{-\infty}^{\infty} - \int_{-\infty}^{\infty} H(x-x_0) \left[\frac{d}{dx}\varphi(x) \right] dx \\
&= \underbrace{H(\infty)\varphi(\infty)}_{\varphi(\infty)=0} - \underbrace{H(-\infty)\varphi(-\infty)}_{0^2=0} - \int_{x_0}^{\infty} \left[\frac{d}{dx}\varphi(x) \right] dx \\
&= -\int_{x_0}^{\infty} \left[\frac{d}{dx}\varphi(x) \right] dx \\
&= -\varphi(x)\Big|_{x=x_0}^{x=\infty} \\
&= -[\varphi(\infty) - \varphi(x_0)] \\
&= \varphi(x_0).
\end{aligned}
\tag{7.114}
$$

7.12.2 Useful formulæ involving H

Some other formulæ involving the Heaviside step function are

$$
H(\pm x) = \lim_{\epsilon \to 0^+} \frac{\mp i}{2\pi} \int_{-\infty}^{+\infty} \frac{e^{ikx}}{k \mp i\epsilon}\,dk,
\tag{7.115}
$$

and

$$
H(x) = \frac{1}{2} + \sum_{l=0}^{\infty} (-1)^l \frac{(2l)!(4l+3)}{2^{2l+2} l!(l+1)!} P_{2l+1}(x),
\tag{7.116}
$$

where $P_{2l+1}(x)$ is a Legendre polynomial. Furthermore,

$$
\delta(x) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} H\left(\frac{\epsilon}{2} - |x| \right).
\tag{7.117}
$$

An integral representation of $H(x)$ is

$$
H(x) = \lim_{\epsilon \downarrow 0^+} \mp \frac{1}{2\pi i} \int_{-\infty}^{\infty} \frac{1}{t \pm i\epsilon} e^{\mp ixt}\,dt.
\tag{7.118}
$$

One commonly used limit form of the Heaviside step function is

$$
H(x) = \lim_{\epsilon \to 0} H_\epsilon(x) = \lim_{\epsilon \to 0} \left[\frac{1}{2} + \frac{1}{\pi} \tan^{-1} \frac{x}{\epsilon} \right].
\tag{7.119}
$$

respectively.

Another limit representation of the Heaviside function is in terms of the *Dirichlet's discontinuity factor* as follows:

$$H(x) = \lim_{t \to \infty} H_t(x)$$

$$= \frac{1}{2} + \frac{1}{\pi} \lim_{t \to \infty} \int_0^t \frac{\sin(kx)}{k} \, dk \qquad (7.120)$$

$$= \frac{1}{2} + \frac{1}{\pi} \int_0^\infty \frac{\sin(kx)}{k} \, dk.$$

A proof [13] uses a variant of the *sine integral function*

$$\mathrm{Si}(y) = \int_0^y \frac{\sin t}{t} \, dt \qquad (7.121)$$

which in the limit of large argument y converges towards the *Dirichlet integral* (no proof is given here)

$$\mathrm{Si}(\infty) = \int_0^\infty \frac{\sin t}{t} \, dt = \frac{\pi}{2}. \qquad (7.122)$$

Suppose we replace t with $t = kx$ in the Dirichlet integral (7.122), whereby $x \neq 0$ is a nonzero constant; that is,

$$\int_0^\infty \frac{\sin(kx)}{kx} \, d(kx) = \int_0^{\pm\infty} \frac{\sin(kx)}{k} \, dk. \qquad (7.123)$$

Note that the integration border $\pm\infty$ changes, depending on whether x is positive or negative, respectively.

If x is positive, we leave the integral (7.123) as is, and we recover the original Dirichlet integral (7.122), which is $\frac{\pi}{2}$. If x is negative, in order to recover the original Dirichlet integral form with the upper limit ∞, we have to perform yet another substitution $k \to -k$ on (7.123), resulting in

$$= \int_0^{-\infty} \frac{\sin(-kx)}{-k} \, d(-k) = -\int_0^\infty \frac{\sin(kx)}{k} \, dk = -\mathrm{Si}(\infty) = -\frac{\pi}{2}, \qquad (7.124)$$

since the sine function is an odd function; that is, $\sin(-\varphi) = -\sin\varphi$.

The Dirichlet's discontinuity factor (7.120) is obtained by normalizing the absolute value of (7.123) [and thus also (7.124)] to $\frac{1}{2}$ by multiplying it with $1/\pi$, and by adding $\frac{1}{2}$.

7.12.3 $H[\varphi]$ distribution

The distribution associated with the Heaviside functiom $H(x)$ is defined by

$$H[\varphi] = \int_{-\infty}^\infty H(x)\varphi(x)\,dx. \qquad (7.125)$$

$H[\varphi]$ can be evaluated and represented as follows:

$$H[\varphi] = \int_{-\infty}^\infty H(x)\varphi(x)\,dx$$

$$= \int_0^\infty \varphi(x)\,dx. \qquad (7.126)$$

7.12.4   *Regularized Heaviside function*

In order to be able to define the distribution associated with the Heaviside function (and its Fourier transform), we sometimes consider the distribution of the *regularized Heaviside function*

$$H_\varepsilon(x) = H(x)e^{-\varepsilon x}, \qquad (7.127)$$

[13] Eli Maor. *Trigonometric Delights.* Princeton University Press, Princeton, 1998. URL http://press.princeton.edu/books/maor/

with $\varepsilon > 0$, such that $\lim_{\varepsilon \to 0^+} H_\varepsilon(x) = H(x)$.

7.12.5 Fourier transform of Heaviside (unit step) function

The Fourier transform of the Heaviside (unit step) function cannot be directly obtained by insertion into Eq. (6.19), because the associated integrals do not exist. For a derivation of the Fourier transform of the Heaviside (unit step) function we shall thus use the regularized Heaviside function (7.127), and arrive at *Sokhotsky's formula* (also known as the *Plemelj's formula*, or the *Plemelj-Sokhotsky formula*)

$$
\begin{aligned}
\mathscr{F}[H(x)] = \widetilde{H}(k) &= \int_{-\infty}^{\infty} H(x) e^{-ikx} dx \\
&= \pi \delta(k) - i \mathscr{P} \frac{1}{k} \\
&= -i \left(i\pi \delta(k) + \mathscr{P} \frac{1}{k} \right) \\
&= \lim_{\varepsilon \to 0^+} -\frac{i}{k - i\varepsilon}
\end{aligned}
\tag{7.128}
$$

We shall compute the Fourier transform of the regularized Heaviside function $H_\varepsilon(x) = H(x) e^{-\varepsilon x}$, with $\varepsilon > 0$, of Eq. (7.127); that is [14],

$$
\begin{aligned}
\mathscr{F}[H_\varepsilon(x)] = \mathscr{F}[H(x) e^{-\varepsilon x}] &= \widetilde{H_\varepsilon}(k) \\
&= \int_{-\infty}^{\infty} H_\varepsilon(x) e^{-ikx} dx \\
&= \int_{-\infty}^{\infty} H(x) e^{-\varepsilon x} e^{-ikx} dx \\
&= \int_{-\infty}^{\infty} H(x) e^{-ikx + (-i)^2 \varepsilon x} dx \\
&= \int_{-\infty}^{\infty} H(x) e^{-i(k - i\varepsilon)x} dx \\
&= \int_{0}^{\infty} e^{-i(k - i\varepsilon)x} dx \\
&= \left[-\frac{e^{-i(k - i\varepsilon)x}}{i(k - i\varepsilon)} \right] \Bigg|_{x=0}^{x=\infty} \\
&= \left[-\frac{e^{-ik} e^{-\varepsilon)x}}{i(k - i\varepsilon)} \right] \Bigg|_{x=0}^{x=\infty} \\
= \left[-\frac{e^{-ik\infty} e^{-\varepsilon \infty}}{i(k - i\varepsilon)} \right] &- \left[-\frac{e^{-ik0} e^{-\varepsilon 0}}{i(k - i\varepsilon)} \right] \\
= 0 - \frac{(-1)}{i(k - i\varepsilon)} &= -\frac{i}{(k - i\varepsilon)}.
\end{aligned}
\tag{7.129}
$$

By using Sokhotsky's formula (7.109) we conclude that

$$
\mathscr{F}[H(x)] = \mathscr{F}[H_{0^+}(x)] = \lim_{\varepsilon \to 0^+} \mathscr{F}[H_\varepsilon(x)] = \pi \delta(k) - i \mathscr{P}\left(\frac{1}{k} \right).
\tag{7.130}
$$

7.13   The sign function

7.13.1   Definition

The *sign function* is defined by

$$\mathrm{sgn}(x - x_0) = \begin{cases} -1 & \text{for } x < x_0 \\ 0 & \text{for } x = x_0 \\ +1 & \text{for } x > x_0 \end{cases}.$$

(7.131)

It is plotted in Fig. 7.5.

7.13.2   Connection to the Heaviside function

In terms of the Heaviside step function, in particular, with $H(0) = \frac{1}{2}$ as in Eq. (7.111), the sign function can be written by "stretching" the former (the Heaviside step function) by a factor of two, and shifting it by one negative unit as follows

$$\mathrm{sgn}(x - x_0) = 2H(x - x_0) - 1,$$

$$H(x - x_0) = \frac{1}{2}\left[\mathrm{sgn}(x - x_0) + 1\right];$$

and also

$$\mathrm{sgn}(x - x_0) = H(x - x_0) - H(x_0 - x).$$

(7.132)

Therefore, the derivative of the sign function is

$$\frac{d}{dx}\mathrm{sgn}(x - x_0) = \frac{d}{dx}[2H(x - x_0) - 1] = 2\delta(x - x_0).$$

(7.133)

Note also that $\mathrm{sgn}(x - x_0) = -\mathrm{sgn}(x_0 - x)$.

7.13.3   Sign sequence

The sequence of functions

$$\mathrm{sgn}_n(x - x_0) = \begin{cases} -e^{-\frac{x}{n}} & \text{for } x < x_0 \\ +e^{\frac{-x}{n}} & \text{for } x > x_0 \end{cases}$$

(7.134)

is a limiting sequence of $\mathrm{sgn}(x - x_0) \stackrel{x \neq x_0}{=} \lim_{n \to \infty} \mathrm{sgn}_n(x - x_0)$.

We can also use the Dirichlet integral to express a limiting sequence for the singn function, in a similar way as the derivation of Eqs. (7.120); that is,

$$\mathrm{sgn}(x) = \lim_{t \to \infty} \mathrm{sgn}_t(x)$$

$$= \frac{2}{\pi} \lim_{t \to \infty} \int_0^t \frac{\sin(kx)}{k} dk$$

$$= \frac{2}{\pi} \int_0^\infty \frac{\sin(kx)}{k} dk.$$

(7.135)

Note (without proof) that

$$\mathrm{sgn}(x) = \frac{4}{\pi} \sum_{n=0}^{\infty} \frac{\sin[(2n+1)x]}{(2n+1)}$$

(7.136)

$$= \frac{4}{\pi} \sum_{n=0}^{\infty} (-1)^n \frac{\cos[(2n+1)(x - \pi/2)]}{(2n+1)}, \quad -\pi < x < \pi.$$

(7.137)
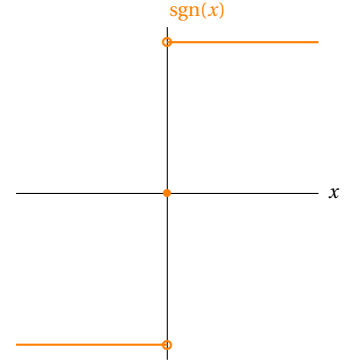
Figure 7.5: Plot of the sign function $\mathrm{sgn}(x)$.

7.13.4   Fourier transform of sgn

Since the Fourier transform is linear, we may use the connection between the sign and the Heaviside functions $\operatorname{sgn}(x) = 2H(x) - 1$, Eq. (7.132), together with the Fourier transform of the Heaviside function $\mathscr{F}[H(x)] = \pi\delta(k) - i\mathscr{P}\left(\frac{1}{k}\right)$, Eq. (7.130) and the Dirac delta function $\mathscr{F}[1] = 2\pi\delta(k)$, Eq. (7.85), to compose and compute the Fourier transform of sgn:

$$\mathscr{F}[\operatorname{sgn}(x)] = \mathscr{F}[2H(x) - 1] = 2\mathscr{F}[H(x)] - \mathscr{F}[1]$$
$$= 2\left[\pi\delta(k) - i\mathscr{P}\left(\frac{1}{k}\right)\right] - 2\pi\delta(k) \qquad (7.138)$$
$$= -2i\mathscr{P}\left(\frac{1}{k}\right).$$

7.14   Absolute value function (or modulus)

7.14.1   Definition

The *absolute value* (or *modulus*) of x is defined by

$$|x - x_0| = \begin{cases} x - x_0 & \text{for } x > x_0 \\ 0 & \text{for } x = x_0 \\ x_0 - x & \text{for } x < x_0 \end{cases} \qquad (7.139)$$

It is plotted in Fig. 7.6.

7.14.2   Connection of absolute value with sign and Heaviside functions

Its relationship to the sign function is twofold: on the one hand, there is

$$|x| = x\operatorname{sgn}(x), \qquad (7.140)$$

and thus, for $x \neq 0$,

$$\operatorname{sgn}(x) = \frac{|x|}{x} = \frac{x}{|x|}. \qquad (7.141)$$

On the other hand, the derivative of the absolute value function is the sign function, at least up to a singular point at $x = 0$, and thus the absolute value function can be interpreted as the integral of the sign function (in the distributional sense); that is,

$$\frac{d}{dx}|x| = \begin{cases} 1 & \text{for } x > 0 \\ 0 & \text{for } x = 0 \\ -1 & \text{for } x < 0 \end{cases} = \operatorname{sgn}(x),$$
$$|x| = \int \operatorname{sgn}(x)\,dx. \qquad (7.142)$$

This can be proven by inserting $|x| = x\operatorname{sgn}(x)$; that is,

$$\frac{d}{dx}|x| = \frac{d}{dx}x\operatorname{sgn}(x) = \operatorname{sgn}(x) + x\frac{d}{dx}\operatorname{sgn}(x)$$
$$= \operatorname{sgn}(x) + x\frac{d}{dx}[2H(x) - 1] = \operatorname{sgn}(x) - 2\underbrace{x\delta(x)}_{=0}. \qquad (7.143)$$

Figure 7.6: Plot of the absolute value $|x|$.

7.15   Some examples

Let us compute some concrete examples related to distributions.

1. For a start, let us prove that

$$\lim_{\epsilon \to 0} \frac{\epsilon \sin^2 \frac{x}{\epsilon}}{\pi x^2} = \delta(x). \tag{7.144}$$

As a hint, take $\int_{-\infty}^{+\infty} \frac{\sin^2 x}{x^2} dx = \pi$.

Let us prove this conjecture by integrating over a good test function φ

$$\frac{1}{\pi} \lim_{\epsilon \to 0} \int_{-\infty}^{+\infty} \frac{\epsilon \sin^2 \left(\frac{x}{\epsilon} \right)}{x^2} \varphi(x) dx$$

$$[\text{variable substitution } y = \frac{x}{\epsilon}, \frac{dy}{dx} = \frac{1}{\epsilon}, dx = \epsilon dy]$$

$$= \frac{1}{\pi} \lim_{\epsilon \to 0} \int_{-\infty}^{+\infty} \varphi(\epsilon y) \frac{\epsilon^2 \sin^2(y)}{\epsilon^2 y^2} dy \tag{7.145}$$

$$= \frac{1}{\pi} \varphi(0) \int_{-\infty}^{+\infty} \frac{\sin^2(y)}{y^2} dy$$

$$= \varphi(0).$$

Hence we can identify

$$\lim_{\epsilon \to 0} \frac{\epsilon \sin^2 \left(\frac{x}{\epsilon} \right)}{\pi x^2} = \delta(x). \tag{7.146}$$

2. In order to prove that $\frac{1}{\pi} \frac{n e^{-x^2}}{1 + n^2 x^2}$ is a δ-sequence we proceed again by integrating over a good test function φ, and with the hint that $\int_{-\infty}^{+\infty} dx / (1 + x^2) = \pi$ we obtain

$$\lim_{n \to \infty} \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{n e^{-x^2}}{1 + n^2 x^2} \varphi(x) dx$$

$$[\text{variable substitution } y = xn, x = \frac{y}{n}, \frac{dy}{dx} = n, dx = \frac{dy}{n}]$$

$$= \lim_{n \to \infty} \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{n e^{-\left(\frac{y}{n} \right)^2}}{1 + y^2} \varphi \left(\frac{y}{n} \right) \frac{dy}{n}$$

$$= \frac{1}{\pi} \int_{-\infty}^{+\infty} \lim_{n \to \infty} \left[e^{-\left(\frac{y}{n} \right)^2} \varphi \left(\frac{y}{n} \right) \right] \frac{1}{1 + y^2} dy$$

$$= \frac{1}{\pi} \int_{-\infty}^{+\infty} \left[e^0 \varphi(0) \right] \frac{1}{1 + y^2} dy$$

$$= \frac{\varphi(0)}{\pi} \int_{-\infty}^{+\infty} \frac{1}{1 + y^2} dy$$

$$= \frac{\varphi(0)}{\pi} \pi$$

$$= \varphi(0). \tag{7.147}$$

Hence we can identify

$$\lim_{n\to\infty} \frac{1}{\pi}\frac{ne^{-x^2}}{1+n^2x^2} = \delta(x). \tag{7.148}$$

3. Let us prove that $x^n\delta^{(n)}(x) = C\delta(x)$ and determine the constant C. We proceed again by integrating over a good test function φ. First note that if $\varphi(x)$ is a good test function, then so is $x^n\varphi(x)$.

$$
\begin{aligned}
\int dx\, x^n\delta^{(n)}(x)\varphi(x) &= \int dx\,\delta^{(n)}(x)\big[x^n\varphi(x)\big] = \\
&= (-1)^n\int dx\,\delta(x)\big[x^n\varphi(x)\big]^{(n)} = \\
&= (-1)^n\int dx\,\delta(x)\big[nx^{n-1}\varphi(x) + x^n\varphi'(x)\big]^{(n-1)} = \\
&\quad\cdots \\
&= (-1)^n\int dx\,\delta(x)\left[\sum_{k=0}^{n}\binom{n}{k}(x^n)^{(k)}\varphi^{(n-k)}(x)\right] = \\
&= (-1)^n\int dx\,\delta(x)\big[n!\varphi(x) + n\cdot n!\, x\varphi'(x) + \cdots + x^n\varphi^{(n)}(x)\big] = \\
&= (-1)^n n!\int dx\,\delta(x)\varphi(x);
\end{aligned}
$$

hence, $C = (-1)^n n!$. Note that $\varphi(x)$ is a good test function then so is $x^n\varphi(x)$.

4. Let us simplify $\int_{-\infty}^{\infty}\delta(x^2-a^2)g(x)\,dx$. First recall Eq. (7.66) stating that

$$\delta(f(x)) = \sum_i \frac{\delta(x-x_i)}{|f'(x_i)|},$$

whenever x_i are simple roots of $f(x)$, and $f'(x_i)\neq 0$. In our case, $f(x) = x^2 - a^2 = (x-a)(x+a)$, and the roots are $x = \pm a$. Furthermore,

$$f'(x) = (x-a) + (x+a);$$

hence

$$|f'(a)| = 2|a|, \qquad |f'(-a)| = |-2a| = 2|a|.$$

As a result,

$$\delta(x^2-a^2) = \delta\big((x-a)(x+a)\big) = \frac{1}{|2a|}\big(\delta(x-a)+\delta(x+a)\big).$$

Taking this into account we finally obtain

$$
\begin{aligned}
&\int_{-\infty}^{+\infty}\delta(x^2-a^2)g(x)\,dx \\
&= \int_{-\infty}^{+\infty}\frac{\delta(x-a)+\delta(x+a)}{2|a|}g(x)\,dx \\
&= \frac{g(a)+g(-a)}{2|a|}.
\end{aligned}
\tag{7.149}
$$

5. Let us evaluate

$$I = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\delta(x_1^2+x_2^2+x_3^2-R^2)\,d^3x \tag{7.150}$$

for $R \in \mathbb{R}, R > 0$. We may, of course, remain tn the standard Cartesian coordinate system and evaluate the integral by "brute force." Alternatively, a more elegant way is to use the spherical symmetry of the problem and use spherical coordinates $r, \Omega(\theta, \varphi)$ by rewriting I into

$$I = \int_{r,\Omega} r^2 \delta(r^2 - R^2) d\Omega dr. \tag{7.151}$$

As the integral kernel $\delta(r^2 - R^2)$ just depends on the radial coordinate r the angular coordinates just integrate to 4π. Next we make use of Eq. (7.66), eliminate the solution for $r = -R$, and obtain

$$
\begin{aligned}
I &= 4\pi \int_0^\infty r^2 \delta(r^2 - R^2) dr \\
&= 4\pi \int_0^\infty r^2 \frac{\delta(r + R) + \delta(r - R)}{2R} dr \\
&= 4\pi \int_0^\infty r^2 \frac{\delta(r - R)}{2R} dr \\
&= 2\pi R.
\end{aligned}
\tag{7.152}
$$

6. Let us compute

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \delta(x^3 - y^2 + 2y)\delta(x + y)H(y - x - 6)f(x, y)\, dx\, dy. \tag{7.153}$$

First, in dealing with $\delta(x + y)$, we evaluate the y integration at $x = -y$ or $y = -x$:

$$\int_{-\infty}^{\infty} \delta(x^3 - x^2 - 2x)H(-2x - 6)f(x, -x)dx$$

Use of Eq. (7.66)

$$\delta(f(x)) = \sum_i \frac{1}{|f'(x_i)|}\delta(x - x_i),$$

at the roots

$$
\begin{aligned}
x_1 &= 0 \\
x_{2,3} &= \frac{1 \pm \sqrt{1 + 8}}{2} = \frac{1 \pm 3}{2} = \begin{cases} 2 \\ -1 \end{cases}
\end{aligned}
$$

of the argument $f(x) = x^3 - x^2 - 2x = x(x^2 - x - 2) = x(x - 2)(x + 1)$ of the remaining δ-function, together with

$$f'(x) = \frac{d}{dx}(x^3 - x^2 - 2x) = 3x^2 - 2x - 2;$$

yields

$$
\begin{aligned}
&\int_{-\infty}^{\infty} dx \frac{\delta(x) + \delta(x - 2) + \delta(x + 1)}{|3x^2 - 2x - 2|} H(-2x - 6)f(x, -x) = \\
&= \frac{1}{|-2|} \underbrace{H(-6)}_{=0} f(0, -0) + \frac{1}{|12 - 4 - 2|} \underbrace{H(-4 - 6)}_{=0} f(2, -2) + \\
&\quad + \frac{1}{|3 + 2 - 2|} \underbrace{H(2 - 6)}_{=0} f(-1, 1) \\
&= 0
\end{aligned}
$$

7. When simplifying derivatives of generalized functions it is always useful to evaluate their properties – such as $x\delta(x) = 0$, $f(x)\delta(x - x_0) = f(x_0)\delta(x - x_0)$, or $\delta(-x) = \delta(x)$ – first and before proceeding with the next differentiation or evaluation. We shall present some applications of this "rule" next.

First, simplify

$$\left(\frac{d}{dx} - \omega\right) H(x)e^{\omega x} \tag{7.154}$$

as follows

$$\frac{d}{dx}\left[H(x)e^{\omega x}\right] - \omega H(x)e^{\omega x}$$
$$= \delta(x)e^{\omega x} + \omega H(x)e^{\omega x} - \omega H(x)e^{\omega x} \tag{7.155}$$
$$= \delta(x)e^0$$
$$= \delta(x)$$

8. Next, simplify

$$\left(\frac{d^2}{dx^2} + \omega^2\right)\frac{1}{\omega}H(x)\sin(\omega x) \tag{7.156}$$

as follows

$$\frac{d^2}{dx^2}\left[\frac{1}{\omega}H(x)\sin(\omega x)\right] + \omega H(x)\sin(\omega x)$$
$$= \frac{1}{\omega}\frac{d}{dx}\left[\underbrace{\delta(x)\sin(\omega x)}_{=0} + \omega H(x)\cos(\omega x)\right] + \omega H(x)\sin(\omega x)$$
$$= \frac{1}{\omega}\left[\omega\underbrace{\delta(x)\cos(\omega x)}_{\delta(x)} - \omega^2 H(x)\sin(\omega x)\right] + \omega H(x)\sin(\omega x) = \delta(x)$$
$$\tag{7.157}$$

9. Let us compute the nth derivative of

$$f(x) = \begin{cases} 0 & \text{for } x < 0, \\ x & \text{for } 0 \le x \le 1, \\ 0 & \text{for } x > 1. \end{cases} \tag{7.158}$$

As depicted in Fig. 7.7, f can be composed from two functions $f(x) = f_2(x) \cdot f_1(x)$; and this composition can be done in at least two ways.

Decomposition (i)

$$f(x) = x\left[H(x) - H(x - 1)\right] = xH(x) - xH(x - 1)$$
$$f'(x) = H(x) + x\delta(x) - H(x - 1) - x\delta(x - 1)$$

Because of $x\delta(x - a) = a\delta(x - a)$,

$$f'(x) = H(x) - H(x - 1) - \delta(x - 1)$$
$$f''(x) = \delta(x) - \delta(x - 1) - \delta'(x - 1)$$

and hence by induction

$$f^{(n)}(x) = \delta^{(n-2)}(x) - \delta^{(n-2)}(x - 1) - \delta^{(n-1)}(x - 1)$$

$f_1(x) = H(x) - H(x - 1)$
$\quad\ \ = H(x)H(1 - x)$

$f_2(x) = x$

Figure 7.7: Composition of $f(x)$

for $n > 1$.

Decomposition (ii)

$$
\begin{aligned}
f(x) &= xH(x)H(1-x) \\
f'(x) &= H(x)H(1-x) + \underbrace{x\delta(x)}_{=0}H(1-x) + \underbrace{xH(x)(-1)\delta(1-x)}_{-H(x)\delta(1-x)} = \\
&= H(x)H(1-x) - \delta(1-x) = [\delta(x) = \delta(-x)] = H(x)H(1-x) - \delta(x-1) \\
f''(x) &= \underbrace{\delta(x)H(1-x)}_{=\delta(x)} + \underbrace{(-1)H(x)\delta(1-x)}_{-\delta(1-x)} - \delta'(x-1) = \\
&= \delta(x) - \delta(x-1) - \delta'(x-1)
\end{aligned}
$$

and hence by induction

$$
f^{(n)}(x) = \delta^{(n-2)}(x) - \delta^{(n-2)}(x-1) - \delta^{(n-1)}(x-1)
$$

for $n > 1$.

10. Let us compute the nth derivative of

$$
f(x) = \begin{cases} |\sin x| & \text{for } -\pi \le x \le \pi, \\ 0 & \text{for } |x| > \pi. \end{cases} \tag{7.159}
$$

$$
f(x) = |\sin x| H(\pi + x) H(\pi - x)
$$

$$
|\sin x| = \sin x \, \text{sgn}(\sin x) = \sin x \, \text{sgn} \, x \qquad \text{für} \quad -\pi < x < \pi;
$$

hence we start from

$$
f(x) = \sin x \, \text{sgn} \, x \, H(\pi + x) H(\pi - x),
$$

Note that

$$
\begin{aligned}
\text{sgn} \, x &= H(x) - H(-x), \\
(\text{sgn} \, x)' &= H'(x) - H'(-x)(-1) = \delta(x) + \delta(-x) = \delta(x) + \delta(x) = 2\delta(x).
\end{aligned}
$$

$$
\begin{aligned}
f'(x) &= \cos x \, \text{sgn} \, x \, H(\pi + x) H(\pi - x) + \sin x \, 2\delta(x) H(\pi + x) H(\pi - x) + \\
&\quad + \sin x \, \text{sgn} \, x \, \delta(\pi + x) H(\pi - x) + \sin x \, \text{sgn} \, x \, H(\pi + x)\delta(\pi - x)(-1) = \\
&= \cos x \, \text{sgn} \, x \, H(\pi + x) H(\pi - x) \\
f''(x) &= -\sin x \, \text{sgn} \, x \, H(\pi + x) H(\pi - x) + \cos x \, 2\delta(x) H(\pi + x) H(\pi - x) + \\
&\quad + \cos x \, \text{sgn} \, x \, \delta(\pi + x) H(\pi - x) + \cos x \, \text{sgn} \, x \, H(\pi + x)\delta(\pi - x)(-1) = \\
&= -\sin x \, \text{sgn} \, x \, H(\pi + x) H(\pi - x) + 2\delta(x) + \delta(\pi + x) + \delta(\pi - x) \\
f'''(x) &= -\cos x \, \text{sgn} \, x \, H(\pi + x) H(\pi - x) - \sin x \, 2\delta(x) H(\pi + x) H(\pi - x) - \\
&\quad - \sin x \, \text{sgn} \, x \, \delta(\pi + x) H(\pi - x) - \sin x \, \text{sgn} \, x \, H(\pi + x)\delta(\pi - x)(-1) + \\
&\quad + 2\delta'(x) + \delta'(\pi + x) - \delta'(\pi - x) = \\
&= -\cos x \, \text{sgn} \, x \, H(\pi + x) H(\pi - x) + 2\delta'(x) + \delta'(\pi + x) - \delta'(\pi - x) \\
f^{(4)}(x) &= \sin x \, \text{sgn} \, x \, H(\pi + x) H(\pi - x) - \cos x \, 2\delta(x) H(\pi + x) H(\pi - x) - \\
&\quad - \cos x \, \text{sgn} \, x \, \delta(\pi + x) H(\pi - x) - \cos x \, \text{sgn} \, x \, H(\pi + x)\delta(\pi - x)(-1) + \\
&\quad + 2\delta''(x) + \delta''(\pi + x) + \delta''(\pi - x) = \\
&= \sin x \, \text{sgn} \, x \, H(\pi + x) H(\pi - x) - 2\delta(x) - \delta(\pi + x) - \delta(\pi - x) + \\
&\quad + 2\delta''(x) + \delta''(\pi + x) + \delta''(\pi - x);
\end{aligned}
$$

hence

$$f^{(4)} \;=\; f(x) - 2\delta(x) + 2\delta''(x) - \delta(\pi + x) + \delta''(\pi + x) - \delta(\pi - x) + \delta''(\pi - x),$$

$$f^{(5)} \;=\; f'(x) - 2\delta'(x) + 2\delta'''(x) - \delta'(\pi + x) + \delta'''(\pi + x) + \delta'(\pi - x) - \delta'''(\pi - x);$$

and thus by induction

$$f^{(n)} \;=\; f^{(n-4)}(x) - 2\delta^{(n-4)}(x) + 2\delta^{(n-2)}(x) - \delta^{(n-4)}(\pi + x) +$$
$$+ \delta^{(n-2)}(\pi + x) + (-1)^{n-1}\delta^{(n-4)}(\pi - x) + (-1)^{n}\delta^{(n-2)}(\pi - x)$$
$$(n = 4, 5, 6, \dots)$$

8

Green's function

This chapter is the beginning of a series of chapters dealing with the solution to differential equations of theoretical physics. These differential equations are *linear*; that is, the "sought after" function $\Psi(x), y(x), \phi(t)$ *et cetera* occur only as a polynomial of degree zero and one, and *not* of any higher degree, such as, for instance, $[y(x)]^2$.

8.1 *Elegant way to solve linear differential equations*

Green's functions present a very elegant way of solving linear differential equations of the form

$$\mathscr{L}_x y(x) = f(x), \text{ with the differential operator}$$

$$\mathscr{L}_x = a_n(x)\frac{d^n}{dx^n} + a_{n-1}(x)\frac{d^{n-1}}{dx^{n-1}} + \ldots + a_1(x)\frac{d}{dx} + a_0(x) \tag{8.1}$$

$$= \sum_{j=0}^{n} a_j(x)\frac{d^j}{dx^j},$$

where $a_i(x), 0 \le i \le n$ are functions of x. The idea is quite straightforward: if we are able to obtain the "inverse" G of the differential operator $\mathscr{L}$ defined by

$$\mathscr{L}_x G(x, x') = \delta(x - x'), \tag{8.2}$$

with δ representing Dirac's delta function, then the solution to the inhomogenuous differential equation (8.1) can be obtained by integrating $G(x, x')$ alongside with the inhomogenuous term $f(x')$; that is,

$$y(x) = \int_{-\infty}^{\infty} G(x, x') f(x') dx'. \tag{8.3}$$

This claim, as posted in Eq. (8.3), can be verified by explicitly applying the differential operator $\mathscr{L}_x$ to the solution $y(x)$,

$$\mathscr{L}_x y(x)$$

$$= \mathscr{L}_x \int_{-\infty}^{\infty} G(x, x') f(x') dx'$$

$$= \int_{-\infty}^{\infty} \mathscr{L}_x G(x, x') f(x') dx' \tag{8.4}$$

$$= \int_{-\infty}^{\infty} \delta(x - x') f(x') dx'$$

$$= f(x).$$

Let us check whether $G(x, x') = H(x - x') \sinh(x - x')$ is a Green's function of the differential operator $\mathscr{L}_x = \frac{d^2}{dx^2} - 1$. In this case, all we have to do is to verify that $\mathscr{L}_x$, applied to $G(x, x')$, actually renders $\delta(x - x')$, as required by Eq. (8.2).

$$\mathscr{L}_x G(x, x') = \delta(x - x')$$

$$\left(\frac{d^2}{dx^2} - 1 \right) H(x - x') \sinh(x - x') \overset{?}{=} \delta(x - x') \tag{8.5}$$

Note that $\dfrac{d}{dx} \sinh x = \cosh x, \quad \dfrac{d}{dx} \cosh x = \sinh x$; and hence

$$\frac{d}{dx} \left(\underbrace{\delta(x - x') \sinh(x - x')}_{= 0} + H(x - x') \cosh(x - x') \right) - H(x - x') \sinh(x - x') =$$

$$\delta(x - x') \cosh(x - x') + H(x - x') \sinh(x - x') - H(x - x') \sinh(x - x') = \delta(x - x').$$

8.2   Nonuniqueness of solution

The solution (8.4) so obtained is *not unique*, as it is only a special solution to the inhomogenuous equation (8.1). The general solution to (8.1) can be found by adding the general solution $y_0(x)$ of the corresponding *homogenuous* differential equation

$$\mathscr{L}_x y(x) = 0 \tag{8.6}$$

to one special solution – say, the one obtained in Eq. (8.4) through Green's function techniques.

Indeed, the most general solution

$$Y(x) = y(x) + y_0(x) \tag{8.7}$$

clearly is a solution of the inhomogenuous differential equation (8.4), as

$$\mathscr{L}_x Y(x) = \mathscr{L}_x y(x) + \mathscr{L}_x y_0(x) = f(x) + 0 = f(x). \tag{8.8}$$

Conversely, any two distinct special solutions $y_1(x)$ and $y_2(x)$ of the inhomogenuous differential equation (8.4) differ only by a function which is a solution of the homogenuous differential equation (8.6), because due to linearity of $\mathscr{L}_x$, their difference $y_d(x) = y_1(x) - y_2(x)$ can be parameterized by some function in y_0

$$\mathscr{L}_x [y_1(x) - y_2(x)] = \mathscr{L}_x y_1(x) + \mathscr{L}_x y_2(x) = f(x) - f(x) = 0. \tag{8.9}$$

8.3   Green's functions of translational invariant differential operators

From now on, we assume that the coefficients $a_j(x) = a_j$ in Eq. (8.1) are constants, and thus are *translational invariant*. Then the differential

operator $\mathscr{L}_x$, as well as the entire *Ansatz* (8.2) for $G(x, x')$, is translation invariant, because derivatives are defined only by relative distances, and $\delta(x - x')$ is translation invariant for the same reason. Hence,

$$G(x, x') = G(x - x'). \tag{8.10}$$

For such translation invariant systems, the Fourier analysis presents an excellent way of analyzing the situation.

Let us see why translanslation invariance of the coefficients $a_j(x) = a_j(x + \xi) = a_j$ under the translation $x \to x + \xi$ with arbitrary ξ – that is, independence of the coefficients a_j on the "coordinate" or "parameter" x – and thus of the Green's function, implies a simple form of the latter. Translanslation invariance of the Green's function really means

$$G(x + \xi, x' + \xi) = G(x, x'). \tag{8.11}$$

Now set $\xi = -x'$; then we can define a new Green's function that just depends on one argument (instead of previously two), which is the difference of the old arguments

$$G(x - x', x' - x') = G(x - x', 0) \to G(x - x'). \tag{8.12}$$

8.4   Solutions with fixed boundary or initial values

For applications it is important to adapt the solutions of some inhomenuous differential equation to boundary and initial value problems. In particular, a properly chosen $G(x - x')$, in its dependence on the parameter x, "inherits" some behaviour of the solution $y(x)$. Suppose, for instance, we would like to find solutions with $y(x_i) = 0$ for some parameter values x_i, $i = 1, \ldots, k$. Then, the Green's function G must vanish there also

$$G(x_i - x') = 0 \text{ for } i = 1, \ldots, k. \tag{8.13}$$

8.5   Finding Green's functions by spectral decompositions

It has been mentioned earlier (cf. Section 7.6.5 on page 154) that the δ-function can be expressed in terms of various *eigenfunction expansions*. We shall make use of these expansions here [1].

Suppose $\psi_i(x)$ are *eigenfunctions* of the differential operator $\mathscr{L}_x$, and λ_i are the associated *eigenvalues*; that is,

$$\mathscr{L}_x \psi_i(x) = \lambda_i \psi_i(x). \tag{8.14}$$

Suppose further that $\mathscr{L}_x$ is of degree n, and therefore (we assume without proof) that we know all (a complete set of) the n eigenfunctions $\psi_1(x), \psi_2(x), \ldots, \psi_n(x)$ of $\mathscr{L}_x$. In this case, orthogonality of the system of eigenfunctions holds, such that

$$\int_{-\infty}^{\infty} \psi_i(x) \overline{\psi_j(x)} dx = \delta_{ij}, \tag{8.15}$$

[1] Dean G. Duffy. *Green's Functions with Applications*. Chapman and Hall/CRC, Boca Raton, 2001

as well as completeness, such that

$$\sum_{i=1}^{n} \psi_i(x)\overline{\psi_i(x')} = \delta(x-x').\tag{8.16}$$

$\overline{\psi_i(x')}$ stands for the complex conjugate of $\psi_i(x')$. The sum in Eq. (8.16) stands for an integral in the case of continuous spectrum of $\mathcal{L}_x$. In this case, the Kronecker δ_{ij} in (8.15) is replaced by the Dirac delta function $\delta(k-k')$. It has been mentioned earlier that the δ-function can be expressed in terms of various *eigenfunction expansions*.

The Green's function of $\mathcal{L}_x$ can be written as the spectral sum of the absolute squares of the eigenfunctions, divided by the eigenvalues λ_j; that is,

$$G(x-x') = \sum_{j=1}^{n} \frac{\psi_j(x)\overline{\psi_j(x')}}{\lambda_j}.\tag{8.17}$$

For the sake of proof, apply the differential operator $\mathcal{L}_x$ to the Green's function *Ansatz* G of Eq. (8.17) and verify that it satisfies Eq. (8.2):

$$\begin{aligned}
&\mathcal{L}_x G(x-x')\\
&= \mathcal{L}_x \sum_{j=1}^{n} \frac{\psi_j(x)\overline{\psi_j(x')}}{\lambda_j}\\
&= \sum_{j=1}^{n} \frac{[\mathcal{L}_x\psi_j(x)]\overline{\psi_j(x')}}{\lambda_j}\\
&= \sum_{j=1}^{n} \frac{[\lambda_j\psi_j(x)]\overline{\psi_j(x')}}{\lambda_j}\\
&= \sum_{j=1}^{n} \psi_j(x)\overline{\psi_j(x')}\\
&= \delta(x-x').
\end{aligned}\tag{8.18}$$

1. For a demonstration of completeness of systems of eigenfunctions, consider, for instance, the differential equation corresponding to the harmonic vibration [please do not confuse this with the harmonic oscillator (6.29)]

$$\mathcal{L}_t\phi(t) = \frac{d^2}{dt^2}\phi(t) = k^2,\tag{8.19}$$

with $k \in \mathbb{R}$.

Without any boundary conditions the associated eigenfunctions are

$$\psi_\omega(t) = e^{\pm i\omega t},\tag{8.20}$$

with $0 \le \omega \le \infty$, and with eigenvalue $-\omega^2$. Taking the complex conjugate and integrating over ω yields [modulo a constant factor which depends on the choice of Fourier transform parameters; see also Eq. (7.85)]

$$\begin{aligned}
&\int_{-\infty}^{\infty} \psi_\omega(t)\overline{\psi_\omega(t')}\,d\omega\\
&= \int_{-\infty}^{\infty} e^{-i\omega t}e^{i\omega t'}\,d\omega\\
&= \int_{-\infty}^{\infty} e^{-i\omega(t-t')}\,d\omega\\
&= \delta(t-t').
\end{aligned}\tag{8.21}$$

The associated Green's function is

$$G(t - t') = \int_{-\infty}^{\infty} \frac{e^{\pm i\omega(t-t')}}{(-\omega^2)} d\omega. \qquad (8.22)$$

The solution is obtained by integrating over the constant k^2; that is,

$$\phi(t) = \int_{-\infty}^{\infty} G(t - t') k^2 dt' = - \int_{-\infty}^{\infty} \left(\frac{k}{\omega}\right)^2 e^{\pm i\omega(t-t')} d\omega dt'. \qquad (8.23)$$

Suppose that, additionally, we impose boundary conditions; e.g., $\phi(0) = \phi(L) = 0$, representing a string "fastened" at positions 0 and L. In this case the eigenfunctions change to

$$\psi_n(t) = \sin(\omega_n t) = \sin\left(\frac{n\pi}{L} t\right), \qquad (8.24)$$

with $\omega_n = \frac{n\pi}{L}$ and $n \in \mathbb{Z}$. We can deduce orthogonality and completeness by listening to the orthogonality relations for sines (6.11).

2. For the sake of another example suppose, from the Euler-Bernoulli bending theory, we know (no proof is given here) that the equation for the quasistatic bending of slender, isotropic, homogeneous beams of constant cross-section under an applied transverse load $q(x)$ is given by

$$\mathcal{L}_x y(x) = \frac{d^4}{dx^4} y(x) = q(x) \approx c, \qquad (8.25)$$

with constant $c \in \mathbb{R}$. Let us further assume the boundary conditions

$$y(0) = y(L) = \frac{d^2}{dx^2} y(0) = \frac{d^2}{dx^2} y(L) = 0. \qquad (8.26)$$

Also, we require that $y(x)$ vanishes everywhere except inbetween 0 and L; that is, $y(x) = 0$ for $x = (-\infty, 0)$ and for $x = (l, \infty)$. Then in accordance with these boundary conditions, the system of eigenfunctions $\{\psi_j(x)\}$ of $\mathcal{L}_x$ can be written as

$$\psi_j(x) = \sqrt{\frac{2}{L}} \sin\left(\frac{\pi j x}{L}\right) \qquad (8.27)$$

for $j = 1, 2, \ldots$. The associated eigenvalues

$$\lambda_j = \left(\frac{\pi j}{L}\right)^4$$

can be verified through explicit differentiation

$$\begin{aligned}
\mathcal{L}_x \psi_j(x) = \mathcal{L}_x \sqrt{\frac{2}{L}} \sin\left(\frac{\pi j x}{L}\right) \\
= \mathcal{L}_x \sqrt{\frac{2}{L}} \sin\left(\frac{\pi j x}{L}\right) \\
= \left(\frac{\pi j}{L}\right)^4 \sqrt{\frac{2}{L}} \sin\left(\frac{\pi j x}{L}\right) \\
= \left(\frac{\pi j}{L}\right)^4 \psi_j(x).
\end{aligned} \qquad (8.28)$$

The cosine functions which are also solutions of the Euler-Bernoulli equations (8.25) do not vanish at the origin $x = 0$.

Hence,

$$
\begin{aligned}
G(x - x')(x) &= \frac{2}{L} \sum_{j=1}^{\infty} \frac{\sin\left(\frac{\pi j x}{L}\right) \sin\left(\frac{\pi j x'}{L}\right)}{\left(\frac{\pi j}{L}\right)^4} \\
&= \frac{2L^3}{\pi^4} \sum_{j=1}^{\infty} \frac{1}{j^4} \sin\left(\frac{\pi j x}{L}\right) \sin\left(\frac{\pi j x'}{L}\right)
\end{aligned}
\tag{8.29}
$$

Finally we are in a good shape to calculate the solution explicitly by

$$
\begin{aligned}
y(x) &= \int_0^L G(x - x') g(x') dx' \\
&\approx \int_0^L c \left[\frac{2L^3}{\pi^4} \sum_{j=1}^{\infty} \frac{1}{j^4} \sin\left(\frac{\pi j x}{L}\right) \sin\left(\frac{\pi j x'}{L}\right) \right] dx' \\
&\approx \frac{2cL^3}{\pi^4} \sum_{j=1}^{\infty} \frac{1}{j^4} \sin\left(\frac{\pi j x}{L}\right) \left[\int_0^L \sin\left(\frac{\pi j x'}{L}\right) dx' \right] \\
&\approx \frac{4cL^4}{\pi^5} \sum_{j=1}^{\infty} \frac{1}{j^5} \sin\left(\frac{\pi j x}{L}\right) \sin^2\left(\frac{\pi j}{2}\right)
\end{aligned}
\tag{8.30}
$$

8.6   Finding Green's functions by Fourier analysis

If one is dealing with translation invariant systems of the form

$$
\mathscr{L}_x y(x) = f(x), \text{ with the differential operator}
$$

$$
\begin{aligned}
\mathscr{L}_x = a_n \frac{d^n}{dx^n} + a_{n-1} \frac{d^{n-1}}{dx^{n-1}} + \ldots + a_1 \frac{d}{dx} + a_0 \\
= \sum_{j=0}^{n} a_j \frac{d^j}{dx^j},
\end{aligned}
\tag{8.31}
$$

with constant coefficients a_j, then one can apply the following strategy using Fourier analysis to obtain the Green's function.

First, recall that, by Eq. (7.84) on page 153 the Fourier transform of the delta function $\widetilde{\delta}(k) = 1$ is just a constant. Therefore, δ can be written as

$$
\delta(x - x') = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ik(x - x')} dk
\tag{8.32}
$$

Next, consider the Fourier transform of the Green's function

$$
\widetilde{G}(k) = \int_{-\infty}^{\infty} G(x) e^{-ikx} dx
\tag{8.33}
$$

and its back transform

$$
G(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{G}(k) e^{ikx} dk.
\tag{8.34}
$$

Insertion of Eq. (8.34) into the *Ansatz* $\mathscr{L}_x G(x - x') = \delta(x - x')$ yields

$$
\begin{aligned}
\mathscr{L}_x G(x) = \mathscr{L}_x \frac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{G}(k) e^{ikx} dk = \frac{1}{2\pi} \int_{-\infty}^{\infty} \widetilde{G}(k) \left(\mathscr{L}_x e^{ikx} \right) dk \\
= \delta(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{ikx} dk.
\end{aligned}
\tag{8.35}
$$

and thus

$$\frac{1}{2\pi}\int_{-\infty}^{\infty}\left[\widetilde{G}(k)\mathscr{L}_x - 1\right]e^{ikx}dk = 0. \tag{8.36}$$

Therefore, the bracketed part of the integral kernel needs to vanish; and we obtain

$$\widetilde{G}(k)\mathscr{L}_k - 1 \equiv 0, \text{ or } \widetilde{G}(k) \equiv (\mathscr{L}_k)^{-1}, \tag{8.37}$$

where $\mathscr{L}_k$ is obtained from $\mathscr{L}_x$ by substituting every derivative $\frac{d}{dx}$ in the latter by ik in the former. As a result, the Fourier transform $\widetilde{G}(k)$ is obtained as a polynomial of degree n, the same degree as the highest order of derivative in $\mathscr{L}_x$.

In order to obtain the Green's function $G(x)$, and to be able to integrate over it with the inhomogenuous term $f(x)$, we have to Fourier transform $\widetilde{G}(k)$ back to $G(x)$.

Then we have to make sure that the solution obeys the initial conditions, and, if necessary, we have to add solutions of the homogenuos equation $\mathscr{L}_x G(x - x') = 0$. That is all.

Let us consider a few examples for this procedure.

1. First, let us solve the differential operator $y' - y = t$ on the intervall $[0, \infty)$ with the boundary conditions $y(0) = 0$.

We observe that the associated differential operator is given by

$$\mathscr{L}_t = \frac{d}{dt} - 1,$$

and the inhomogenuous term can be identified with $f(t) = t$.

We use the *Ansatz* $G_1(t, t') = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\tilde{G}_1(k)e^{ik(t-t')}dk$; hence

$$\begin{aligned}
\mathscr{L}_t G_1(t, t') &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\tilde{G}_1(k)\underbrace{\left(\frac{d}{dt}-1\right)e^{ik(t-t')}}_{=(ik-1)e^{ik(t-t')}}dk \\
&\qquad\qquad\qquad = (ik-1)e^{ik(t-t')} \tag{8.38}\\
&= \delta(t-t') = \frac{1}{2\pi}\int_{-\infty}^{+\infty}e^{ik(t-t')}dk
\end{aligned}$$

Now compare the kernels of the Fourier integrals of $\mathscr{L}_t G_1$ and δ:

$$\begin{aligned}
\tilde{G}_1(k)(ik-1) = 1 &\Longrightarrow \tilde{G}_1(k) = \frac{1}{ik-1} = \frac{1}{i(k+i)} \\
G_1(t, t') &= \frac{1}{2\pi}\int_{-\infty}^{+\infty}\frac{e^{ik(t-t')}}{i(k+i)}dk \tag{8.39}
\end{aligned}$$

The paths in the upper and lower integration plain are drawn in Frig. 8.1.

The "closures" throught the respective half-circle paths vanish. The residuum theorem yields

$$G_1(t, t') = \begin{cases} 0 & \text{for } t > t' \\ -2\pi i\,\mathrm{Res}\left(\frac{1}{2\pi i}\frac{e^{ik(t-t')}}{k+i}; -i\right) = -e^{t-t'} & \text{for } t < t'. \end{cases} \tag{8.40}$$

Note that $\int_{-\infty}^{\infty} f(x)\cos(kx)dk = -i\int_{-\infty}^{\infty} f(x)\sin(kx)dk$ cannot be satisfied for arbitrary x unless $f(x) = 0$.
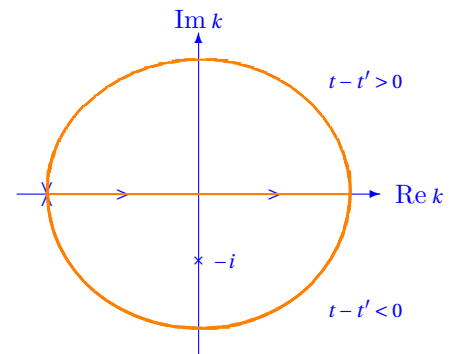
Figure 8.1: Plot of the two paths reqired for solving the Fourier integral.

Hence we obtain a Green's function for the inhomogenuous differential equation

$$G_1(t, t') = -H(t' - t)e^{t-t'}$$

However, this Green's function and its associated (special) solution does not obey the boundary conditions $G_1(0, t') = -H(t')e^{-t'} \neq 0$ for $t' \in [0, \infty)$.

Therefore, we have to fit the Green's function by adding an appropriately weighted solution to the homogenuos differential equation. The homogenuous Green's function is found by $\mathcal{L}_t G_0(t, t') = 0$, and thus, in particular, $\frac{d}{dt} G_0 = G_0 \Longrightarrow G_0 = ae^{t-t'}$. with the *Ansatz*

$$G(0, t') = G_1(0, t') + G_0(0, t'; a) = -H(t')e^{-t'} + ae^{-t'}$$

for the general solution we can choose the constant coefficient a so that

$$G(0, t') = G_1(0, t') + G_0(0, t'; a) = -H(t')e^{-t'} + ae^{-t'} = 0$$

For $a = 1$, the Green's function and thus the solution obeys the boundary value conditions; that is,

$$G(t, t') = \left[1 - H(t' - t)\right] e^{t-t'}.$$

Since $H(-x) = 1 - H(x)$, $G(t, t')$ can be rewritten as

$$G(t, t') = H(t - t')e^{t-t'}.$$

In the final step we obtain the solution through integration of G over the inhomogenuous term t:

$$
\begin{aligned}
y(t) &= \int_0^\infty \underbrace{H(t - t')}_{= 1 \text{ for } t' < t} e^{t-t'} t' dt' = \int_0^t e^{t-t'} t' dt' \\
&= e^t \int_0^t t' e^{-t'} dt' == e^t \left(-t'e^{-t'} \Big|_0^t - \int_0^t (-e^{-t'}) dt' \right) \\
&= e^t \left[(-te^{-t}) - e^{-t'} \Big|_0^t \right] = e^t \left(-te^{-t} - e^{-t} + 1 \right) = e^t - 1 - t.
\end{aligned}
\tag{8.41}
$$

It is prudent to check whether this is indeed a solution of the differential equation satisfying the boundary conditions:

$$
\mathcal{L}_t y(t) = \left(\frac{d}{dt} - 1 \right) \left(e^t - 1 - t \right) = e^t - 1 - \left(e^t - 1 - t \right) = t,
\tag{8.42}
$$
$$
\text{and } y(0) = e^0 - 1 - 0 = 0.
$$

2. Next, let us solve the differential equation $\frac{d^2 y}{dt^2} + y = \cos t$ on the interval $t \in [0, \infty)$ with the boundary conditions $y(0) = y'(0) = 0$.

First, observe that $\mathcal{L} = \frac{d^2}{dt^2} + 1$. The Fourier *Ansatz* for the Green's

function is

$$G_1(t, t') = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} \tilde{G}(k) e^{ik(t-t')} \, dk$$

$$\mathcal{L} G_1 = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} \tilde{G}(k) \left(\frac{d^2}{dt^2} + 1 \right) e^{ik(t-t')} \, dk$$

$$= \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} \tilde{G}(k) ((ik)^2 + 1) e^{ik(t-t')} \, dk \qquad (8.43)$$

$$= \delta(t - t') = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} e^{ik(t-t')} \, dk$$

Hence $\tilde{G}(k)(1 - k^2) = 1$ and thus $\tilde{G}(k) = \frac{1}{(1-k^2)} = \frac{-1}{(k+1)(k-1)}$. The Fourier transformation is

$$G_1(t, t') = -\frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} \frac{e^{ik(t-t')}}{(k+1)(k-1)} \, dk$$

$$= -\frac{1}{2\pi} 2\pi i \left[\mathrm{Res} \left(\frac{e^{ik(t-t')}}{(k+1)(k-1)}; k = 1 \right) \right. \qquad (8.44)$$

$$\left. + \mathrm{Res} \left(\frac{e^{ik(t-t')}}{(k+1)(k-1)}; k = -1 \right) \right] H(t - t')$$

The path in the upper integration plain is drawn in Fig. 8.2.
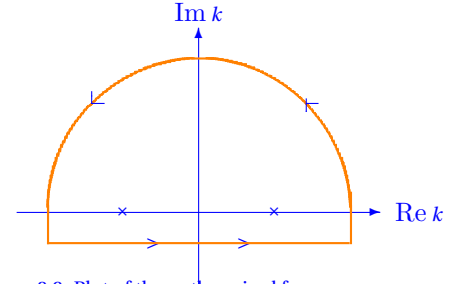
Figure 8.2: Plot of the path reqired for solving the Fourier integral.

$$G_1(t, t') = -\frac{i}{2} \left(e^{i(t-t')} - e^{-i(t-t')} \right) H(t - t')$$

$$= \frac{e^{i(t-t')} - e^{-i(t-t')}}{2i} H(t - t') = \sin(t - t') H(t - t')$$

$$G_1(0, t') = \sin(-t') H(-t') = 0 \text{ since } \quad t' > 0 \qquad (8.45)$$

$$G_1'(t, t') = \cos(t - t') H(t - t') + \underbrace{\sin(t - t') \delta(t - t')}_{= 0}$$

$$G_1'(0, t') = \cos(-t') H(-t') = 0.$$

G_1 already satisfies the boundary conditions; hence we do not need to find the Green's function G_0 of the homogenuous equation.

$$y(t) = \int\limits_0^\infty G(t, t') f(t') \, dt' = \int\limits_0^\infty \sin(t - t') \underbrace{H(t - t')}_{= 1 \text{ for } t > t'} \cos t' \, dt'$$

$$= \int\limits_0^t \sin(t - t') \cos t' \, dt' = \int\limits_0^t (\sin t \cos t' - \cos t \sin t') \cos t' \, dt'$$

$$= \int\limits_0^t \left[\sin t (\cos t')^2 - \cos t \sin t' \cos t' \right] dt' = \qquad (8.46)$$

$$= \sin t \int\limits_0^t (\cos t')^2 \, dt' - \cos t \int\limits_0^t \sin t' \cos t' \, dt'$$

$$= \sin t \left[\frac{1}{2} (t' + \sin t' \cos t') \right] \Big|_0^t - \cos t \left[\frac{\sin^2 t'}{2} \right] \Big|_0^t$$

$$= \frac{t \sin t}{2} + \frac{\sin^2 t \cos t}{2} - \frac{\cos t \sin^2 t}{2} = \frac{t \sin t}{2}.$$

Part III
Differential equations

9
Sturm-Liouville theory

This is only a very brief "dive into Sturm-Liouville theory," which has many fascinating aspects and connections to Fourier analysis, the special functions of mathematical physics, operator theory, and linear algebra [1]. In physics, many formalizations involve second order linear differential equations, which, in their most general form, can be written as [2]

$$\mathscr{L}_x y(x) = a_0(x) y(x) + a_1(x) \frac{d}{dx} y(x) + a_2(x) \frac{d^2}{dx^2} y(x) = f(x). \tag{9.1}$$

The differential operator associated with this differential equation is defined by

$$\mathscr{L}_x = a_0(x) + a_1(x) \frac{d}{dx} + a_2(x) \frac{d^2}{dx^2}. \tag{9.2}$$

The solutions $y(x)$ are often subject to boundary conditions of various forms:

- *Dirichlet boundary conditions* are of the form $y(a) = y(b) = 0$ for some a, b.

- *Neumann boundary conditions* are of the form $y'(a) = y'(b) = 0$ for some a, b.

- *Periodic boundary conditions* are of the form $y(a) = y(b)$ and $y'(a) = y'(b)$ for some a, b.

9.1 Sturm-Liouville form

Any second order differential equation of the general form (9.1) can be rewritten into a differential equation of the *Sturm-Liouville form*

$$\mathscr{S}_x y(x) = \frac{d}{dx} \left[p(x) \frac{d}{dx} \right] y(x) + q(x) y(x) = F(x),$$

$$\text{with } p(x) = e^{\int \frac{a_1(x)}{a_2(x)} dx},$$

$$q(x) = p(x) \frac{a_0(x)}{a_2(x)} = \frac{a_0(x)}{a_2(x)} e^{\int \frac{a_1(x)}{a_2(x)} dx},$$

$$F(x) = p(x) \frac{f(x)}{a_2(x)} = \frac{f(x)}{a_2(x)} e^{\int \frac{a_1(x)}{a_2(x)} dx} \tag{9.3}$$

[1] Garrett Birkhoff and Gian-Carlo Rota. *Ordinary Differential Equations.* John Wiley & Sons, New York, Chichester, Brisbane, Toronto, fourth edition, 1959, 1960, 1962, 1969, 1978, and 1989; M. A. Al-Gwaiz. *Sturm-Liouville Theory and its Applications.* Springer, London, 2008; and William Norrie Everitt. A catalogue of Sturm-Liouville differential equations. In Werner O. Amrein, Andreas M. Hinz, and David B. Pearson, editors, *Sturm-Liouville Theory, Past and Present*, pages 271–331. Birkhäuser Verlag, Basel, 2005. URL http://www.math.niu.edu/SL2/papers/birk0.pdf

[2] Russell Herman. *A Second Course in Ordinary Differential Equations: Dynamical Systems and Boundary Value Problems.* University of North Carolina Wilmington, Wilmington, NC, 2008. URL http://people.uncw.edu/hermanr/mat463/ODEBook/Book/ODE_LargeFont.pdf. Creative Commons Attribution-NoncommercialShare Alike 3.0 United States License

The associated differential operator

$$\mathscr{S}_x = \frac{d}{dx}\left[p(x)\frac{d}{dx}\right] + q(x)$$

$$= p(x)\frac{d^2}{dx^2} + p'(x)\frac{d}{dx} + q(x) \tag{9.4}$$

is called *Sturm-Liouville differential operator*.

For a proof, we insert $p(x)$, $q(x)$ and $F(x)$ into the Sturm-Liouville form of Eq. (9.3) and compare it with Eq. (9.1).

$$\left\{\frac{d}{dx}\left[e^{\int \frac{a_1(x)}{a_2(x)}dx}\frac{d}{dx}\right] + \frac{a_0(x)}{a_2(x)}e^{\int \frac{a_1(x)}{a_2(x)}dx}\right\}y(x) = \frac{f(x)}{a_2(x)}e^{\int \frac{a_1(x)}{a_2(x)}dx}$$

$$e^{\int \frac{a_1(x)}{a_2(x)}dx}\left\{\frac{d^2}{dx^2} + \frac{a_1(x)}{a_2(x)}\frac{d}{dx} + \frac{a_0(x)}{a_2(x)}\right\}y(x) = \frac{f(x)}{a_2(x)}e^{\int \frac{a_1(x)}{a_2(x)}dx} \tag{9.5}$$

$$\left\{a_2(x)\frac{d^2}{dx^2} + a_1(x)\frac{d}{dx} + a_0(x)\right\}y(x) = f(x).$$

9.2   *Sturm-Liouville eigenvalue problem*

The Sturm-Liouville eigenvalue problem is given by the differential equation

$$\mathscr{S}_x\phi(x) = -\lambda\rho(x)\phi(x), \text{ or}$$

$$\frac{d}{dx}\left[p(x)\frac{d}{dx}\right]\phi(x) + [q(x) + \lambda\rho(x)]\phi(x) = 0 \tag{9.6}$$

for $x \in (a, b)$ and continuous $p'(x)$, $q(x)$ and $p(x) > 0$, $\rho(x) > 0$.

We mention without proof (for proofs, see, for instance, Ref. [3]) that we can formulate a spectral theorem as follows

- the eigenvalues λ turn out to be real, countable, and ordered, and that there is a smallest eigenvalue λ_1 such that $\lambda_1 < \lambda_2 < \lambda_3 < \cdots$;

- for each eigenvalue λ_j there exists an eigenfunction $\phi_j(x)$ with $j - 1$ zeroes on (a, b);

- eigenfunctions corresponding to different eigenvalues are *orthogonal*, and can be normalized, with respect to the weight function $\rho(x)$; that is,

$$\langle \phi_j \mid \phi_k \rangle = \int_a^b \phi_j(x)\phi_k(x)\rho(x)dx = \delta_{jk} \tag{9.7}$$

- the set of eigenfunctions is *complete*; that is, any piecewise smooth function can be represented by

$$f(x) = \sum_{k=1}^{\infty} c_k\phi_k(x),$$

$$\text{with} \tag{9.8}$$

$$c_k = \frac{\langle f \mid \phi_k \rangle}{\langle \phi_k \mid \phi_k \rangle} = \langle f \mid \phi_k \rangle.$$

- the orthonormal (with respect to the weight ρ) set $\{\phi_j(x) \mid j \in \mathbb{N}\}$ is a *basis* of a Hilbert space with the inner product

$$\langle f \mid g \rangle = \int_a^b f(x)g(x)\rho(x)dx. \tag{9.9}$$

[3] M. A. Al-Gwaiz. *Sturm-Liouville Theory and its Applications*. Springer, London, 2008

9.3   Adjoint and self-adjoint operators

In operator theory, just as in matrix theory, we can define an *adjoint operator* (for finite dimensional Hilbert space, see Sec. 1.19 on page 43) *via* the scalar product defined in Eq. (9.9). In this formalization, the Sturm-Liouville differential operator $\mathscr{S}$ is self-adjoint.

Let us first define the *domain* of a differential operator $\mathscr{L}$ as the set of all square integrable (with respect to the weight $\rho(x)$) functions φ satisfying boundary conditions.

$$\int_a^b |\varphi(x)|^2 \rho(x)\,dx < \infty. \tag{9.10}$$

Then, the adjoint operator $\mathscr{L}^\dagger$ is defined by satisfying

$$\begin{aligned}
\langle \psi \mid \mathscr{L}\varphi \rangle &= \int_a^b \psi(x)[\mathscr{L}\varphi(x)]\rho(x)\,dx \\
&= \langle \mathscr{L}^\dagger \psi \mid \varphi \rangle = \int_a^b [\mathscr{L}^\dagger \psi(x)]\varphi(x)\rho(x)\,dx
\end{aligned} \tag{9.11}$$

for all $\psi(x)$ in the domain of $\mathscr{L}^\dagger$ and $\varphi(x)$ in the domain of $\mathscr{L}$.

Note that in the case of second order differential operators in the standard form (9.2) and with $\rho(x) = 1$, we can move the differential quotients and the entire differential operator in

$$\begin{aligned}
\langle \psi \mid \mathscr{L}\varphi \rangle &= \int_a^b \psi(x)[\mathscr{L}_x\varphi(x)]\rho(x)\,dx \\
&= \int_a^b \psi(x)[a_2(x)\varphi''(x) + a_1(x)\varphi'(x) + a_0(x)\varphi(x)]\,dx
\end{aligned} \tag{9.12}$$

from φ to ψ by one and two partial integrations.

Integrating the kernel $a_1(x)\varphi'(x)$ by parts yields

$$\int_a^b \psi(x)a_1(x)\varphi'(x)\,dx = \psi(x)a_1(x)\varphi(x)\big|_a^b - \int_a^b (\psi(x)a_1(x))'\varphi(x)\,dx. \tag{9.13}$$

Integrating the kernel $a_2(x)\varphi''(x)$ by parts twice yields

$$\begin{aligned}
\int_a^b \psi(x)a_2(x)\varphi''(x)\,dx &= \psi(x)a_2(x)\varphi'(x)\big|_a^b - \int_a^b (\psi(x)a_2(x))'\varphi'(x)\,dx \\
&= \psi(x)a_2(x)\varphi'(x)\big|_a^b - (\psi(x)a_2(x))'\varphi(x)\big|_a^b + \int_a^b (\psi(x)a_2(x))''\varphi(x)\,dx \\
&= \psi(x)a_2(x)\varphi'(x) - (\psi(x)a_2(x))'\varphi(x)\big|_a^b + \int_a^b (\psi(x)a_2(x))''\varphi(x)\,dx.
\end{aligned} \tag{9.14}$$

Combining these two calculations yields

$$\begin{aligned}
\langle \psi \mid \mathscr{L}\varphi \rangle &= \int_a^b \psi(x)[\mathscr{L}_x\varphi(x)]\rho(x)\,dx \\
&= \int_a^b \psi(x)[a_2(x)\varphi''(x) + a_1(x)\varphi'(x) + a_0(x)\varphi(x)]\,dx \\
&= \psi(x)a_1(x)\varphi(x) + \psi(x)a_2(x)\varphi'(x) - (\psi(x)a_2(x))'\varphi(x)\big|_a^b \\
&\quad + \int_a^b [(a_2(x)\psi(x))'' - (a_1(x)\psi(x))' + a_0(x)\psi(x)]\varphi(x)\,dx.
\end{aligned} \tag{9.15}$$

If the terms with no integral vanish because of boundary conditions or other reasons, such as $\varphi(x) = \psi(x)$ and $a_1(x) = a_2'(x)$ in the case of the Sturm-Liouville operator $\mathscr{S}_x$; that is, if

$$\psi(x)a_1(x)\varphi(x) + \psi(x)a_2(x)\varphi'(x) - (\psi(x)a_2(x))'\varphi(x)\big|_a^b = 0, \qquad (9.16)$$

then Eq. (9.15) reduces to

$$\langle \psi \mid \mathscr{L}\varphi \rangle = \int_a^b [(a_2(x)\psi(x))'' - (a_1(x)\psi(x))' + a_0(x)\psi(x)]\varphi(x)\,dx, \quad (9.17)$$

and we can identify the adjoint differential operator of $\mathscr{L}_x$ with

$$
\begin{aligned}
\mathscr{L}_x^\dagger &= \frac{d^2}{dx^2} a_2(x) - \frac{d}{dx} a_1(x) + a_0(x) \\
&= \frac{d}{dx}\left[a_2(x)\frac{d}{dx} + a_2'(x)\right] - a_1'(x) - a_1(x)\frac{d}{dx} + a_0(x) \\
&= a_2'(x)\frac{d}{dx} + a_2(x)\frac{d^2}{dx^2} + a_2''(x) + a_2'(x)\frac{d}{dx} - a_1'(x) - a_1(x)\frac{d}{dx} + a_0(x) \\
&= a_2(x)\frac{d^2}{dx^2} + [2a_2'(x) - a_1(x)]\frac{d}{dx} + a_2''(x) - a_1'(x) + a_0(x).
\end{aligned}
$$
$$(9.18)$$

The operator $\mathscr{L}_x$ is called self-adjoint if

$$\mathscr{L}_x^\dagger = \mathscr{L}_x. \qquad (9.19)$$

Next we shall show that, in particular, the Sturm-Liouville differential operator (9.4) is self-adjoint, and that all second order differential operators [with the boundary condition (9.16)] which are self-adjoint are of the Sturm-Liouville form.

In order to prove that the Sturm-Liouville differential operator

$$\mathscr{S} = \frac{d}{dx}\left[p(x)\frac{d}{dx}\right] + q(x) = p(x)\frac{d^2}{dx^2} + p'(x)\frac{d}{dx} + q(x) \qquad (9.20)$$

from Eq. (9.4) is self-adjoint, we verify Eq. (9.19) with $\mathscr{S}^\dagger$ taken from Eq. (9.18). Thereby, we identify $a_2(x) = p(x)$, $a_1(x) = p'(x)$, and $a_0(x) = q(x)$; hence

$$
\begin{aligned}
\mathscr{S}_x^\dagger &= a_2(x)\frac{d^2}{dx^2} + [2a_2'(x) - a_1(x)]\frac{d}{dx} + a_2''(x) - a_1'(x) + a_0(x) \\
&= p(x)\frac{d^2}{dx^2} + [2p'(x) - p'(x)]\frac{d}{dx} + p''(x) - p''(x) + q(x) \\
&= p(x)\frac{d^2}{dx^2} + p'(x)\frac{d}{dx} q(x) \\
&= \mathscr{S}_x.
\end{aligned}
$$
$$(9.21)$$

Alternatively we could argue from Eqs. (9.18) and (9.19), noting that a differential operator is self-adjoint if and only if

$$\mathscr{L}_x = a_2(x)\frac{d^2}{dx^2} + a_1(x)\frac{d}{dx} + a_0(x)$$
$$(9.22)$$
$$= \mathscr{L}_x^\dagger = a_2(x)\frac{d^2}{dx^2} + [2a_2'(x) - a_1(x)]\frac{d}{dx} + a_2''(x) - a_1'(x) + a_0(x).$$

By comparison of the coefficients,

$$a_2(x) = a_2(x),$$
$$a_1(x) = 2a_2'(x) - a_1(x), \tag{9.23}$$
$$a_0(x) = +a_2''(x) - a_1'(x) + a_0(x),$$

and hence,

$$a_2'(x) = a_1(x), \tag{9.24}$$

which is exactly the form of the Sturm-Liouville differential operator.

9.4 Sturm-Liouville transformation into Liouville normal form

Let, for $x \in [a, b]$,

$$[\mathscr{S}_x + \lambda \rho(x)] y(x) = 0,$$
$$\frac{d}{dx}\left[p(x)\frac{d}{dx}\right]y(x) + [q(x) + \lambda\rho(x)]y(x) = 0,$$
$$\left[p(x)\frac{d^2}{dx^2} + p'(x)\frac{d}{dx} + q(x) + \lambda\rho(x)\right]y(x) = 0, \tag{9.25}$$
$$\left[\frac{d^2}{dx^2} + \frac{p'(x)}{p(x)}\frac{d}{dx} + \frac{q(x) + \lambda\rho(x)}{p(x)}\right]y(x) = 0$$

be a second order differential equation of the Sturm-Liouville form [4].

This equation (9.25) can be written in the *Liouville normal form* containing no first order differentiation term

$$-\frac{d^2}{dt^2}w(t) + [\hat{q}(t) - \lambda]w(t) = 0, \text{ with } t \in [t(a), t(b)]. \tag{9.26}$$

It is obtained *via* the *Sturm-Liouville transformation*

$$\xi = t(x) = \int_a^x \sqrt{\frac{\rho(s)}{p(s)}}\,ds, \tag{9.27}$$
$$w(t) = \sqrt[4]{p(x(t))\rho(x(t))}\,y(x(t)),$$

where

$$\hat{q}(t) = \frac{1}{\rho}\left[-q - \sqrt[4]{p\rho}\left(p\left(\frac{1}{\sqrt[4]{p\rho}}\right)'\right)'\right]. \tag{9.28}$$

The apostrophe represents derivation with respect to x.

For the sake of an example, suppose we want to know the normalized eigenfunctions of

$$x^2 y'' + 3x y' + y = -\lambda y, \text{ with } x \in [1, 2] \tag{9.29}$$

with the boundary conditions $y(1) = y(2) = 0$.

The first thing we have to do is to transform this differential equation into its Sturm-Liouville form by identifying $a_2(x) = x^2$, $a_1(x) = 3x$, $a_0 = 1$, $\rho = 1$ such that $f(x) = -\lambda y(x)$; and hence

$$p(x) = e^{\int \frac{3x}{x^2}dx} = e^{\int \frac{3}{x}dx} = e^{3\log x} = x^3,$$
$$q(x) = p(x)\frac{1}{x^2} = x, \tag{9.30}$$
$$F(x) = p(x)\frac{\lambda y}{(-x^2)} = -\lambda x y, \text{ and hence } \rho(x) = x.$$

[4] Garrett Birkhoff and Gian-Carlo Rota. *Ordinary Differential Equations.* John Wiley & Sons, New York, Chichester, Brisbane, Toronto, fourth edition, 1959, 1960, 1962, 1969, 1978, and 1989

As a result we obtain the Sturm-Liouville form

$$\frac{1}{x}\left((x^3 y')' + xy\right) = -\lambda y. \tag{9.31}$$

In the next step we apply the Sturm-Liouville transformation

$$\xi = t(x) = \int \sqrt{\frac{\rho(x)}{p(x)}}\,dx = \int \frac{dx}{x} = \log x,$$

$$w(t(x)) = \sqrt[4]{p(x(t))\rho(x(t))}\,y(x(t)) = \sqrt[4]{x^4}\,y(x(t)) = xy, \tag{9.32}$$

$$\hat{q}(t) = \frac{1}{x}\left[-x - \sqrt[4]{x^4}\left(x^3\left(\frac{1}{\sqrt[4]{x^4}}\right)'\right)'\right] = 0.$$

We now take the Ansatz $y = \frac{1}{x}w(t(x)) = \frac{1}{x}w(\log x)$ and finally obtain the Liouville normal form

$$-w'' = \lambda w. \tag{9.33}$$

As an *Ansatz* for solving the Liouville normal form we use

$$w(\xi) = a\sin(\sqrt{\lambda}\xi) + b\cos(\sqrt{\lambda}\xi) \tag{9.34}$$

The boundary conditions translate into $x = 1 \to \xi = 0$, and $x = 2 \to \xi = \log 2$. From $w(0) = 0$ we obtain $b = 0$. From $w(\log 2) = a\sin(\sqrt{\lambda}\log 2) = 0$ we obtain $\sqrt{\lambda_n}\log 2 = n\pi$.

Thus the eigenvalues are

$$\lambda_n = \left(\frac{n\pi}{\log 2}\right)^2. \tag{9.35}$$

The associated eigenfunctions are

$$w_n(\xi) = a\sin\left[\frac{n\pi}{\log 2}\xi\right], \tag{9.36}$$

and thus

$$y_n = \frac{1}{x}a\sin\left[\frac{n\pi}{\log 2}\log x\right]. \tag{9.37}$$

We can check that they are orthonormal by inserting into Eq. (9.7) and verifying it; that is,

$$\int_1^2 \rho(x)y_n(x)y_m(x)\,dx = \delta_{nm}; \tag{9.38}$$

more explicitly,

$$\int_1^2 dx\,x\left(\frac{1}{x^2}\right)a^2\sin\left(n\pi\frac{\log x}{\log 2}\right)\sin\left(m\pi\frac{\log x}{\log 2}\right)$$

$$[\text{variable substitution } u = \frac{\log x}{\log 2}$$

$$\frac{du}{dx} = \frac{1}{\log 2}\frac{1}{x},\ u = \frac{dx}{x\log 2}]$$

$$= \int_{u=0}^{u=1} du\,\log 2\,a^2\sin(n\pi u)\sin(m\pi u) \tag{9.39}$$

$$= \underbrace{a^2\left(\frac{\log 2}{2}\right)}_{=1}\underbrace{2\int_0^1 du\,\sin(n\pi u)\sin(m\pi u)}_{=\delta_{nm}}$$

$$= \delta_{nm}.$$

Finally, with $a = \sqrt{\frac{2}{\log 2}}$ we obtain the solution

$$y_n = \sqrt{\frac{2}{\log 2}} \frac{1}{x} \sin\left(n\pi \frac{\log x}{\log 2}\right). \qquad (9.40)$$

9.5   Varieties of Sturm-Liouville differential equations

A catalogue of Sturm-Liouville differential equations comprises the following *species*, among many others [5]. Some of these cases are tabellated as functions p, q, λ and ρ appearing in the general form of the Sturm-Liouville eigenvalue problem (9.6)

$$\mathcal{S}_x \phi(x) = -\lambda \rho(x)\phi(x), \text{ or}$$

$$\frac{d}{dx}\left[p(x)\frac{d}{dx}\right]\phi(x) + [q(x) + \lambda\rho(x)]\phi(x) = 0 \qquad (9.41)$$

in Table 9.1.

Equation	$p(x)$	$q(x)$	$-\lambda$	$\rho(x)$
Hypergeometric	$x^{\alpha+1}(1-x)^{\beta+1}$	0	μ	$x^\alpha(1-x)^\beta$
Legendre	$1-x^2$	0	$l(l+1)$	1
Shifted Legendre	$x(1-x)$	0	$l(l+1)$	1
Associated Legendre	$1-x^2$	$-\frac{m^2}{1-x^2}$	$l(l+1)$	1
Chebyshev I	$\sqrt{1-x^2}$	0	n^2	$\frac{1}{\sqrt{1-x^2}}$
Shifted Chebyshev I	$\sqrt{x(1-x)}$	0	n^2	$\frac{1}{\sqrt{x(1-x)}}$
Chebyshev II	$(1-x^2)^{\frac{3}{2}}$	0	$n(n+2)$	$\sqrt{1-x^2}$
Ultraspherical (Gegenbauer)	$(1-x^2)^{\alpha+\frac{1}{2}}$	0	$n(n+2\alpha)$	$(1-x^2)^{\alpha-\frac{1}{2}}$
Bessel	x	$-\frac{n^2}{x}$	a^2	x
Laguerre	xe^{-x}	0	α	e^{-x}
Associated Laguerre	$x^{k+1}e^{-x}$	0	$\alpha - k$	$x^k e^{-x}$
Hermite	xe^{-x^2}	0	2α	e^{-x}
Fourier (harmonic oscillator)	1	0	k^2	1
Schrödinger (hydrogen atom)	1	$l(l+1)x^{-2}$	μ	1

Table 9.1: Some varieties of differential equations expressible as Sturm-Liouville differential equations

10
Separation of variables

This chapter deals with the ancient alchemic suspicion of *"solve et co-agula"* that it is possible to solve a problem by splitting it up into partial problems, solving these issues separately; and consecutively joining together the partial solutions, thereby yielding the full answer to the problem – translated into the context of *partial differential equations;* that is, equations with derivatives of more than one variable. Thereby, solving the separate partial problems is not dissimilar to applying subprograms from some program library.

Already Descartes mentioned this sort of method in his *Discours de la méthode pour bien conduire sa raison et chercher la verité dans les sciences* (English translation: *Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth*) [1] stating that (in a newer translation [2])

> [Rule Five:] *The whole method consists entirely in the ordering and arranging of the objects on which we must concentrate our mind's eye if we are to discover some truth . We shall be following this method exactly if we first reduce complicated and obscure propositions step by step to simpler ones, and then, starting with the intuition of the simplest ones of all, try to ascend through the same steps to a knowledge of all the rest. … [Rule Thirteen:] If we perfectly understand a problem we must abstract it from every superfluous conception, reduce it to its simplest terms and, by means of an enumeration, divide it up into the smallest possible parts.*

[1] Rene Descartes. *Discours de la méthode pour bien conduire sa raison et chercher la verité dans les sciences (Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth)*. 1637. URL http://www.gutenberg.org/etext/59

[2] Rene Descartes. *The Philosophical Writings of Descartes. Volume 1.* Cambridge University Press, Cambridge, 1985. translated by John Cottingham, Robert Stoothoff and Dugald Murdoch

The method of separation of variables is one among a couple of strategies to solve differential equations [3], and it is a very important one in physics.

Separation of variables can be applied whenever we have no "mixtures of derivatives and functional dependencies;" more specifically, whenever the partial differential equation can be written as a sum

[3] Lawrence C. Evans. *Partial differential equations. Graduate Studies in Mathematics, volume 19.* American Mathematical Society, Providence, Rhode Island, 1998; and Klaus Jänich. *Analysis für Physiker und Ingenieure. Funktionentheorie, Differentialgleichungen, Spezielle Funktionen.* Springer, Berlin, Heidelberg, fourth edition, 2001. URL http://www.springer.com/mathematics/analysis/book/978-3-540-41985-3

$$\mathscr{L}_{x,y}\psi(x,y) = (\mathscr{L}_x + \mathscr{L}_y)\psi(x,y) = 0, \text{ or}$$
$$\mathscr{L}_x\psi(x,y) = -\mathscr{L}_y\psi(x,y). \tag{10.1}$$

Because in this case we may make a *multiplicative Ansatz*

$$\psi(x,y) = v(x)u(y). \tag{10.2}$$

Inserting (10.2) into (10) effectively separates the variable dependencies

$$
\begin{aligned}
\mathcal{L}_x v(x) u(y) &= -\mathcal{L}_y v(x) u(y), \\
u(y) \left[\mathcal{L}_x v(x) \right] &= -v(x) \left[\mathcal{L}_y u(y) \right], \\
\frac{1}{v(x)} \mathcal{L}_x v(x) &= -\frac{1}{u(y)} \mathcal{L}_y u(y) = a,
\end{aligned}
\tag{10.3}
$$

with constant a, because $\frac{\mathcal{L}_x v(x)}{v(x)}$ does not depend on x, and $\frac{\mathcal{L}_y u(y)}{u(y)}$ does not depend on y. Therefore, neither side depends on x or y; hence both sides are constants.

As a result, we can treat and integrate both sides separately; that is,

$$
\begin{aligned}
\frac{1}{v(x)} \mathcal{L}_x v(x) &= a, \\
\frac{1}{u(y)} \mathcal{L}_y u(y) &= -a,
\end{aligned}
\tag{10.4}
$$

or

$$
\begin{aligned}
\mathcal{L}_x v(x) - a v(x) &= 0, \\
\mathcal{L}_y u(y) + a u(y) &= 0.
\end{aligned}
\tag{10.5}
$$

This separation of variable *Ansatz* can be often used when the *Laplace operator* $\Delta = \nabla \cdot \nabla$ is involved, since there the partial derivatives with respect to different variables occur in different summands.

The general solution is a linear combination (superposition) of the products of all the linear independent solutions – that is, the sum of the products of all separate (linear independent) solutions, weighted by an arbitrary scalar factor.

For the sake of demonstration, let us consider a few examples.

If we would just consider a single product of all general one parameter solutions we would run into the same problem as in the entangled case on page 24 – we could not cover all the solutions of the original equation.

1. Let us separate the homogenuous Laplace differential equation

$$
\Delta \Phi = \frac{1}{u^2 + v^2} \left(\frac{\partial^2 \Phi}{\partial u^2} + \frac{\partial^2 \Phi}{\partial v^2} \right) + \frac{\partial^2 \Phi}{\partial z^2} = 0
\tag{10.6}
$$

in parabolic cylinder coordinates (u, v, z) with $\mathbf{x} = \left(\frac{1}{2}(u^2 - v^2), uv, z \right)$.

The separation of variables *Ansatz* is

$$
\Phi(u, v, z) = \Phi_1(u) \Phi_2(v) \Phi_3(z).
$$

Then,

$$
\frac{1}{u^2 + v^2} \left(\Phi_2 \Phi_3 \frac{\partial^2 \Phi_1}{\partial u^2} + \Phi_1 \Phi_3 \frac{\partial^2 \Phi_2}{\partial v^2} \right) + \Phi_1 \Phi_2 \frac{\partial^2 \Phi}{\partial z^2} = 0
$$

$$
\frac{1}{u^2 + v^2} \left(\frac{\Phi_1''}{\Phi_1} + \frac{\Phi_2''}{\Phi_2} \right) = -\frac{\Phi_3''}{\Phi_3} = \lambda = \text{const.}
$$

λ is constant because it does neither depend on u, v [because of the right hand side $\Phi_3''(z)/\Phi_3(z)$], nor on z (because of the left hand side). Furthermore,

$$
\frac{\Phi_1''}{\Phi_1} - \lambda u^2 = -\frac{\Phi_2''}{\Phi_2} + \lambda v^2 = l^2 = \text{const.}
$$

with constant l for analoguous reasons. The three resulting differential equations are

$$
\begin{aligned}
\Phi_1'' - (\lambda u^2 + l^2) \Phi_1 &= 0, \\
\Phi_2'' - (\lambda v^2 - l^2) \Phi_2 &= 0, \\
\Phi_3'' + \lambda \Phi_3 &= 0.
\end{aligned}
$$

2. Let us separate the homogenuous (i) Laplace, (ii) wave, and (iii) diffusion equations, in elliptic cylinder coordinates (u, v, z) with $\vec{x} = (a \cosh u \cos v, a \sinh u \sin v, z)$ and

$$\Delta = \frac{1}{a^2 (\sinh^2 u + \sin^2 v)} \left[\frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2} \right] + \frac{\partial^2}{\partial z^2}.$$

ad (i):

Again the separation of variables *Ansatz* is $\Phi(u, v, z) = \Phi_1(u)\Phi_2(v)\Phi_3(z)$. Hence,

$$\frac{1}{a^2 (\sinh^2 u + \sin^2 v)} \left(\Phi_2 \Phi_3 \frac{\partial^2 \Phi_1}{\partial u^2} + \Phi_1 \Phi_3 \frac{\partial^2 \Phi_2}{\partial v^2} \right) = -\Phi_1 \Phi_2 \frac{\partial^2 \Phi}{\partial z^2},$$

$$\frac{1}{a^2 (\sinh^2 u + \sin^2 v)} \left(\frac{\Phi_1''}{\Phi_1} + \frac{\Phi_2''}{\Phi_2} \right) = -\frac{\Phi_3''}{\Phi_3} = k^2 = \text{const.} \Longrightarrow \Phi_3'' + k^2 \Phi_3 = 0$$

$$\frac{\Phi_1''}{\Phi_1} + \frac{\Phi_2''}{\Phi_2} = k^2 a^2 (\sinh^2 u + \sin^2 v),$$

$$\frac{\Phi_1''}{\Phi_1} - k^2 a^2 \sinh^2 u = -\frac{\Phi_2''}{\Phi_2} + k^2 a^2 \sin^2 v = l^2,$$

$$(10.7)$$

and finally,

$$\begin{aligned} \Phi_1'' &- (k^2 a^2 \sinh^2 u + l^2)\Phi_1 &= 0, \\ \Phi_2'' &- (k^2 a^2 \sin^2 v - l^2)\Phi_2 &= 0. \end{aligned}$$

ad (ii):

the wave equation is given by

$$\Delta \Phi = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2}.$$

Hence,

$$\frac{1}{a^2 (\sinh^2 u + \sin^2 v)} \left(\frac{\partial^2}{\partial u^2} + \frac{\partial^2}{\partial v^2} + \right) \Phi + \frac{\partial^2 \Phi}{\partial z^2} = \frac{1}{c^2} \frac{\partial^2 \Phi}{\partial t^2}.$$

The separation of variables *Ansatz* is $\Phi(u, v, z, t) = \Phi_1(u)\Phi_2(v)\Phi_3(z)T(t)$

$$\Longrightarrow \frac{1}{a^2 (\sinh^2 u + \sin^2 v)} \left(\frac{\Phi_1''}{\Phi_1} + \frac{\Phi_2''}{\Phi_2} \right) + \frac{\Phi_3''}{\Phi_3} = \frac{1}{c^2} \frac{T''}{T} = -\omega^2 = \text{const.},$$

$$\frac{1}{c^2} \frac{T''}{T} = -\omega^2 \Longrightarrow T'' + c^2 \omega^2 T = 0,$$

$$\frac{1}{a^2 (\sinh^2 u + \sin^2 v)} \left(\frac{\Phi_1''}{\Phi_1} + \frac{\Phi_2''}{\Phi_2} \right) = -\frac{\Phi_3''}{\Phi_3} - \omega^2 = k^2, \quad (10.8)$$

$$\Phi_3'' + (\omega^2 + k^2)\Phi_3 = 0$$

$$\frac{\Phi_1''}{\Phi_1} + \frac{\Phi_2''}{\Phi_2} = k^2 a^2 (\sinh^2 u + \sin^2 v)$$

$$\frac{\Phi_1''}{\Phi_1} - a^2 k^2 \sinh^2 u = -\frac{\Phi_2''}{\Phi_2} + a^2 k^2 \sin^2 v = l^2,$$

and finally,

$$\begin{aligned} \Phi_1'' - (k^2 a^2 \sinh^2 u + l^2)\Phi_1 = 0, \\ \Phi_2'' - (k^2 a^2 \sin^2 v - l^2)\Phi_2 = 0. \end{aligned} \quad (10.9)$$

ad (iii):

The diffusion equation is $\Delta\Phi = \frac{1}{D}\frac{\partial\Phi}{\partial t}$.

The separation of variables *Ansatz* is $\Phi(u,v,z,t) = \Phi_1(u)\Phi_2(v)\Phi_3(z)T(t)$. Let us take the result of (i), then

$$\frac{1}{a^2(\sinh^2 u + \sin^2 v)}\left(\frac{\Phi_1''}{\Phi_1} + \frac{\Phi_2''}{\Phi_2}\right) + \frac{\Phi_3''}{\Phi_3} = \frac{1}{D}\frac{T'}{T} = -\alpha^2 = \text{const.}$$

$$T = Ae^{-\alpha^2 Dt} \quad (10.10)$$

$$\Phi_3'' + (\alpha^2 + k^2)\Phi_3 = 0 \implies \Phi_3'' = -(\alpha^2 + k^2)\Phi_3 \implies \Phi_3 = Be^{i\sqrt{\alpha^2+k^2}\,z}$$

and finally,

$$\begin{aligned}
\Phi_1'' - (\alpha^2 k^2 \sinh^2 u + l^2)\Phi_1 &= 0 \\
\Phi_2'' - (\alpha^2 k^2 \sin^2 v - l^2)\Phi_2 &= 0.
\end{aligned} \quad (10.11)$$

11

Special functions of mathematical physics

Special functions often arise as solutions of differential equations; for instance as eigenfunctions of differential operators in quantum mechanics. Sometimes they occur after several *separation of variables* and substitution steps have transformed the physical problem into something manageable. For instance, we might start out with some linear partial differential equation like the wave equation, then separate the space from time coordinates, then separate the radial from the angular components, and finally separate the two angular parameters. After we have done that, we end up with several separate differential equations of the Liouville form; among them the Legendre differential equation leading us to the Legendre polynomials.

In what follows, a particular class of special functions will be considered. These functions are all special cases of the *hypergeometric function*, which is the solution of the *hypergeometric equation.* The hypergeometric function exhibits a high degree of "plasticity," as many elementary analytic functions can be expressed by it.

First, as a prerequisite, let us define the gamma function. Then we proceed to second order Fuchsian differential equations; followed by rewriting a Fuchsian differential equation into a hypergeometric equation. Then we study the hypergeometric function as a solution to the hypergeometric equation. Finally, we mention some particular hypergeometric functions, such as the Legendre orthogonal polynomials, and others.

Again, if not mentioned otherwise, we shall restrict our attention to second order differential equations. Sometimes – such as for the Fuchsian class – a generalization is possible but not very relevant for physics.

11.1 Gamma function

The gamma function $\Gamma(x)$ is an extension of the factorial function $n!$, because it generalizes the factorial to real or complex arguments (different from the negative integers and from zero); that is,

$$\Gamma(n+1) = n! \text{ for } n \in \mathbb{N}, \text{ or } \Gamma(n) = (n-1)! \text{ for } n \in \mathbb{N} - 0. \tag{11.1}$$

Let us first define the *shifted factorial* or, by another naming, the

This chapter follows several approaches:

N. N. Lebedev. *Special Functions and Their Applications*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1965. R. A. Silverman, translator and editor; reprinted by Dover, New York, 1972; Herbert S. Wilf. *Mathematics for the physical sciences*. Dover, New York, 1962. URL http://www.math.upenn.edu/~wilf/website/Mathematics_for_the_Physical_Sciences.html; W. W. Bell. *Special Functions for Scientists and Engineers*. D. Van Nostrand Company Ltd, London, 1968; Nico M. Temme. *Special functions: an introduction to the classical functions of mathematical physics*. John Wiley & Sons, Inc., New York, 1996. ISBN 0-471-11313-1; Nico M. Temme. Numerical aspects of special functions. *Acta Numerica*, 16:379–478, 2007. ISSN 0962-4929. DOI: 10.1017/S0962492904000077. URL http://doi.org/10.1017/S0962492904000077; George E. Andrews, Richard Askey, and Ranjan Roy. *Special Functions*, volume 71 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, Cambridge, 1999. ISBN 0-521-62321-9; Vadim Kuznetsov. Special functions and their symmetries. Part I: Algebraic and analytic methods. Postgraduate Course in Applied Analysis, May 2003. URL http://www1.maths.leeds.ac.uk/~kisilv/courses/sp-funct.pdf; and Vladimir Kisil. Special functions and their symmetries. Part II: Algebraic and symmetry methods. Postgraduate Course in Applied Analysis, May 2003. URL http://www1.maths.leeds.ac.uk/~kisilv/courses/sp-repr.pdf
For reference, consider

Milton Abramowitz and Irene A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Number 55 in National Bureau of Standards Applied Mathematics Series. U.S. Government Printing Office, Washington, D.C., 1964. URL http://www.math.sfu.ca/~cbm/aands/; Yuri Alexandrovich Brychkov and Anatolii Platonovich Prudnikov. *Handbook of special functions: derivatives, integrals, series and other formulas*. CRC/Chapman & Hall Press, Boca Raton, London, New York, 2008; and I. S. Gradshteyn and I. M. Ryzhik. *Tables of Integrals, Series, and Products, 6th ed*. Academic Press, San

Pochhammer symbol

$$(a)_0 \stackrel{\text{def}}{=} 1,$$

$$(a)_n \stackrel{\text{def}}{=} a(a+1)\cdots(a+n-1) = \frac{\Gamma(a+n)}{\Gamma(a)}, \tag{11.2}$$

where $n > 0$ and a can be any real or complex number.

With this definition,

$$\begin{aligned}
z!(z+1)_n &= 1 \cdot 2 \cdots z \cdot (z+1)((z+1)+1)\cdots((z+1)+n-1) \\
&= 1 \cdot 2 \cdots z \cdot (z+1)(z+2)\cdots(z+n) \\
&= (z+n)!, \\
\text{or } z! &= \frac{(z+n)!}{(z+1)_n}.
\end{aligned} \tag{11.3}$$

Since

$$\begin{aligned}
(z+n)! &= (n+z)! \\
&= 1 \cdot 2 \cdots n \cdot (n+1)(n+2)\cdots(n+z) \\
&= n! \cdot (n+1)(n+2)\cdots(n+z) \\
&= n!(n+1)_z,
\end{aligned} \tag{11.4}$$

we can rewrite Eq. (11.3) into

$$z! = \frac{n!(n+1)_z}{(z+1)_n} = \frac{n! n^z}{(z+1)_n} \frac{(n+1)_z}{n^z}. \tag{11.5}$$

Since the latter factor, for large n, converges as ["$O(x)$" means "of the order of x"]

$$\begin{aligned}
\frac{(n+1)_z}{n^z} &= \frac{(n+1)((n+1)+1)\cdots((n+1)+z-1)}{n^z} \\
&= \frac{n^z + O(n^{z-1})}{n^z} \\
&= \frac{n^z}{n^z} + \frac{O(n^{z-1})}{n^z} \\
&= 1 + O(n^{-1}) \stackrel{n \to \infty}{\longrightarrow} 1,
\end{aligned} \tag{11.6}$$

in this limit, Eq. (11.5) can be written as

$$z! = \lim_{n \to \infty} z! = \lim_{n \to \infty} \frac{n! n^z}{(z+1)_n}. \tag{11.7}$$

Hence, for all $z \in \mathbb{C}$ which are not equal to a negative integer – that is, $z \notin \{-1, -2, \ldots\}$ – we can, in analogy to the "classical factorial," define a "factorial function shifted by one" as

$$\Gamma(z+1) \stackrel{\text{def}}{=} \lim_{n \to \infty} \frac{n! n^z}{(z+1)_n}, \tag{11.8}$$

and thus, because for very large n and constant z (i.e., $z \ll n$), $(z + n) \approx n$,

$$
\begin{aligned}
\Gamma(z) &= \lim_{n \to \infty} \frac{n! \, n^{z-1}}{(z)_n} \\
&= \lim_{n \to \infty} \frac{n! \, n^{z-1}}{z(z+1) \cdots (z+n-1)} \\
&= \lim_{n \to \infty} \frac{n! \, n^{z-1}}{z(z+1) \cdots (z+n-1)} \underbrace{\left(\frac{z+n}{z+n} \right)}_{1} \\
&= \lim_{n \to \infty} \frac{n! \, n^{z-1}(z+n)}{z(z+1) \cdots (z+n)} \\
&= \frac{1}{z} \lim_{n \to \infty} \frac{n! \, n^z}{(z+1)_n}.
\end{aligned}
\tag{11.9}
$$

$\Gamma(z + 1)$ has thus been redefined in terms of $z!$ in Eq. (11.3), which, by comparing Eqs. (11.8) and (11.9), also implies that

$$
\Gamma(z+1) = z\Gamma(z). \tag{11.10}
$$

Note that, since

$$
(1)_n = 1(1+1)(1+2) \cdots (1+n-1) = n!, \tag{11.11}
$$

Eq. (11.8) yields

$$
\Gamma(1) = \lim_{n \to \infty} \frac{n! \, n^0}{(1)_n} = \lim_{n \to \infty} \frac{n!}{n!} = 1. \tag{11.12}
$$

By induction, Eqs. (11.12) and (11.10) yield $\Gamma(n+1) = n!$ for $n \in \mathbb{N}$.

We state without proof that, for complex numbers z with positive real parts $\Re z > 0$, $\Gamma(z)$ can be defined by an integral representation as

$$
\Gamma(z) \stackrel{\text{def}}{=} \int_0^\infty t^{z-1} e^{-t} dt. \tag{11.13}
$$

Note that Eq. (11.10) can be derived from this integral representation of $\Gamma(z)$ by partial integration; that is,

$$
\begin{aligned}
\Gamma(z+1) &= \int_0^\infty t^z e^{-t} dt \\
&= \underbrace{-t^z e^{-t} \big|_0^\infty}_{0} - \left[-\int_0^\infty \left(\frac{d}{dt} t^z \right) e^{-t} dt \right] \\
&= \int_0^\infty z t^{z-1} e^{-t} dt \\
&= z \int_0^\infty t^{z-1} e^{-t} dt = z\Gamma(z).
\end{aligned}
\tag{11.14}
$$

We also mention without proof the following formulæ:

$$
\Gamma\left(\frac{1}{2} \right) = \sqrt{\pi}, \tag{11.15}
$$

or, more generally,

$$
\Gamma\left(\frac{n}{2} \right) = \sqrt{\pi} \frac{(n-2)!!}{2^{(n-1)/2}}, \text{ for } n > 0; \text{ and} \tag{11.16}
$$

Euler's reflection formula $\Gamma(x)\Gamma(1-x) = \dfrac{\pi}{\sin(\pi x)}.$ \hfill (11.17)

Here, the *double factorial* is defined by

$$
n!! =
\begin{cases}
1 \text{ for } n = -1, 0, \text{ and} \\
2 \cdot 4 \cdots (n-2) \cdot n \\
\quad = (2k)!! = \displaystyle\prod_{i=1}^{k} (2i) \\
\quad = 2^{n/2} \left(1 \cdot 2 \cdots \dfrac{(n-2)}{2} \cdot \dfrac{n}{2} \right) \\
\quad = k! 2^k \text{ for positive even } n = 2k, k \geq 1, \text{ and} \\
1 \cdot 3 \cdots (n-2) \cdot n \\
\quad = (2k-1)!! = \displaystyle\prod_{i=1}^{k} (2i-1) \\
\quad = \dfrac{1 \cdot 2 \cdots (2k-2) \cdot (2k-1) \cdot (2k)}{(2k)!!} \\
\quad = \dfrac{(2k)!}{k! 2^k} \text{ for odd positive } n = 2k-1, k \geq 1.
\end{cases}
\tag{11.18}
$$

Stirling's formula [again, $O(x)$ means "of the order of x"]

$$
\log n! = n \log n - n + O(\log(n)), \text{ or}
$$
$$
n! \overset{n \to \infty}{\longrightarrow} \sqrt{2\pi n} \left(\frac{n}{e} \right)^n, \text{ or, more generally,}
$$
$$
\Gamma(x) = \sqrt{\frac{2\pi}{x}} \left(\frac{x}{e} \right)^x \left(1 + O\left(\frac{1}{x} \right) \right)
\tag{11.19}
$$

is stated without proof.

11.2   Beta function

The *beta function,* also called the *Euler integral of the first kind,* is a special function defined by

$$
B(x, y) = \int_0^1 t^{x-1} (1-t)^{y-1} dt = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \text{ for } \Re x, \Re y > 0
\tag{11.20}
$$

No proof of the identity of the two representations in terms of an integral, and of Γ-functions is given.

11.3   Fuchsian differential equations

Many differential equations of theoretical physics are Fuchsian equations. We shall therefore study this class in some generality.

11.3.1   Regular, regular singular, and irregular singular point

Consider the homogenuous differential equation [Eq. (9.1) on page 185 is inhomogenuos]

$$
\mathscr{L}_x y(x) = a_2(x) \frac{d^2}{dx^2} y(x) + a_1(x) \frac{d}{dx} y(x) + a_0(x) y(x) = 0.
\tag{11.21}
$$

If $a_0(x)$, $a_1(x)$ and $a_2(x)$ are analytic at some point x_0 and in its neighborhood, and if $a_2(x_0) \neq 0$ at x_0, then x_0 is called an *ordinary point,* or *regular point.* We state without proof that in this case the solutions

around x_0 can be expanded as power series. In this case we can divide equation (11.21) by $a_2(x)$ and rewrite it

$$\frac{1}{a_2(x)} \mathcal{L}_x y(x) = \frac{d^2}{dx^2} y(x) + p_1(x) \frac{d}{dx} y(x) + p_2(x) y(x) = 0, \qquad (11.22)$$

with $p_1(x) = a_1(x)/a_2(x)$ and $p_2(x) = a_0(x)/a_2(x)$.

If, however, $a_2(x_0) = 0$ and $a_1(x_0)$ or $a_0(x_0)$ are nonzero, then the x_0 is called *singular point* of (11.21). The simplest case is if $a_0(x)$ has a *simple zero* at x_0; then both $p_1(x)$ and $p_2(x)$ in (11.22) have at most simple poles.

Furthermore, for reasons disclosed later – mainly motivated by the possibility to write the solutions as power series – a point x_0 is called a *regular singular point* of Eq. (11.21) if

$$\lim_{x \to x_0} (x - x_0) \frac{a_1(x)}{a_2(x)}, \text{ as well as } \lim_{x \to x_0} (x - x_0)^2 \frac{a_0(x)}{a_2(x)} \qquad (11.23)$$

both exist. If any one of these limits does not exist, the singular point is an *irregular singular point*.

A linear ordinary differential equation is called *Fuchsian,* or *Fuchsian differential equation* generalizable to arbitrary order n of differentiation

$$\left[\frac{d^n}{dx^n} + p_1(x) \frac{d^{n-1}}{dx^{n-1}} + \cdots + p_{n-1}(x) \frac{d}{dx} + p_n(x) \right] y(x) = 0, \qquad (11.24)$$

if every singular point, including infinity, is regular, meaning that $p_k(x)$ has at most poles of order k.

A very important case is a Fuchsian of the second order (up to second derivatives occur). In this case, we suppose that the coefficients in (11.22) satisfy the following conditions:

- $p_1(x)$ has at most *single poles*, and

- $p_2(x)$ has at most *double poles*.

The simplest realization of this case is for $a_2(x) = a(x - x_0)^2$, $a_1(x) = b(x - x_0)$, $a_0(x) = c$ for some constant $a, b, c \in \mathbb{C}$.

11.3.2   Behavior at infinity

In order to cope with infinity $z = \infty$ let us transform the Fuchsian equation $w'' + p_1(z) w' + p_2(z) w = 0$ into the new variable $t = \frac{1}{z}$.

$$t = \frac{1}{z}, \ z = \frac{1}{t}, \ u(t) \stackrel{\text{def}}{=} w\left(\frac{1}{t}\right) = w(z)$$

$$\frac{dz}{dt} = -\frac{1}{t^2}, \text{ and thus } \frac{d}{dz} = -t^2 \frac{d}{dt}$$

$$\frac{d^2}{dz^2} = -t^2 \frac{d}{dt} \left(-t^2 \frac{d}{dt} \right) = -t^2 \left(-2t \frac{d}{dt} - t^2 \frac{d^2}{dt^2} \right) = 2t^3 \frac{d}{dt} + t^4 \frac{d^2}{dt^2} \quad (11.25)$$

$$w'(z) = \frac{d}{dz} w(z) = -t^2 \frac{d}{dt} u(t) = -t^2 u'(t)$$

$$w''(z) = \frac{d^2}{dz^2} w(z) = \left(2t^3 \frac{d}{dt} + t^4 \frac{d^2}{dt^2} \right) u(t) = 2t^3 u'(t) + t^4 u''(t)$$

Insertion into the Fuchsian equation $w'' + p_1(z) w' + p_2(z) w = 0$ yields

$$2t^3 u' + t^4 u'' + p_1\left(\frac{1}{t}\right)(-t^2 u') + p_2\left(\frac{1}{t}\right) u = 0, \qquad (11.26)$$

and hence,

$$u'' + \left[\frac{2}{t} - \frac{p_1\left(\frac{1}{t}\right)}{t^2} \right] u' + \frac{p_2\left(\frac{1}{t}\right)}{t^4} u = 0. \tag{11.27}$$

From

$$\tilde{p}_1(t) \overset{\text{def}}{=} \left[\frac{2}{t} - \frac{p_1\left(\frac{1}{t}\right)}{t^2} \right] \tag{11.28}$$

and

$$\tilde{p}_2(t) \overset{\text{def}}{=} \frac{p_2\left(\frac{1}{t}\right)}{t^4} \tag{11.29}$$

follows the form of the rewritten differential equation

$$u'' + \tilde{p}_1(t) u' + \tilde{p}_2(t) u = 0. \tag{11.30}$$

A necessary criterion for this equation to be Fuchsian is that 0 is an ordinary, or at least a regular singular, point.

Note that, for infinity to be a regular singular point, $\tilde{p}_1(t)$ must have at most a pole of the order of t^{-1}, and $\tilde{p}_2(t)$ must have at most a pole of the order of t^{-2} at $t = 0$. Therefore, $(1/t)p_1(1/t) = zp_1(z)$ as well as $(1/t^2)p_2(1/t) = z^2 p_2(z)$ must both be analytic functions as $t \to 0$, or $z \to \infty$. This will be an important finding for the following arguments.

11.3.3 Functional form of the coefficients in Fuchsian differential equations

The functional form of the coefficients $p_1(x)$ and $p_2(x)$, resulting from the assumption of merely regular singular poles, can be estimated as follows.

First, let us start with *poles at finite complex numbers.* Suppose there are k finite poles. [The behavior of $p_1(x)$ and $p_2(x)$ at infinity will be treated later.] Therefore, in Eq. (11.22), the coefficients must be of the form

$$p_1(x) = \frac{P_1(x)}{\prod_{j=1}^{k}(x - x_j)},$$

$$\text{and } p_2(x) = \frac{P_2(x)}{\prod_{j=1}^{k}(x - x_j)^2}, \tag{11.31}$$

where the $x_1, \ldots, x_k$ are k the (regular singular) points of the poles, and $P_1(x)$ and $P_2(x)$ are *entire functions;* that is, they are analytic (or, by another wording, holomorphic) over the whole complex plane formed by $\{x \mid x \in \mathbb{C}\}$.

Second, consider possible *poles at infinity.* Note that the requirement that infinity is regular singular will restrict the possible growth of $p_1(x)$ as well as $p_2(x)$ and thus, to a lesser degree, of $P_1(x)$ as well as $P_2(x)$.

As has been shown earlier, because of the requirement that infinity is regular singular, as x approaches infinity, $p_1(x)x$ as well as $p_2(x)x^2$ must both be analytic. Therefore, $p_1(x)$ cannot grow faster than $|x|^{-1}$, and $p_2(x)$ cannot grow faster than $|x|^{-2}$.

Consequently, by (11.31), as x approaches infinity, $P_1(x) = p_1(x) \prod_{j=1}^{k}(x - x_j)$ does not grow faster than $|x|^{k-1}$ (meaning that it is bounded by some constant times $|x|^{k-1}$), and $P_2(x) = p_2(x) \prod_{j=1}^{k}(x - x_j)^2$

does not grow faster than $|x|^{2k-2}$ (meaning that it is bounded by some constant times $|x|^{k-2}$).

Recall that both $P_1(x)$ and $P_2(x)$ are *entire functions*. Therefore, because of the *generalized Liouville theorem*[1] (mentioned on page 130), both $P_1(x)$ and $P_2(x)$ must be polynomials of degree of at most $k-1$ and $2k-2$, respectively.

Moreover, by using *partial fraction decomposition*[2] of the rational functions (that is, the quotient of polynomials of the form $\frac{R(x)}{Q(x)}$, where $Q(x)$ is not identically zero) in terms of their pole factors $(x - x_j)$, we obtain the general form of the coefficients

$$p_1(x) = \sum_{j=1}^{k} \frac{A_j}{x - x_j},$$

$$\text{and } p_2(x) = \sum_{j=1}^{k} \left[\frac{B_j}{(x - x_j)^2} + \frac{C_j}{x - x_j} \right], \tag{11.32}$$

with constant $A_j, B_j, C_j \in \mathbb{C}$. The resulting Fuchsian differential equation is called *Riemann differential equation.*

Although we have considered an arbitrary finite number of poles, for reasons that are unclear to this author, in physics we are mainly concerned with two poles (i.e., $k = 2$) at finite points, and one at infinity.

The *hypergeometric differential equation* is a *Fuchsian differential equation* which has at most *three regular singularities*, including infinity, at 0, 1, and ∞[3].

11.3.4 Frobenius method by power series

Now let us get more concrete about the solution of Fuchsian equations by *power series*.

In order to obtain a feeling for power series solutions of differential equations, consider the "first order" Fuchsian equation[4]

$$y' - \lambda y = 0. \tag{11.33}$$

Make the *Ansatz*, also known as known as *Frobenius method*[5], that the solution can be expanded into a power series of the form

$$y(x) = \sum_{j=0}^{\infty} a_j x^j. \tag{11.34}$$

Then, Eq. (11.33) can be written as

$$\left(\frac{d}{dx} \sum_{j=0}^{\infty} a_j x^j \right) - \lambda \sum_{j=0}^{\infty} a_j x^j = 0,$$

$$\sum_{j=0}^{\infty} j a_j x^{j-1} - \lambda \sum_{j=0}^{\infty} a_j x^j = 0,$$

$$\sum_{j=1}^{\infty} j a_j x^{j-1} - \lambda \sum_{j=0}^{\infty} a_j x^j = 0, \tag{11.35}$$

$$\sum_{m=j-1=0}^{\infty} (m+1) a_{m+1} x^m - \lambda \sum_{j=0}^{\infty} a_j x^j = 0,$$

$$\sum_{j=0}^{\infty} (j+1) a_{j+1} x^j - \lambda \sum_{j=0}^{\infty} a_j x^j = 0,$$

[1] Robert E. Greene and Stephen G. Krantz. *Function theory of one complex variable*, volume 40 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, third edition, 2006

[2] Gerhard Kristensson. Equations of Fuchsian type. In *Second Order Differential Equations*, pages 29–42. Springer, New York, 2010. ISBN 978-1-4419-7019-0. DOI: 10.1007/978-1-4419-7020-6. URL http://doi.org/10.1007/978-1-4419-7020-6

[3] Vadim Kuznetsov. Special functions and their symmetries. Part I: Algebraic and analytic methods. Postgraduate Course in Applied Analysis, May 2003. URL http://www1.maths.leeds.ac.uk/~kisilv/courses/sp-funct.pdf

[4] Ron Larson and Bruce H. Edwards. *Calculus*. Brooks/Cole Cengage Learning, Belmont, CA, nineth edition, 2010. ISBN 978-0-547-16702-2

[5] George B. Arfken and Hans J. Weber. *Mathematical Methods for Physicists*. Elsevier, Oxford, sixth edition, 2005. ISBN 0-12-059876-0;0-12-088584-0

and hence, by comparing the coefficients of x^j, for $n \geq 0$,

$$(j+1)a_{j+1} = \lambda a_j, \text{ or}$$

$$a_{j+1} = \frac{\lambda a_j}{j+1} = a_0 \frac{\lambda^{j+1}}{(j+1)!}, \text{ and} \qquad (11.36)$$

$$a_j = a_0 \frac{\lambda^j}{j!}.$$

Therefore,

$$y(x) = \sum_{j=0}^{\infty} a_0 \frac{\lambda^j}{j!} x^j = a_0 \sum_{j=0}^{\infty} \frac{(\lambda x)^j}{j!} = a_0 e^{\lambda x}. \qquad (11.37)$$

In the Fuchsian case let us consider the following *Frobenius Ansatz* to expand the solution as a *generalized power series* around a regular singular point x_0, which can be motivated by Eq. (11.31), and by the *Laurent series expansion* (5.28)–(5.30) on page 125:

$$p_1(x) = \frac{A_1(x)}{x - x_0} = \sum_{j=0}^{\infty} \alpha_j (x-x_0)^{j-1} \text{ for } 0 < |x - x_0| < r_1,$$

$$p_2(x) = \frac{A_2(x)}{(x-x_0)^2} = \sum_{j=0}^{\infty} \beta_j (x-x_0)^{j-2} \text{ for } 0 < |x - x_0| < r_2, \qquad (11.38)$$

$$y(x) = (x-x_0)^\sigma \sum_{l=0}^{\infty} (x-x_0)^l w_l = \sum_{l=0}^{\infty} (x-x_0)^{l+\sigma} w_l, \text{ with } w_0 \neq 0,$$

where $A_1(x) = [(x-x_0)a_1(x)]/a_2(x)$ and $A_2(x) = [(x-x_0)^2 a_0(x)]/a_2(x)$. Eq. (11.22) then becomes

$$\frac{d^2}{dx^2} y(x) + p_1(x) \frac{d}{dx} y(x) + p_2(x) y(x) = 0,$$

$$\left[\frac{d^2}{dx^2} + \sum_{j=0}^{\infty} \alpha_j (x-x_0)^{j-1} \frac{d}{dx} + \sum_{j=0}^{\infty} \beta_j (x-x_0)^{j-2} \right] \sum_{l=0}^{\infty} w_l (x-x_0)^{l+\sigma} = 0,$$

$$\sum_{l=0}^{\infty} (l+\sigma)(l+\sigma-1) w_l (x-x_0)^{l+\sigma-2}$$

$$+ \left[\sum_{l=0}^{\infty} (l+\sigma) w_l (x-x_0)^{l+\sigma-1} \right] \sum_{j=0}^{\infty} \alpha_j (x-x_0)^{j-1}$$

$$+ \left[\sum_{l=0}^{\infty} w_l (x-x_0)^{l+\sigma} \right] \sum_{j=0}^{\infty} \beta_j (x-x_0)^{j-2} = 0,$$

$$(x-x_0)^{\sigma-2} \sum_{l=0}^{\infty} (x-x_0)^l [(l+\sigma)(l+\sigma-1) w_l$$

$$+ (l+\sigma) w_l \sum_{j=0}^{\infty} \alpha_j (x-x_0)^j + w_l \sum_{j=0}^{\infty} \beta_j (x-x_0)^j \right] = 0,$$

$$(x-x_0)^{\sigma-2} \left[\sum_{l=0}^{\infty} (l+\sigma)(l+\sigma-1) w_l (x-x_0)^l \right.$$

$$+ \sum_{l=0}^{\infty} (l+\sigma) w_l \sum_{j=0}^{\infty} \alpha_j (x-x_0)^{l+j} + \sum_{l=0}^{\infty} w_l \sum_{j=0}^{\infty} \beta_j (x-x_0)^{l+j} \right] = 0.$$

Next, in order to reach a common power of $(x - x_0)$, we perform an index identification in the second and third summands (where the order of the sums change): $l = m$ in the first summand, as well as an index shift $l + j = m$, and thus $j = m - l$. Since $l \geq 0$ and $j \geq 0$, also $m = l + j$ cannot

be negative. Furthermore, $0 \le j = m - l$, so that $l \le m$.

$$(x - x_0)^{\sigma-2} \left[\sum_{l=0}^{\infty} (l + \sigma)(l + \sigma - 1) w_l (x - x_0)^l \right.$$

$$+ \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} (l + \sigma) w_l \alpha_j (x - x_0)^{l+j}$$

$$\left. + \sum_{j=0}^{\infty} \sum_{l=0}^{\infty} w_l \beta_j (x - x_0)^{l+j} \right] = 0,$$

$$(x - x_0)^{\sigma-2} \left[\sum_{m=0}^{\infty} (m + \sigma)(m + \sigma - 1) w_m (x - x_0)^m \right.$$

$$+ \sum_{m=0}^{\infty} \sum_{l=0}^{m} (l + \sigma) w_l \alpha_{m-l} (x - x_0)^{l+m-l}$$

$$\left. + \sum_{m=0}^{\infty} \sum_{l=0}^{m} w_l \beta_{m-l} (x - x_0)^{l+m-l} \right] = 0,$$ (11.39)

$$(x - x_0)^{\sigma-2} \left\{ \sum_{m=0}^{\infty} (x - x_0)^m \left[(m + \sigma)(m + \sigma - 1) w_m \right. \right.$$

$$\left. \left. + \sum_{l=0}^{m} (l + \sigma) w_l \alpha_{m-l} + \sum_{l=0}^{m} w_l \beta_{m-l} \right] \right\} = 0,$$

$$(x - x_0)^{\sigma-2} \left\{ \sum_{m=0}^{\infty} (x - x_0)^m \left[(m + \sigma)(m + \sigma - 1) w_m \right. \right.$$

$$\left. \left. + \sum_{l=0}^{m} w_l \left((l + \sigma) \alpha_{m-l} + \beta_{m-l} \right) \right] \right\} = 0.$$

If we can divide this equation through $(x - x_0)^{\sigma-2}$ and exploit the linear independence of the polynomials $(x - x_0)^m$, we obtain an infinite number of equations for the infinite number of coefficients w_m by requiring that all the terms "inbetween" the $[\cdots]$–brackets in Eq. (11.39) vanish *individually*. In particular, for $m = 0$ and $w_0 \ne 0$,

$$(0 + \sigma)(0 + \sigma - 1) w_0 + w_0 \left((0 + \sigma) \alpha_0 + \beta_0 \right) = 0$$

$$f_0(\sigma) \overset{\text{def}}{=} \sigma(\sigma - 1) + \sigma \alpha_0 + \beta_0 = 0.$$ (11.40)

The *radius of convergence* of the solution will, in accordance with the Laurent series expansion, extend to the next singularity.

Note that in Eq. (11.40) we have defined $f_0(\sigma)$ which we will use now. Furthermore, for successive m, and with the definition of

$$f_k(\sigma) \overset{\text{def}}{=} \alpha_k \sigma + \beta_k,$$ (11.41)

we obtain the sequence of linear equations

$$w_0 f_0(\sigma) = 0$$

$$w_1 f_0(\sigma + 1) + w_0 f_1(\sigma) = 0,$$

$$w_2 f_0(\sigma + 2) + w_1 f_1(\sigma + 1) + w_0 f_2(\sigma) = 0,$$ (11.42)

$$\vdots$$

$$w_n f_0(\sigma + n) + w_{n-1} f_1(\sigma + n - 1) + \cdots + w_0 f_n(\sigma) = 0.$$

which can be used for an inductive determination of the coefficients w_k.

Eq. (11.40) is a quadratic equation $\sigma^2 + \sigma(\alpha_0 - 1) + \beta_0 = 0$ for the *characteristic exponents*

$$\sigma_{1,2} = \frac{1}{2} \left[1 - \alpha_0 \pm \sqrt{(1 - \alpha_0)^2 - 4\beta_0} \right]$$ (11.43)

We state without proof that, if the difference of the characteristic exponents

$$\sigma_1 - \sigma_2 = \sqrt{(1-\alpha_0)^2 - 4\beta_0} \qquad (11.44)$$

is *nonzero* and *not* an integer, then the two solutions found from $\sigma_{1,2}$ through the generalized series Ansatz (11.38) are linear independent.

Intuitively speaking, the Frobenius method "is in obvious trouble" to find the general solution of the Fuchsian equation if the two characteristic exponents coincide (e.g., $\sigma_1 = \sigma_2$), but it "is also in trouble" to find the general solution if $\sigma_1 - \sigma_2 = m \in \mathbb{N}$; that is, if, for some positive integer m, $\sigma_1 = \sigma_2 + m > \sigma_2$. Because in this case, "eventually" at $n = m$ in Eq. (11.42), we obtain as iterative solution for the coefficient w_m the term

$$
\begin{aligned}
w_m &= -\frac{w_{m-1} f_1(\sigma_2 + m - 1) + \cdots + w_0 f_m(\sigma_2)}{f_0(\sigma_2 + m)} \\
&= -\frac{w_{m-1} f_1(\sigma_1 - 1) + \cdots + w_0 f_m(\sigma_2)}{\underbrace{f_0(\sigma_1)}_{=0}},
\end{aligned}
\qquad (11.45)
$$

as the greater critical exponent σ_1 is a solution of Eq. (11.40) and thus vanishes, leaving us with a vanishing denominator.

11.3.5 d'Alambert reduction of order

If $\sigma_1 = \sigma_2 + n$ with $n \in \mathbb{Z}$, then we find only a *single* solution of the Fuchsian equation. In order to obtain another linear independent solution we have to employ a method based on the Wronskian [6], or the d'Alambert reduction [7], which is a general method to obtain another, linear independent solution $y_2(x)$ from an existing particular solution $y_1(x)$ by the *Ansatz* (no proof is presented here)

$$y_2(x) = y_1(x) \int_x v(s)\,ds. \qquad (11.46)$$

[6] George B. Arfken and Hans J. Weber. *Mathematical Methods for Physicists.* Elsevier, Oxford, sixth edition, 2005. ISBN 0-12-059876-0;0-12-088584-0

[7] Gerald Teschl. *Ordinary Differential Equations and Dynamical Systems. Graduate Studies in Mathematics, volume 140.* American Mathematical Society, Providence, Rhode Island, 2012. ISBN ISBN-10: 0-8218-8328-3 / ISBN-13: 978-0-8218-8328-0. URL http://www.mat.univie.ac.at/~gerald/ftp/book-ode/ode.pdf

Inserting $y_2(x)$ from (11.46) into the Fuchsian equation (11.22), and using the fact that by assumption $y_1(x)$ is a solution of it, yields

$$\frac{d^2}{dx^2}y_2(x) + p_1(x)\frac{d}{dx}y_2(x) + p_2(x)y_2(x) = 0,$$

$$\frac{d^2}{dx^2}y_1(x)\int_x v(s)\,ds + p_1(x)\frac{d}{dx}y_1(x)\int_x v(s)\,ds + p_2(x)y_1(x)\int_x v(s)\,ds = 0,$$

$$\frac{d}{dx}\left\{\left[\frac{d}{dx}y_1(x)\right]\int_x v(s)\,ds + y_1(x)v(x)\right\}$$

$$+p_1(x)\left[\frac{d}{dx}y_1(x)\right]\int_x v(s)\,ds + p_1(x)v(x) + p_2(x)y_1(x)\int_x v(s)\,ds = 0,$$

$$\left[\frac{d^2}{dx^2}y_1(x)\right]\int_x v(s)\,ds + \left[\frac{d}{dx}y_1(x)\right]v(x) + \left[\frac{d}{dx}y_1(x)\right]v(x) + y_1(x)\left[\frac{d}{dx}v(x)\right]$$

$$+p_1(x)\left[\frac{d}{dx}y_1(x)\right]\int_x v(s)\,ds + p_1(x)y_1(x)v(x) + p_2(x)y_1(x)\int_x v(s)\,ds = 0,$$

$$\left[\frac{d^2}{dx^2}y_1(x)\right]\int_x v(s)\,ds + p_1(x)\left[\frac{d}{dx}y_1(x)\right]\int_x v(s)\,ds + p_2(x)y_1(x)\int_x v(s)\,ds$$

$$+p_1(x)y_1(x)v(x) + \left[\frac{d}{dx}y_1(x)\right]v(x) + \left[\frac{d}{dx}y_1(x)\right]v(x) + y_1(x)\left[\frac{d}{dx}v(x)\right] = 0,$$

$$\left[\frac{d^2}{dx^2}y_1(x)\right]\int_x v(s)\,ds + p_1(x)\left[\frac{d}{dx}y_1(x)\right]\int_x v(s)\,ds + p_2(x)y_1(x)\int_x v(s)\,ds$$

$$+y_1(x)\left[\frac{d}{dx}v(x)\right] + 2\left[\frac{d}{dx}y_1(x)\right]v(x) + p_1(x)y_1(x)v(x) = 0,$$

$$\underbrace{\left\{\left[\frac{d^2}{dx^2}y_1(x)\right] + p_1(x)\left[\frac{d}{dx}y_1(x)\right] + p_2(x)y_1(x)\right\}}_{=0}\int_x v(s)\,ds$$

$$+y_1(x)\left[\frac{d}{dx}v(x)\right] + \left\{2\left[\frac{d}{dx}y_1(x)\right] + p_1(x)y_1(x)\right\}v(x) = 0,$$

$$y_1(x)\left[\frac{d}{dx}v(x)\right] + \left\{2\left[\frac{d}{dx}y_1(x)\right] + p_1(x)y_1(x)\right\}v(x) = 0,$$

and finally,

$$v'(x) + v(x)\left\{2\frac{y_1'(x)}{y_1(x)} + p_1(x)\right\} = 0. \tag{11.47}$$

11.3.6   Computation of the characteristic exponent

Let $w'' + p_1(z)w' + p_2(z)w = 0$ be a Fuchsian equation. From the Laurent series expansion of $p_1(z)$ and $p_2(z)$ with Cauchy's integral formula we can derive the following equations, which are helpful in determining the characteristic exponent σ:

$$\alpha_0 = \lim_{z\to z_0}(z - z_0)p_1(z),$$
$$\beta_0 = \lim_{z\to z_0}(z - z_0)^2 p_2(z), \tag{11.48}$$

where z_0 is a regular singular point.

In order to find α_0, consider the Laurent series for

$$p_1(z) = \sum_{k=-1}^{\infty}\tilde{a}_k(z - z_0)^k,$$
$$\text{with } \tilde{a}_k = \frac{1}{2\pi i}\oint p_1(s)(s - z_0)^{-(k+1)}\,ds. \tag{11.49}$$

The summands vanish for $k < -1$, because $p_1(z)$ has at most a pole of order one at z_0.

An index change $n = k + 1$, or $k = n - 1$, as well as a redefinition $\alpha_n \stackrel{\text{def}}{=} \tilde{a}_{n-1}$ yields

$$p_1(z) = \sum_{n=0}^{\infty} \alpha_n (z - z_0)^{n-1}, \qquad (11.50)$$

where

$$\alpha_n = \tilde{a}_{n-1} = \frac{1}{2\pi i} \oint p_1(s)(s - z_0)^{-n} ds; \qquad (11.51)$$

and, in particular,

$$\alpha_0 = \frac{1}{2\pi i} \oint p_1(s) ds. \qquad (11.52)$$

Because the equation is Fuchsian, $p_1(z)$ has at most a pole of order one at z_0. Therefore, $(z - z_0) p_1(z)$ is analytic around z_0. By multiplying $p_1(z)$ with unity $1 = (z - z_0)/(z - z_0)$ and insertion into (11.52) we obtain

$$\alpha_0 = \frac{1}{2\pi i} \oint \frac{p_1(s)(s - z_0)}{(s - z_0)} ds. \qquad (11.53)$$

Cauchy's integral formula (5.20) on page 123 yields

$$\alpha_0 = \lim_{s \to z_0} p_1(s)(s - z_0). \qquad (11.54)$$

Alternatively we may consider the *Frobenius Ansatz* (11.38) $p_1(z) = \sum_{n=0}^{\infty} \alpha_n (z - z_0)^{n-1}$ which has again been motivated by the fact that $p_1(z)$ has at most a pole of order one at z_0. Multiplication of this series by $(z - z_0)$ yields

$$(z - z_0) p_1(z) = \sum_{n=0}^{\infty} \alpha_n (z - z_0)^n. \qquad (11.55)$$

In the limit $z \to z_0$,

$$\alpha_0 = \lim_{z \to z_0} (z - z_0) p_1(z). \qquad (11.56)$$

Likewise, let us find the expression for β_0 by considering the Laurent series for

$$p_2(z) = \sum_{k=-2}^{\infty} \tilde{b}_k (z - z_0)^k$$
$$\text{with } \tilde{b}_k = \frac{1}{2\pi i} \oint p_2(s)(s - z_0)^{-(k+1)} ds. \qquad (11.57)$$

The summands vanish for $k < -2$, because $p_2(z)$ has at most a pole of order two at z_0.

An index change $n = k + 2$, or $k = n - 2$, as well as a redefinition $\beta_n \stackrel{\text{def}}{=} \tilde{b}_{n-2}$ yields

$$p_2(z) = \sum_{n=0}^{\infty} \beta_n (z - z_0)^{n-2}, \qquad (11.58)$$

where

$$\beta_n = \tilde{b}_{n-2} = \frac{1}{2\pi i} \oint p_2(s)(s - z_0)^{-(n-1)} ds; \qquad (11.59)$$

and, in particular,

$$\beta_0 = \frac{1}{2\pi i} \oint (s - z_0) p_2(s) ds. \qquad (11.60)$$

Because the equation is Fuchsian, $p_2(z)$ has at most a pole of order two at z_0. Therefore, $(z - z_0)^2 p_2(z)$ is analytic around z_0. By multiplying $p_2(z)$ with unity $1 = (z - z_0)^2/(z - z_0)^2$ and insertion into (11.60) we obtain

$$\beta_0 = \frac{1}{2\pi i} \oint \frac{p_2(s)(s - z_0)^2}{(s - z_0)} ds. \qquad (11.61)$$

Cauchy's integral formula (5.20) on page 123 yields

$$\beta_0 = \lim_{s \to z_0} p_2(s)(s - z_0)^2. \tag{11.62}$$

Again another way to see this is with the *Frobenius Ansatz* (11.38) $p_2(z) = \sum_{n=0}^{\infty} \beta_n (z - z_0)^{n-2}$. Multiplication with $(z - z_0)^2$, and taking the limit $z \to z_0$, yields

$$\lim_{z \to z_0} (z - z_0)^2 p_2(z) = \beta_n. \tag{11.63}$$

11.3.7   Examples

Let us consider some examples involving Fuchsian equations of the second order.

1. First we shall prove that $z^2 y''(z) + z y'(z) - y(z) = 0$ is of the Fuchsian type; and compute the solutions with the Frobenius method.

   Let us first locate the singularities of

   $$y''(z) + \frac{y'(z)}{z} - \frac{y(z)}{z^2} = 0. \tag{11.64}$$

   One singularity is at the finite point finite $z_0 = 0$.

   In order to analyze the singularity at infinity we have to transform the equation by $z = 1/t$. First observe that $p_1(z) = 1/z$ and $p_2(z) = -1/z^2$. Therefore, after the transformation, the new coefficients, computed from (11.28) and (11.29), are

   $$\tilde{p}_1(t) = \left(\frac{2}{t} - \frac{t}{t^2} \right) = \frac{1}{t}, \text{ and}$$
   $$\tilde{p}_2(t) = \frac{-t^2}{t^4} = -\frac{1}{t^2}. \tag{11.65}$$

   Thereby we effectively regain the original type of equation (11.64). We can thus treat both singularities at zero and infinity in the same way.

   Both singularities are regular, as the coefficients p_1 and $\tilde{p}_1$ have poles of order 1, p_2 and $\tilde{p}_2$ have poles of order 2, respectively. Therefore, the differential equation is Fuchsian.

   In order to obtain solutions, let us first compute the characteristic exponents by

   $$\alpha_0 = \lim_{z \to 0} z p_1(z) = \lim_{z \to 0} z \frac{1}{z} = 1, \text{ and}$$
   $$\beta_0 = \lim_{z \to 0} z^2 p_2(z) = \lim_{z \to 0} -z^2 \frac{1}{z^2} = -1, \tag{11.66}$$

   so that, from (11.40),

   $$0 = f_0(\sigma) = \sigma(\sigma - 1) + \sigma \alpha_0 + \beta_0 = \sigma^2 - \sigma + \sigma - 1 =$$
   $$= \sigma^2 - 1, \text{ and thus } \sigma_{1,2} = \pm 1. \tag{11.67}$$

   The first solution is obtained by insertion of the *Frobenius Ansatz* (11.38), in particular, $y_1(x) = \sum_{l=0}^{\infty} (x - x_0)^{l+\sigma} w_l$ with $\sigma_1 = 1$

into (11.64). In this case,

$$z^2 \sum_{l=0}^{\infty} (l+1)l(x-x_0)^{l-1} w_l + z \sum_{l=0}^{\infty} l(x-x_0)^l w_l - \sum_{l=0}^{\infty} (x-x_0)^{l+1} w_l = 0,$$

$$\sum_{l=0}^{\infty} \left[w_l(l+1)l + w_l(l+1) - w_l \right] (x-x_0)^{l+1} w_l = 0,$$

$$\sum_{l=0}^{\infty} w_l l(l+2)(x-x_0)^{l+1} = 0.$$

(11.68)

Since the polynomials are linear independent, we obtain $w_l l(l+2) = 0$ for all $l \geq 0$. Therefore, for constant A,

$$w_0 = A, \text{ and } w_l = 0 \text{ for } l > 0. \tag{11.69}$$

So, the first solution is $y_1(z) = Az$.

The second solution, computed through the *Frobenius Ansatz* (11.38), is obtained by inserting $y_2(x) = \sum_{l=0}^{\infty} (x-x_0)^{l+\sigma} w_l$ with $\sigma_2 = -1$ into (11.64). This yields

$$z^2 \sum_{l=0}^{\infty} (l-1)(l-2)(x-x_0)^{l-3} w_l +$$

$$+ z \sum_{l=0}^{\infty} (l-1)(x-x_0)^{l-2} w_l - \sum_{l=0}^{\infty} (x-x_0)^{l-1} w_l = 0,$$

$$\sum_{l=0}^{\infty} \left[w_l(l-1)(l-2) + w_l(l-1) - w_l \right] (x-x_0)^{l-1} w_l = 0,$$

(11.70)

$$\sum_{l=0}^{\infty} w_l l(l-2)(x-x_0)^{l-1} = 0.$$

Since the polynomials are linear independent, we obtain $w_l l(l-2) = 0$ for all $l \geq 0$. Therefore, for constant B, C,

$$w_0 = B, w_1 = 0, w_2 = C, \text{ and } w_l = 0 \text{ for } l > 2, \tag{11.71}$$

So that the second solution is $y_2(z) = B\frac{1}{z} + Cz$.

Note that $y_2(z)$ already represents the general solution of (11.64). Alternatively we could have started from $y_1(z)$ and applied d'Alambert's *Ansatz* (11.46)–(11.47):

> Most of the coefficients are zero, so no iteration with a "catastrophic divisions by zero" occurs here.

$$v'(z) + v(z)\left(\frac{2}{z} + \frac{1}{z}\right) = 0, \text{ or } v'(z) = -\frac{3v(z)}{z} \tag{11.72}$$

yields

$$\frac{dv}{v} = -3\frac{dz}{z}, \text{ and } \log v = -3\log z, \text{ or } v(z) = z^{-3}. \tag{11.73}$$

Therefore, according to (11.46),

$$y_2(z) = y_1(z) \int_z v(s) ds = Az\left(-\frac{1}{2z^2}\right) = \frac{A'}{z}. \tag{11.74}$$

2. Find out whether the following differential equations are Fuchsian, and enumerate the regular singular points:

$$z w'' + (1-z) w' = 0,$$
$$z^2 w'' + z w' - v^2 w = 0,$$
$$z^2(1+z)^2 w'' + 2z(z+1)(z+2) w' - 4w = 0,$$
$$2z(z+2) w'' + w' - z w = 0.$$

(11.75)

ad 1: $z w'' + (1-z) w' = 0 \implies w'' + \dfrac{(1-z)}{z} w' = 0$

$z = 0$:
$$\alpha_0 = \lim_{z \to 0} z \frac{(1-z)}{z} = 1, \quad \beta_0 = \lim_{z \to 0} z^2 \cdot 0 = 0.$$

The equation for the characteristic exponent is

$$\sigma(\sigma - 1) + \sigma \alpha_0 + \beta_0 = 0 \implies \sigma^2 - \sigma + \sigma = 0 \implies \sigma_{1,2} = 0.$$

$z = \infty$: $z = \frac{1}{t}$

$$\tilde{p}_1(t) = \frac{2}{t} - \frac{\left(1 - \frac{1}{t}\right)}{\frac{1}{t}} = \frac{2}{t} - \frac{\left(1 - \frac{1}{t}\right)}{t} = \frac{1}{t} + \frac{1}{t^2} = \frac{t+1}{t^2}$$

$\implies$ not Fuchsian.

ad 2: $z^2 w'' + z w' - \nu^2 w = 0 \implies w'' + \dfrac{1}{z} w' - \dfrac{\nu^2}{z^2} w = 0.$

$z = 0$:
$$\alpha_0 = \lim_{z \to 0} z \frac{1}{z} = 1, \quad \beta_0 = \lim_{z \to 0} z^2 \left(-\frac{\nu^2}{z^2}\right) = -\nu^2.$$
$$\implies \sigma^2 - \sigma + \sigma - \nu^2 = 0 \implies \sigma_{1,2} = \pm \nu$$

$z = \infty$: $z = \frac{1}{t}$

$$\tilde{p}_1(t) \;=\; \frac{2}{t} - \frac{1}{t^2} t = \frac{1}{t}$$
$$\tilde{p}_2(t) \;=\; \frac{1}{t^4}\left(-t^2 \nu^2\right) = -\frac{\nu^2}{t^2}$$

$$\implies u'' + \frac{1}{t} u' - \frac{\nu^2}{t^2} u = 0 \implies \sigma_{1,2} = \pm \nu$$

$\implies$ Fuchsian equation.

ad 3:

$$z^2(1+z)^2 w'' + 2z(z+1)(z+2) w' - 4w = 0 \implies w'' + \frac{2(z+2)}{z(z+1)} w' - \frac{4}{z^2(1+z)^2} w = 0$$

$z = 0$:
$$\alpha_0 = \lim_{z \to 0} z \frac{2(z+2)}{z(z+1)} = 4, \quad \beta_0 = \lim_{z \to 0} z^2 \left(-\frac{4}{z^2(1+z)^2}\right) = -4.$$

$$\implies \sigma(\sigma - 1) + 4\sigma - 4 = \sigma^2 + 3\sigma - 4 = 0 \implies \sigma_{1,2} = \frac{-3 \pm \sqrt{9 + 16}}{2} = \begin{cases} -4 \\ +1 \end{cases}$$

$z = -1$:
$$\alpha_0 = \lim_{z \to -1} (z+1) \frac{2(z+2)}{z(z+1)} = -2, \quad \beta_0 = \lim_{z \to -1} (z+1)^2 \left(-\frac{4}{z^2(1+z)^2}\right) = -4.$$

$$\implies \sigma(\sigma - 1) - 2\sigma - 4 = \sigma^2 - 3\sigma - 4 = 0 \implies \sigma_{1,2} = \frac{3 \pm \sqrt{9 + 16}}{2} = \begin{cases} +4 \\ -1 \end{cases}$$

$z = \infty$:

$$\tilde{p}_1(t) = \frac{2}{t} - \frac{1}{t^2}\frac{2\left(\frac{1}{t}+2\right)}{\frac{1}{t}\left(\frac{1}{t}+1\right)} = \frac{2}{t} - \frac{2\left(\frac{1}{t}+2\right)}{1+t} = \frac{2}{t}\left(1 - \frac{1+2t}{1+t}\right)$$

$$\tilde{p}_2(t) = \frac{1}{t^4}\left(-\frac{4}{\frac{1}{t^2}\left(1+\frac{1}{t}\right)^2}\right) = -\frac{4}{t^2}\frac{t^2}{(t+1)^2} = -\frac{4}{(t+1)^2}$$

$$\Longrightarrow u'' + \frac{2}{t}\left(1 - \frac{1+2t}{1+t}\right)u' - \frac{4}{(t+1)^2}u = 0$$

$$\alpha_0 = \lim_{t\to 0} t\frac{2}{t}\left(1 - \frac{1+2t}{1+t}\right) = 0, \quad \beta_0 = \lim_{t\to 0} t^2\left(-\frac{4}{(t+1)^2}\right) = 0.$$

$$\Longrightarrow \sigma(\sigma-1) = 0 \Longrightarrow \sigma_{1,2} = \begin{cases} 0 \\ 1 \end{cases}$$

$\Longrightarrow$ Fuchsian equation.

ad 4:

$$2z(z+2)w'' + w' - zw = 0 \Longrightarrow w'' + \frac{1}{2z(z+2)}w' - \frac{1}{2(z+2)}w = 0$$

$z = 0$:

$$\alpha_0 = \lim_{z\to 0} z\frac{1}{2z(z+2)} = \frac{1}{4}, \quad \beta_0 = \lim_{z\to 0} z^2\frac{-1}{2(z+2)} = 0.$$

$$\Longrightarrow \sigma^2 - \sigma + \frac{1}{4}\sigma = 0 \Longrightarrow \sigma^2 - \frac{3}{4}\sigma = 0 \Longrightarrow \sigma_1 = 0, \sigma_2 = \frac{3}{4}.$$

$z = -2$:

$$\alpha_0 = \lim_{z\to -2} (z+2)\frac{1}{2z(z+2)} = -\frac{1}{4}, \quad \beta_0 = \lim_{z\to -2} (z+2)^2\frac{-1}{2(z+2)} = 0.$$

$$\Longrightarrow \sigma_1 = 0, \quad \sigma_2 = \frac{5}{4}.$$

$z = \infty$:

$$\tilde{p}_1(t) = \frac{2}{t} - \frac{1}{t^2}\left(\frac{1}{2\frac{1}{t}\left(\frac{1}{t}+2\right)}\right) = \frac{2}{t} - \frac{1}{2(1+2t)}$$

$$\tilde{p}_2(t) = \frac{1}{t^4}\frac{(-1)}{2\left(\frac{1}{t}+2\right)} = -\frac{1}{2t^3(1+2t)}$$

$\Longrightarrow$ not a Fuchsian.

3. Determine the solutions of

$$z^2 w'' + (3z+1)w' + w = 0$$

around the regular singular points.

The singularities are at $z = 0$ and $z = \infty$.

Singularities at $z = 0$:

$$p_1(z) = \frac{3z+1}{z^2} = \frac{a_1(z)}{z} \text{ with } a_1(z) = 3 + \frac{1}{z}$$

$p_1(z)$ has a pole of higher order than one; hence this is no Fuchsian equation; and $z = 0$ is an irregular singular point.

Singularities at $z = \infty$:

- Transformation $z = \dfrac{1}{t}$, $w(z) \to u(t)$:

$$u''(t) + \left[\frac{2}{t} - \frac{1}{t^2} p_1\left(\frac{1}{t}\right)\right] \cdot u'(t) + \frac{1}{t^4} p_2\left(\frac{1}{t}\right) \cdot u(t) = 0.$$

The new coefficient functions are

$$\tilde{p}_1(t) = \frac{2}{t} - \frac{1}{t^2} p_1\left(\frac{1}{t}\right) = \frac{2}{t} - \frac{1}{t^2}(3t + t^2) = \frac{2}{t} - \frac{3}{t} - 1 = -\frac{1}{t} - 1$$

$$\tilde{p}_2(t) = \frac{1}{t^4} p_2\left(\frac{1}{t}\right) = \frac{t^2}{t^4} = \frac{1}{t^2}$$

- check whether this is a regular singular point:

$$\tilde{p}_1(t) = -\frac{1+t}{t} = \frac{\tilde{a}_1(t)}{t} \qquad \text{with} \quad \tilde{a}_1(t) = -(1+t) \qquad \text{regular}$$
$$\tilde{p}_2(t) = \frac{1}{t^2} = \frac{\tilde{a}_2(t)}{t^2} \qquad \text{with} \quad \tilde{a}_2(t) = 1 \qquad\qquad \text{regular}$$

$\tilde{a}_1$ and $\tilde{a}_2$ are regular at $t = 0$, hence this is a regular singular point.

- *Ansatz* around $t = 0$: the transformed equation is

$$u''(t) + \tilde{p}_1(t)u'(t) + \tilde{p}_2(t)u(t) = 0$$
$$u''(t) - \left(\frac{1}{t} + 1\right)u'(t) + \frac{1}{t^2}u(t) = 0$$
$$t^2 u''(t) - (t + t^2)u'(t) + u(t) = 0$$

The generalized power series is

$$u(t) = \sum_{n=0}^{\infty} w_n t^{n+\sigma}$$
$$u'(t) = \sum_{n=0}^{\infty} w_n (n+\sigma) t^{n+\sigma-1}$$
$$u''(t) = \sum_{n=0}^{\infty} w_n (n+\sigma)(n+\sigma-1) t^{n+\sigma-2}$$

If we insert this into the transformed differential equation we obtain

$$t^2 \sum_{n=0}^{\infty} w_n (n+\sigma)(n+\sigma-1) t^{n+\sigma-2} -$$

$$- (t+t^2) \sum_{n=0}^{\infty} w_n (n+\sigma) t^{n+\sigma-1} + \sum_{n=0}^{\infty} w_n t^{n+\sigma} = 0$$

$$\sum_{n=0}^{\infty} w_n (n+\sigma)(n+\sigma-1) t^{n+\sigma} - \sum_{n=0}^{\infty} w_n (n+\sigma) t^{n+\sigma} -$$

$$- \sum_{n=0}^{\infty} w_n (n+\sigma) t^{n+\sigma+1} + \sum_{n=0}^{\infty} w_n t^{n+\sigma} = 0$$

Change of index: $m = n + 1$, $n = m - 1$ in the third sum yields

$$\sum_{n=0}^{\infty} w_n \left[(n+\sigma)(n+\sigma-2) + 1\right] t^{n+\sigma} - \sum_{m=1}^{\infty} w_{m-1}(m-1+\sigma) t^{m+\sigma} = 0.$$

In the second sum, substitute m for n

$$\sum_{n=0}^{\infty} w_n \left[(n+\sigma)(n+\sigma-2) + 1\right] t^{n+\sigma} - \sum_{n=1}^{\infty} w_{n-1}(n+\sigma-1) t^{n+\sigma} = 0.$$

We write out explicitly the $n = 0$ term of the first sum

$$w_0\Big[\sigma(\sigma-2)+1\Big]t^\sigma + \sum_{n=1}^{\infty} w_n\Big[(n+\sigma)(n+\sigma-2)+1\Big]t^{n+\sigma}-$$

$$-\sum_{n=1}^{\infty} w_{n-1}(n+\sigma-1)t^{n+\sigma} = 0.$$

Now we can combine the two sums

$$w_0\Big[\sigma(\sigma-2)+1\Big]t^\sigma +$$

$$+\sum_{n=1}^{\infty}\Big\{w_n\Big[(n+\sigma)(n+\sigma-2)+1\Big]-w_{n-1}(n+\sigma-1)\Big\}t^{n+\sigma} = 0.$$

The left hand side can only vanish for all t if the coefficients vanish; hence

$$w_0\Big[\sigma(\sigma-2)+1\Big] \quad = \quad 0, \quad (11.76)$$

$$w_n\Big[(n+\sigma)(n+\sigma-2)+1\Big]-w_{n-1}(n+\sigma-1) \quad = \quad 0. \quad (11.77)$$

ad (11.76) for w_0:

$$\sigma(\sigma-2)+1 \quad = \quad 0$$

$$\sigma^2 - 2\sigma + 1 \quad = \quad 0$$

$$(\sigma-1)^2 \quad = \quad 0 \quad \Longrightarrow \sigma_\infty^{(1,2)} = 1$$

The charakteristic exponent is $\sigma_\infty^{(1)} = \sigma_\infty^{(2)} = 1$.

ad (11.77) for w_n: For the coefficients w_n we obtain the recursion formula

$$w_n\Big[(n+\sigma)(n+\sigma-2)+1\Big] \quad = \quad w_{n-1}(n+\sigma-1)$$

$$\Longrightarrow w_n \quad = \quad \frac{n+\sigma-1}{(n+\sigma)(n+\sigma-2)+1}w_{n-1}.$$

Let us insert $\sigma = 1$:

$$w_n = \frac{n}{(n+1)(n-1)+1}w_{n-1} = \frac{n}{n^2-1+1}w_{n-1} = \frac{n}{n^2}w_{n-1} = \frac{1}{n}w_{n-1}.$$

We can fix $w_0 = 1$, hence:

$$w_0 \quad = \quad 1 = \frac{1}{1} = \frac{1}{0!}$$

$$w_1 \quad = \quad \frac{1}{1} = \frac{1}{1!}$$

$$w_2 \quad = \quad \frac{1}{1\cdot 2} = \frac{1}{2!}$$

$$w_3 \quad = \quad \frac{1}{1\cdot 2\cdot 3} = \frac{1}{3!}$$

$$\vdots$$

$$w_n \quad = \quad \frac{1}{1\cdot 2\cdot 3\cdot\cdots\cdot n} = \frac{1}{n!}$$

And finally,

$$u_1(t) = t^\sigma \sum_{n=0}^{\infty} w_n t^n = t\sum_{n=0}^{\infty}\frac{t^n}{n!} = te^t.$$

- Notice that both characteristic exponents are equal; hence we have to employ the d'Alambert reduction

$$u_2(t) = u_1(t) \int_0^t v(s) \, ds$$

with

$$v'(t) + v(t) \left[2 \frac{u_1'(t)}{u_1(t)} + \tilde{p}_1(t) \right] = 0.$$

Insertion of u_1 and $\tilde{p}_1$,

$$
\begin{aligned}
u_1(t) &= t e^t \\
u_1'(t) &= e^t (1 + t) \\
\tilde{p}_1(t) &= -\left(\frac{1}{t} + 1 \right),
\end{aligned}
$$

yields

$$
\begin{aligned}
v'(t) + v(t) \left(2 \frac{e^t(1+t)}{t e^t} - \frac{1}{t} - 1 \right) &= 0 \\
v'(t) + v(t) \left(2 \frac{(1+t)}{t} - \frac{1}{t} - 1 \right) &= 0 \\
v'(t) + v(t) \left(\frac{2}{t} + 2 - \frac{1}{t} - 1 \right) &= 0 \\
v'(t) + v(t) \left(\frac{1}{t} + 1 \right) &= 0 \\
\frac{dv}{dt} &= -v \left(1 + \frac{1}{t} \right) \\
\frac{dv}{v} &= -\left(1 + \frac{1}{t} \right) dt
\end{aligned}
$$

Upon integration of both sides we obtain

$$
\begin{aligned}
\int \frac{dv}{v} &= -\int \left(1 + \frac{1}{t} \right) dt \\
\log v &= -(t + \log t) = -t - \log t \\
v &= \exp(-t - \log t) = e^{-t} e^{-\log t} = \frac{e^{-t}}{t},
\end{aligned}
$$

and hence an explicit form of $v(t)$:

$$v(t) = \frac{1}{t} e^{-t}.$$

If we insert this into the equation for u_2 we obtain

$$u_2(t) = t e^t \int_0^t \frac{1}{s} e^{-s} \, ds.$$

- Therefore, with $t = \frac{1}{z}$, $u(t) = w(z)$, the two linear independent solutions around the regular singular point at $z = \infty$ are

$$w_1(z) = \frac{1}{z} \exp\left(\frac{1}{z} \right), \text{ and}$$

$$w_2(z) = \frac{1}{z} \exp\left(\frac{1}{z} \right) \int_0^{\frac{1}{z}} \frac{1}{t} e^{-t} \, dt. \tag{11.78}$$

11.4   Hypergeometric function

11.4.1   Definition

A *hypergeometric series* is a series

$$\sum_{j=0}^{\infty} c_j, \tag{11.79}$$

where the quotients $\frac{c_{j+1}}{c_j}$ are *rational functions* (that is, the quotient of two polynomials $\frac{R(x)}{Q(x)}$, where $Q(x)$ is not identically zero) of j, so that they can be factorized by

$$
\begin{aligned}
\frac{c_{j+1}}{c_j} &= \frac{(j+a_1)(j+a_2)\cdots(j+a_p)}{(j+b_1)(j+b_2)\cdots(j+b_q)}\left(\frac{x}{j+1}\right), \\
\text{or } c_{j+1} &= c_j \frac{(j+a_1)(j+a_2)\cdots(j+a_p)}{(j+b_1)(j+b_2)\cdots(j+b_q)}\left(\frac{x}{j+1}\right) \\
&= c_{j-1}\frac{(j-1+a_1)(j-1+a_2)\cdots(j-1+a_p)}{(j-1+b_1)(j-1+b_2)\cdots(j-1+b_q)}\times \\
&\quad \times \frac{(j+a_1)(j+a_2)\cdots(j+a_p)}{(j+b_1)(j+b_2)\cdots(j+b_q)}\left(\frac{x}{j}\right)\left(\frac{x}{j+1}\right) \\
&= c_0 \frac{a_1 a_2 \cdots a_p}{b_1 b_2 \cdots b_q}\cdots\frac{(j-1+a_1)(j-1+a_2)\cdots(j-1+a_p)}{(j-1+b_1)(j-1+b_2)\cdots(j-1+b_q)}\times \\
&\quad \times \frac{(j+a_1)(j+a_2)\cdots(j+a_p)}{(j+b_1)(j+b_2)\cdots(j+b_q)}\left(\frac{x}{1}\right)\cdots\left(\frac{x}{j}\right)\left(\frac{x}{j+1}\right) \\
&= c_0 \frac{(a_1)_{j+1}(a_2)_{j+1}\cdots(a_p)_{j+1}}{(b_1)_{j+1}(b_2)_{j+1}\cdots(b_q)_{j+1}}\left(\frac{x^{j+1}}{(j+1)!}\right).
\end{aligned}
\tag{11.80}
$$

The factor $j+1$ in the denominator of the first line of (11.80) on the right yields $(j+1)!$. If it not there "naturally" we may obtain it by compensating it with a factor $j+1$ in the numerator.

With this iterated ratio (11.80), the hypergeometric series (11.79) can be written in terms of *shifted factorials*, or, by another naming, the *Pochhammer symbol*, as

$$
\begin{aligned}
\sum_{j=0}^{\infty} c_j &= c_0 \sum_{j=0}^{\infty} \frac{(a_1)_j(a_2)_j\cdots(a_p)_j}{(b_1)_j(b_2)_j\cdots(b_q)_j}\frac{x^j}{j!} \\
&= c_0\, {}_pF_q\left(\begin{array}{c} a_1,\ldots,a_p \\ b_1,\ldots,b_q \end{array};x\right), \text{ or} \\
&= c_0\, {}_pF_q\left(a_1,\ldots,a_p;b_1,\ldots,b_q;x\right).
\end{aligned}
\tag{11.81}
$$

Apart from this definition *via* hypergeometric series, the Gauss *hypergeometric function*, or, used synonymuously, the *Gauss series*

$$
\begin{aligned}
{}_2F_1\left(\begin{array}{c} a,b \\ c \end{array};x\right) &= {}_2F_1\left(a,b;c;x\right) = \sum_{j=0}^{\infty}\frac{(a)_j(b)_j}{(c)_j}\frac{x^j}{j!} \\
&= 1 + \frac{ab}{c}x + \frac{1}{2!}\frac{a(a+1)b(b+1)}{c(c+1)}x^2 + \cdots
\end{aligned}
\tag{11.82}
$$

can be defined as a solution of a *Fuchsian differential equation* which has at most *three regular singularities* at 0, 1, and ∞.

Indeed, any Fuchsian equation with finite regular singularities at x_1 and x_2 can be rewritten into the *Riemann differential equation* (11.32),

which in turn can be rewritten into the *Gaussian differential equation* or *hypergeometric differential equation* with regular singularities at 0, 1, and ∞ [8]. This can be demonstrated by rewriting any such equation of the form

$$
w''(x) + \left(\frac{A_1}{x - x_1} + \frac{A_2}{x - x_2} \right) w'(x)
$$
$$
+ \left(\frac{B_1}{(x - x_1)^2} + \frac{B_2}{(x - x_2)^2} + \frac{C_1}{x - x_1} + \frac{C_2}{x - x_2} \right) w(x) = 0
$$
(11.83)

through transforming Eq. (11.83) into the *hypergeometric equation*

$$
\left[\frac{d^2}{dx^2} + \frac{(a + b + 1)x - c}{x(x - 1)} \frac{d}{dx} + \frac{ab}{x(x - 1)} \right] {}_2F_1(a, b; c; x) = 0,
$$
(11.84)

where the solution is proportional to the Gauss hypergeometric function

$$
w(x) \longrightarrow (x - x_1)^{\sigma_1^{(1)}} (x - x_2)^{\sigma_2^{(2)}} {}_2F_1(a, b; c; x),
$$
(11.85)

and the variable transform as

$$
x \longrightarrow x = \frac{x - x_1}{x_2 - x_1}, \text{ with}
$$
$$
a = \sigma_1^{(1)} + \sigma_2^{(1)} + \sigma_\infty^{(1)},
$$
$$
b = \sigma_1^{(1)} + \sigma_2^{(1)} + \sigma_\infty^{(2)},
$$
$$
c = 1 + \sigma_1^{(1)} - \sigma_1^{(2)}.
$$
(11.86)

where $\sigma_j^{(i)}$ stands for the ith characteristic exponent of the jth singularity.

Whereas the full transformation from Eq. (11.83) to the hypergeometric equation (11.84) will not been given, we shall show that the Gauss hypergeometric function ${}_2F_1$ satisfies the hypergeometric equation (11.84).

First, define the differential operator

$$
\vartheta = x \frac{d}{dx},
$$
(11.87)

and observe that

$$
\vartheta(\vartheta + c - 1)x^n = x \frac{d}{dx} \left(x \frac{d}{dx} + c - 1 \right) x^n
$$
$$
= x \frac{d}{dx} \left(xnx^{n-1} + cx^n - x^n \right)
$$
$$
= x \frac{d}{dx} \left(nx^n + cx^n - x^n \right)
$$
$$
= x \frac{d}{dx} (n + c - 1) x^n
$$
$$
= n(n + c - 1) x^n.
$$
(11.88)

The Bessel equation has a regular singular point at 0, and an irregular singular point at infinity.

Thus, if we apply $\vartheta(\vartheta + c - 1)$ to $_2F_1$, then

$$\vartheta(\vartheta + c - 1)\,_2F_1(a, b; c; x) = \vartheta(\vartheta + c - 1) \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \frac{x^j}{j!}$$

$$= \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \frac{j(j + c - 1)x^j}{j!} = \sum_{j=1}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \frac{j(j + c - 1)x^j}{j!}$$

$$= \sum_{j=1}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \frac{(j + c - 1)x^j}{(j - 1)!}$$

[index shift: $j \to n + 1, n = j - 1, n \geq 0$]

$$= \sum_{n=0}^{\infty} \frac{(a)_{n+1} (b)_{n+1}}{(c)_{n+1}} \frac{(n + 1 + c - 1)x^{n+1}}{n!}$$

$$= x \sum_{n=0}^{\infty} \frac{(a)_n (a + n)(b)_n (b + n)}{(c)_n (c + n)} \frac{(n + c)x^n}{n!}$$

$$= x \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{(a + n)(b + n)x^n}{n!}$$

$$= x(\vartheta + a)(\vartheta + b) \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{x^n}{n!} = x(\vartheta + a)(\vartheta + b)\,_2F_1(a, b; c; x),$$

(11.89)

where we have used

$$(a + n)x^n = (a + \vartheta)x^n, \text{ and}$$

$$(a)_{n+1} = a(a + 1) \cdots (a + n - 1)(a + n) = (a)_n (a + n).$$

(11.90)

Writing out ϑ in Eq. (11.89) explicitly yields

$$\{\vartheta(\vartheta + c - 1) - x(\vartheta + a)(\vartheta + b)\}\,_2F_1(a, b; c; x) = 0,$$

$$\left\{ x\frac{d}{dx}\left(x\frac{d}{dx} + c - 1\right) - x\left(x\frac{d}{dx} + a\right)\left(x\frac{d}{dx} + b\right)\right\} _2F_1(a, b; c; x) = 0,$$

$$\left\{ \frac{d}{dx}\left(x\frac{d}{dx} + c - 1\right) - \left(x\frac{d}{dx} + a\right)\left(x\frac{d}{dx} + b\right)\right\} _2F_1(a, b; c; x) = 0,$$

$$\left\{ \frac{d}{dx} + x\frac{d^2}{dx^2} + (c - 1)\frac{d}{dx} - \left(x^2\frac{d^2}{dx^2} + x\frac{d}{dx} + bx\frac{d}{dx}\right.\right.$$
$$\left.\left. + ax\frac{d}{dx} + ab\right)\right\} _2F_1(a, b; c; x) = 0,$$

$$\left\{ (x - x^2)\frac{d^2}{dx^2} + (1 + c - 1 - x - x(a + b))\frac{d}{dx} + ab\right\} _2F_1(a, b; c; x) = 0,$$

$$\left\{ -x(x - 1)\frac{d^2}{dx^2} - (c - x(1 + a + b))\frac{d}{dx} - ab\right\} _2F_1(a, b; c; x) = 0,$$

$$\left\{ \frac{d^2}{dx^2} + \frac{x(1 + a + b) - c}{x(x - 1)}\frac{d}{dx} + \frac{ab}{x(x - 1)}\right\} _2F_1(a, b; c; x) = 0.$$

(11.91)

11.4.2 Properties

There exist many properties of the hypergeometric series. In the following we shall mention a few.

$$\frac{d}{dz}\,_2F_1(a, b; c; z) = \frac{ab}{c}\,_2F_1(a + 1, b + 1; c + 1; z).$$

(11.92)

$$\frac{d}{dz} \, _2F_1(a,b;c;z) \quad = \quad \frac{d}{dz} \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^n}{n!} =$$

$$= \quad \sum_{n=0}^{\infty} \frac{(a)_n (b)_n}{(c)_n} n \frac{z^{n-1}}{n!}$$

$$= \quad \sum_{n=1}^{\infty} \frac{(a)_n (b)_n}{(c)_n} \frac{z^{n-1}}{(n-1)!}$$

An index shift $n \to m+1$, $m = n-1$, and a subsequent renaming $m \to n$, yields

$$\frac{d}{dz} \, _2F_1(a,b;c;z) = \sum_{n=0}^{\infty} \frac{(a)_{n+1}(b)_{n+1}}{(c)_{n+1}} \frac{z^n}{n!}.$$

As

$$(x)_{n+1} \quad = \quad x(x+1)(x+2) \cdots (x+n-1)(x+n)$$

$$(x+1)_n \quad = \quad (x+1)(x+2) \cdots (x+n-1)(x+n)$$

$$(x)_{n+1} \quad = \quad x(x+1)_n$$

holds, we obtain

$$\frac{d}{dz} \, _2F_1(a,b;c;z) = \sum_{n=0}^{\infty} \frac{ab}{c} \frac{(a+1)_n (b+1)_n}{(c+1)_n} \frac{z^n}{n!} = \frac{ab}{c} \, _2F_1(a+1,b+1;c+1;z).$$

We state *Euler's integral representation* for $\Re c > 0$ and $\Re b > 0$ without proof:

$$_2F_1(a,b;c;x) = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)} \int_0^1 t^{b-1}(1-t)^{c-b-1}(1-xt)^{-a} dt. \quad (11.93)$$

For $\Re(c-a-b) > 0$, we also state Gauss' theorem

$$_2F_1(a,b;c;1) = \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{j!(c)_j} = \frac{\Gamma(c)\Gamma(c-a-b)}{\Gamma(c-a)\Gamma(c-b)}. \quad (11.94)$$

For a proof, we can set $x = 1$ in Euler's integral representation, and the Beta function defined in Eq. (11.20).

11.4.3   Plasticity

Some of the most important elementary functions can be expressed as hypergeometric series; most importantly the Gaussian one $_2F_1$, which is

sometimes denoted by just F. Let us enumerate a few.

$$e^x = {}_0F_0(-;-;x) \tag{11.95}$$

$$\cos x = {}_0F_1(-;\frac{1}{2};-\frac{x^2}{4}) \tag{11.96}$$

$$\sin x = x\,{}_0F_1(-;\frac{3}{2};-\frac{x^2}{4}) \tag{11.97}$$

$$(1-x)^{-a} = {}_1F_0(a;-;x) \tag{11.98}$$

$$\sin^{-1}x = x\,{}_2F_1(\frac{1}{2},\frac{1}{2};\frac{3}{2};x^2) \tag{11.99}$$

$$\tan^{-1}x = x\,{}_2F_1(\frac{1}{2},1;\frac{3}{2};-x^2) \tag{11.100}$$

$$\log(1+x) = x\,{}_2F_1(1,1;2;-x) \tag{11.101}$$

$$H_{2n}(x) = \frac{(-1)^n(2n)!}{n!}\,{}_1F_1(-n;\frac{1}{2};x^2) \tag{11.102}$$

$$H_{2n+1}(x) = 2x\frac{(-1)^n(2n+1)!}{n!}\,{}_1F_1(-n;\frac{3}{2};x^2) \tag{11.103}$$

$$L_n^\alpha(x) = \binom{n+\alpha}{n}\,{}_1F_1(-n;\alpha+1;x) \tag{11.104}$$

$$P_n(x) = P_n^{(0,0)}(x) = {}_2F_1(-n,n+1;1;\frac{1-x}{2}), \tag{11.105}$$

$$C_n^\gamma(x) = \frac{(2\gamma)_n}{(\gamma+\frac{1}{2})_n}P_n^{(\gamma-\frac{1}{2},\gamma-\frac{1}{2})}(x), \tag{11.106}$$

$$T_n(x) = \frac{n!}{(\frac{1}{2})_n}P_n^{(-\frac{1}{2},-\frac{1}{2})}(x), \tag{11.107}$$

$$J_\alpha(x) = \frac{(\frac{x}{2})^\alpha}{\Gamma(\alpha+1)}\,{}_0F_1(-;\alpha+1;-\frac{1}{4}x^2), \tag{11.108}$$

where H stands for *Hermite polynomials*, L for *Laguerre polynomials*,

$$P_n^{(\alpha,\beta)}(x) = \frac{(\alpha+1)_n}{n!}\,{}_2F_1(-n,n+\alpha+\beta+1;\alpha+1;\frac{1-x}{2}) \tag{11.109}$$

for *Jacobi polynomials*, C for *Gegenbauer polynomials*, T for *Chebyshev polynomials*, P for *Legendre polynomials*, and J for the *Bessel functions of the first kind*, respectively.

1. Let us prove that
$$\log(1-z) = -z\,{}_2F_1(1,1;2;z).$$

Consider

$${}_2F_1(1,1,2;z) = \sum_{m=0}^{\infty}\frac{[(1)_m]^2}{(2)_m}\frac{z^m}{m!} = \sum_{m=0}^{\infty}\frac{[1\cdot2\cdot\cdots\cdot m]^2}{2\cdot(2+1)\cdot\cdots\cdot(2+m-1)}\frac{z^m}{m!}$$

With

$$(1)_m = 1\cdot2\cdot\cdots\cdot m = m!, \qquad (2)_m = 2\cdot(2+1)\cdot\cdots\cdot(2+m-1) = (m+1)!$$

follows

$${}_2F_1(1,1,2;z) = \sum_{m=0}^{\infty}\frac{[m!]^2}{(m+1)!}\frac{z^m}{m!} = \sum_{m=0}^{\infty}\frac{z^m}{m+1}.$$

Index shift $k = m+1$

$${}_2F_1(1,1,2;z) = \sum_{k=1}^{\infty}\frac{z^{k-1}}{k}$$

and hence

$$-z\,{}_2F_1(1,1,2;z) = -\sum_{k=1}^{\infty}\frac{z^k}{k}.$$

Compare with the series

$$\log(1+x) = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k} \qquad \text{for} \quad -1 < x \le 1$$

If one substitutes $-x$ for x, then

$$\log(1-x) = -\sum_{k=1}^{\infty}\frac{x^k}{k}.$$

The identity follows from the analytic continuation of x to the complex z plane.

2. Let us prove that, because of $(a+z)^n = \sum_{k=0}^{n}\binom{n}{k}z^k a^{n-k}$,

$$(1-z)^n = {}_2F_1(-n,1,1;z).$$

$$_2F_1(-n,1,1;z) = \sum_{i=0}^{\infty}\frac{(-n)_i(1)_i}{(1)_i}\frac{z^i}{i!} = \sum_{i=0}^{\infty}(-n)_i\frac{z^i}{i!}.$$

Consider $(-n)_i$

$$(-n)_i = (-n)(-n+1)\cdots(-n+i-1).$$

For even $n \ge 0$ the series stops after a finite number of terms, because the factor $-n+i-1 = 0$ for $i = n+1$ vanishes; hence the sum of i extends only from 0 to n. Hence, if we collect the factors (-1) which yield $(-1)^i$ we obtain

$$(-n)_i = (-1)^i\,n(n-1)\cdots[n-(i-1)] = (-1)^i\frac{n!}{(n-i)!}.$$

Hence, insertion into the Gauss hypergeometric function yields

$$_2F_1(-n,1,1;z) = \sum_{i=0}^{n}(-1)^i z^i\frac{n!}{i!(n-i)!} = \sum_{i=0}^{n}\binom{n}{i}(-z)^i.$$

This is the binomial series

$$(1+x)^n = \sum_{k=0}^{n}\binom{n}{k}x^k$$

with $x = -z$; and hence,

$$_2F_1(-n,1,1;z) = (1-z)^n.$$

3. Let us prove that, because of $\arcsin x = \sum_{k=0}^{\infty}\frac{(2k)!\,x^{2k+1}}{2^{2k}(k!)^2(2k+1)}$,

$$_2F_1\left(\frac{1}{2},\frac{1}{2},\frac{3}{2};\sin^2 z\right) = \frac{z}{\sin z}.$$

Consider

$$_2F_1\left(\frac{1}{2},\frac{1}{2},\frac{3}{2};\sin^2 z\right) = \sum_{m=0}^{\infty}\frac{\left[\left(\frac{1}{2}\right)_m\right]^2}{\left(\frac{3}{2}\right)_m}\frac{(\sin z)^{2m}}{m!}.$$

We take

$$
\begin{aligned}
(2n)!! &= 2\cdot 4\cdot\cdots\cdot(2n) = n!2^n \\
(2n-1)!! &= 1\cdot 3\cdot\cdots\cdot(2n-1) = \frac{(2n)!}{2^n n!}
\end{aligned}
$$

Hence

$$
\begin{aligned}
\left(\frac{1}{2}\right)_m &= \frac{1}{2}\cdot\left(\frac{1}{2}+1\right)\cdots\left(\frac{1}{2}+m-1\right) = \frac{1\cdot 3\cdot 5\cdots(2m-1)}{2^m} = \frac{(2m-1)!!}{2^m} \\
\left(\frac{3}{2}\right)_m &= \frac{3}{2}\cdot\left(\frac{3}{2}+1\right)\cdots\left(\frac{3}{2}+m-1\right) = \frac{3\cdot 5\cdot 7\cdots(2m+1)}{2^m} = \frac{(2m+1)!!}{2^m}
\end{aligned}
$$

Therefore,

$$
\frac{\left(\frac{1}{2}\right)_m}{\left(\frac{3}{2}\right)_m} = \frac{1}{2m+1}.
$$

On the other hand,

$$
\begin{aligned}
(2m)! &= 1\cdot 2\cdot 3\cdot\cdots\cdot(2m-1)(2m) = (2m-1)!!(2m)!! = \\
&= 1\cdot 3\cdot 5\cdot\cdots\cdot(2m-1)\cdot 2\cdot 4\cdot 6\cdot\cdots\cdot(2m) = \\
&= \left(\frac{1}{2}\right)_m 2^m\cdot 2^m m! = 2^{2m} m!\left(\frac{1}{2}\right)_m \Longrightarrow \left(\frac{1}{2}\right)_m = \frac{(2m)!}{2^{2m}m!}
\end{aligned}
$$

Upon insertion one obtains

$$
F\left(\frac{1}{2},\frac{1}{2},\frac{3}{2};\sin^2 z\right) = \sum_{m=0}^{\infty} \frac{(2m)!(\sin z)^{2m}}{2^{2m}(m!)^2(2m+1)}.
$$

Comparing with the series for arcsin one finally obtains

$$
\sin z\, F\left(\frac{1}{2},\frac{1}{2},\frac{3}{2};\sin^2 z\right) = \arcsin(\sin z) = z.
$$

11.4.4 Four forms

We state without proof the four forms of the Gauss hypergeometric function [9].

$$
\begin{aligned}
{}_2F_1(a,b;c;x) &= (1-x)^{c-a-b}\,{}_2F_1(c-a,c-b;c;x) &\text{(11.110)} \\
&= (1-x)^{-a}\,{}_2F_1\left(a,c-b;c;\frac{x}{x-1}\right) &\text{(11.111)} \\
&= (1-x)^{-b}\,{}_2F_1\left(b,c-a;c;\frac{x}{x-1}\right). &\text{(11.112)}
\end{aligned}
$$

[9] T. M. MacRobert. *Spherical Harmonics. An Elementary Treatise on Harmonic Functions with Applications*, volume 98 of *International Series of Monographs in Pure and Applied Mathematics*. Pergamon Press, Oxford, third edition, 1967

11.5 Orthogonal polynomials

Many systems or sequences of functions may serve as a *basis of linearly independent functions* which are capable to "cover" – that is, to approximate – certain functional classes [10]. We have already encountered at least two such prospective bases [cf. Eq. (6.12)]:

$$
\{1, x, x^2, \ldots, x^k, \ldots\} \text{ with } f(x) = \sum_{k=0}^{\infty} c_k x^k, \tag{11.113}
$$

and

$$
\left\{e^{ikx} \mid k\in\mathbb{Z}\right\} \quad \text{for } f(x+2\pi) = f(x)
$$

$$
\text{with } f(x) = \sum_{k=-\infty}^{\infty} c_k e^{ikx}, \tag{11.114}
$$

$$
\text{where } c_k = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-ikx}dx.
$$

[10] Russell Herman. *A Second Course in Ordinary Differential Equations: Dynamical Systems and Boundary Value Problems*. University of North Carolina Wilmington, Wilmington, NC, 2008. URL http://people.uncw.edu/hermanr/mat463/ODEBook/Book/ODE_LargeFont.pdf. Creative Commons Attribution-NoncommercialShare Alike 3.0 United States License; and Francisco Marcellán and Walter Van Assche. *Orthogonal Polynomials and Special Functions*, volume 1883 of *Lecture Notes in Mathematics*. Springer, Berlin, 2006. ISBN 3-540-31062-2

In order to claim existence of such functional basis systems, let us first define what *orthogonality* means in the functional context. Just as for linear vector spaces, we can define an *inner product* or *scalar product* [cf. also Eq. (6.4)] of two real-valued functions $f(x)$ and $g(x)$ by the integral [11]

$$\langle f \mid g \rangle = \int_a^b f(x)g(x)\rho(x)dx \qquad (11.115)$$

for some suitable *weight function* $\rho(x) \geq 0$. Very often, the weight function is set to the identity; that is, $\rho(x) = \rho = 1$. We notice without proof that $\langle f|g \rangle$ satisfies all requirements of a scalar product. A system of functions $\{\psi_0, \psi_1, \psi_2, \ldots, \psi_k, \ldots\}$ is orthogonal if, for $j \neq k$,

$$\langle \psi_j \mid \psi_k \rangle = \int_a^b \psi_j(x)\psi_k(x)\rho(x)dx = 0. \qquad (11.116)$$

Suppose, in some generality, that $\{f_0, f_1, f_2, \ldots, f_k, \ldots\}$ is a sequence of nonorthogonal functions. Then we can apply a *Gram-Schmidt orthogonalization process* to these functions and thereby obtain orthogonal functions $\{\phi_0, \phi_1, \phi_2, \ldots, \phi_k, \ldots\}$ by

$$\phi_0(x) = f_0(x),$$
$$\phi_k(x) = f_k(x) - \sum_{j=0}^{k-1} \frac{\langle f_k \mid \phi_j \rangle}{\langle \phi_j \mid \phi_j \rangle} \phi_j(x). \qquad (11.117)$$

Note that the proof of the Gram-Schmidt process in the functional context is analoguous to the one in the vector context.

11.6   *Legendre polynomials*

The polynomial functions in $\{1, x, x^2, \ldots, x^k, \ldots\}$ are not mutually orthogonal because, for instance, with $\rho = 1$ and $b = -a = 1$,

$$\langle 1 \mid x^2 \rangle = \int_{a=-1}^{b=1} x^2 dx = \left.\frac{x^3}{3}\right|_{x=-1}^{x=1} = \frac{2}{3}. \qquad (11.118)$$

Hence, by the Gram-Schmidt process we obtain

$$\phi_0(x) = 1,$$
$$\phi_1(x) = x - \frac{\langle x \mid 1 \rangle}{\langle 1 \mid 1 \rangle} 1$$
$$= x - 0 = x,$$
$$\phi_2(x) = x^2 - \frac{\langle x^2 \mid 1 \rangle}{\langle 1 \mid 1 \rangle} 1 - \frac{\langle x^2 \mid x \rangle}{\langle x \mid x \rangle} x \qquad (11.119)$$
$$= x^2 - \frac{2/3}{2} 1 - 0x = x^2 - \frac{1}{3},$$
$$\vdots$$

If, on top of orthogonality, we are "forcing" a type of "normalization" by defining

$$P_l(x) \stackrel{\text{def}}{=} \frac{\phi_l(x)}{\phi_l(1)}, \qquad (11.120)$$
$$\text{with } P_l(1) = 1,$$

[11] Herbert S. Wilf. *Mathematics for the physical sciences.* Dover, New York, 1962. URL http://www.math.upenn.edu/~wilf/website/Mathematics_for_the_Physical_Sciences.html

then the resulting orthogonal polynomials are the *Legendre polynomials* P_l; in particular,

$$P_0(x) = 1,$$
$$P_1(x) = x,$$
$$P_2(x) = \left(x^2 - \frac{1}{3}\right) \Big/ \frac{2}{3} = \frac{1}{2}\left(3x^2 - 1\right), \qquad (11.121)$$

$$\vdots$$

with $P_l(1) = 1$, $l = \mathbb{N}_0$.

Why should we be interested in orthonormal systems of functions? Because, as pointed out earlier in the contect of hypergeometric functions, they could be alternatively defined as the eigenfunctions and solutions of certain differential equation, such as, for instance, the Schrödinger equation, which may be subjected to a separation of variables. For Legendre polynomials the associated differential equation is the *Legendre equation*

$$\left\{(1 - x^2)\frac{d^2}{dx^2} - 2x\frac{d}{dx} + l(l+1)\right\} P_l(x) = 0,$$
$$\text{or } \left\{\frac{d}{dx}\left((1 - x^2)\frac{d}{dx}\right) + l(l+1)\right\} P_l(x) = 0 \qquad (11.122)$$

for $l \in \mathbb{N}_0$, whose Sturm-Liouville form has been mentioned earlier in Table 9.1 on page 191. For a proof, we refer to the literature.

11.6.1 Rodrigues formula

A third alternative definition of Legendre polynomials is by the Rodrigues formula

$$P_l(x) = \frac{1}{2^l l!} \frac{d^l}{dx^l}(x^2 - 1)^l, \text{ for } l \in \mathbb{N}_0. \qquad (11.123)$$

Again, no proof of equivalence will be given.

For even l, $P_l(x) = P_l(-x)$ is an even function of x, whereas for odd l, $P_l(x) = -P_l(-x)$ is an odd function of x; that is,

$$P_l(-x) = (-1)^l P_l(x). \qquad (11.124)$$

Moreover,

$$P_l(-1) = (-1)^l \qquad (11.125)$$

and

$$P_{2k+1}(0) = 0. \qquad (11.126)$$

This can be shown by the substitution $t = -x$, $dt = -dx$, and insertion into the Rodrigues formula:

$$P_l(-x) = \frac{1}{2^l l!} \frac{d^l}{du^l}(u^2 - 1)^l \Big|_{u=-x} = [u \to -u] =$$

$$= \frac{1}{(-1)^l} \frac{1}{2^l l!} \frac{d^l}{du^l}(u^2 - 1)^l \Big|_{u=x} = (-1)^l P_l(x).$$

Because of the "normalization" $P_l(1) = 1$ we obtain

$$P_l(-1) = (-1)^l P_l(1) = (-1)^l.$$

And as $P_l(-0) = P_l(0) = (-1)^l P_l(0)$, we obtain $P_l(0) = 0$ for odd l.

11.6.2   Generating function

For $|x| < 1$ and $|t| < 1$ the Legendre polynomials $P_l(x)$ are the coefficients in the Taylor series expansion of the following generating function

$$g(x, t) = \frac{1}{\sqrt{1 - 2xt + t^2}} = \sum_{l=0}^{\infty} P_l(x)\, t^l \qquad (11.127)$$

around $t = 0$. No proof is given here.

11.6.3   The three term and other recursion formulae

Among other things, generating functions are useful for the derivation of certain recursion relations involving Legendre polynomials.

For instance, for $l = 1, 2, \ldots$, the three term recursion formula

$$(2l + 1) x P_l(x) = (l + 1) P_{l+1}(x) + l P_{l-1}(x), \qquad (11.128)$$

or, by substituting $l - 1$ for l, for $l = 2, 3 \ldots$,

$$(2l - 1) x P_{l-1}(x) = l P_l(x) + (l - 1) P_{l-2}(x), \qquad (11.129)$$

can be proven as follows.

$$g(x, t) = \frac{1}{\sqrt{1 - 2tx + t^2}} = \sum_{n=0}^{\infty} t^n P_n(x)$$

$$\frac{\partial}{\partial t} g(x, t) = -\frac{1}{2}(1 - 2tx + t^2)^{-\frac{3}{2}}(-2x + 2t) = \frac{1}{\sqrt{1 - 2tx + t^2}} \frac{x - t}{1 - 2tx + t^2}$$

$$\frac{\partial}{\partial t} g(x, t) = \frac{x - t}{1 - 2tx + t^2} \sum_{n=0}^{\infty} t^n P_n(x) = \sum_{n=0}^{\infty} n t^{n-1} P_n(x)$$

$$(x - t) \sum_{n=0}^{\infty} t^n P_n(x) - (1 - 2tx + t^2) \sum_{n=0}^{\infty} n t^{n-1} P_n(x) = 0$$

$$\sum_{n=0}^{\infty} x t^n P_n(x) - \sum_{n=0}^{\infty} t^{n+1} P_n(x) - \sum_{n=1}^{\infty} n t^{n-1} P_n(x) +$$

$$+ \sum_{n=0}^{\infty} 2xn t^n P_n(x) - \sum_{n=0}^{\infty} n t^{n+1} P_n(x) = 0$$

$$\sum_{n=0}^{\infty} (2n + 1) x t^n P_n(x) - \sum_{n=0}^{\infty} (n + 1) t^{n+1} P_n(x) - \sum_{n=1}^{\infty} n t^{n-1} P_n(x) = 0$$

$$\sum_{n=0}^{\infty} (2n + 1) x t^n P_n(x) - \sum_{n=1}^{\infty} n t^n P_{n-1}(x) - \sum_{n=0}^{\infty} (n + 1) t^n P_{n+1}(x) = 0,$$

$$x P_0(x) - P_1(x) + \sum_{n=1}^{\infty} t^n \left[(2n + 1) x P_n(x) - n P_{n-1}(x) - (n + 1) P_{n+1}(x) \right] = 0,$$

hence

$$x P_0(x) - P_1(x) = 0, \qquad (2n + 1) x P_n(x) - n P_{n-1}(x) - (n + 1) P_{n+1}(x) = 0,$$

hence

$$P_1(x) = x P_0(x), \qquad (n + 1) P_{n+1}(x) = (2n + 1) x P_n(x) - n P_{n-1}(x).$$

Let us prove

$$P_{l-1}(x) = P_l'(x) - 2x P_{l-1}'(x) + P_{l-2}'(x). \qquad (11.130)$$

$$g(x,t) = \frac{1}{\sqrt{1-2tx+t^2}} = \sum_{n=0}^{\infty} t^n P_n(x)$$

$$\frac{\partial}{\partial x} g(x,t) = -\frac{1}{2}(1-2tx+t^2)^{-\frac{3}{2}}(-2t) = \frac{1}{\sqrt{1-2tx+t^2}}\frac{t}{1-2tx+t^2}$$

$$\frac{\partial}{\partial x} g(x,t) = \frac{t}{1-2tx+t^2}\sum_{n=0}^{\infty} t^n P_n(x) = \sum_{n=0}^{\infty} t^n P'_n(x)$$

$$\sum_{n=0}^{\infty} t^{n+1} P_n(x) = \sum_{n=0}^{\infty} t^n P'_n(x) - \sum_{n=0}^{\infty} 2xt^{n+1} P'_n(x) + \sum_{n=0}^{\infty} t^{n+2} P'_n(x)$$

$$\sum_{n=1}^{\infty} t^n P_{n-1}(x) = \sum_{n=0}^{\infty} t^n P'_n(x) - \sum_{n=1}^{\infty} 2xt^n P'_{n-1}(x) + \sum_{n=2}^{\infty} t^n P'_{n-2}(x)$$

$$tP_0 + \sum_{n=2}^{\infty} t^n P_{n-1}(x) = P'_0(x) + tP'_1(x) + \sum_{n=2}^{\infty} t^n P'_n(x) -$$

$$- 2xtP'_0 - \sum_{n=2}^{\infty} 2xt^n P'_{n-1}(x) + \sum_{n=2}^{\infty} t^n P'_{n-2}(x)$$

$$P'_0(x) + t\left[P'_1(x) - P_0(x) - 2xP'_0(x)\right] +$$

$$+ \sum_{n=2}^{\infty} t^n [P'_n(x) - 2xP'_{n-1}(x) + P'_{n-2}(x) - P_{n-1}(x)] = 0$$

$$P'_0(x) = 0, \text{ hence } P_0(x) = \text{const.}$$

$$P'_1(x) - P_0(x) - 2xP'_0(x) = 0.$$

Because of $P'_0(x) = 0$ we obtain $P'_1(x) - P_0(x) = 0$, hence $P'_1(x) = P_0(x)$, and

$$P'_n(x) - 2xP'_{n-1}(x) + P'_{n-2}(x) - P_{n-1}(x) = 0.$$

Finally we substitute $n+1$ for n:

$$P'_{n+1}(x) - 2xP'_n(x) + P'_{n-1}(x) - P_n(x) = 0,$$

hence

$$P_n(x) = P'_{n+1}(x) - 2xP'_n(x) + P'_{n-1}(x).$$

Let us prove

$$P'_{l+1}(x) - P'_{l-1}(x) = (2l+1)P_l(x). \tag{11.131}$$

$$(n+1)P_{n+1}(x) \quad = \quad (2n+1)xP_n(x) - nP_{n-1}(x) \quad \left|\frac{d}{dx}\right.$$

$$(n+1)P'_{n+1}(x) \quad = \quad (2n+1)P_n(x) + (2n+1)xP'_n(x) - nP'_{n-1}(x) \quad \left|\cdot 2\right.$$

(i): $\quad (2n+2)P'_{n+1}(x) \quad = \quad 2(2n+1)P_n(x) + 2(2n+1)xP'_n(x) - 2nP'_{n-1}(x)$

$$P'_{n+1}(x) - 2xP'_n(x) + P'_{n-1}(x) = P_n(x) \quad \left|\cdot(2n+1)\right.$$

(ii): $\quad (2n+1)P'_{n+1}(x) - 2(2n+1)xP'_n(x) + (2n+1)P'_{n-1}(x) = (2n+1)P_n(x)$

We subtract (ii) from (i):

$$P'_{n+1}(x) + 2(2n+1)xP'_n(x) - (2n+1)P'_{n-1}(x) =$$

$$= (2n+1)P_n(x) + 2(2n+1)xP'_n(x) - 2nP'_{n-1}(x);$$

hence

$$P'_{n+1}(x) - P'_{n-1}(x) = (2n+1)P_n(x).$$

11.6.4   Expansion in Legendre polynomials

We state without proof that square integrable functions $f(x)$ can be written as series of Legendre polynomials as

$$f(x) = \sum_{l=0}^{\infty} a_l P_l(x),$$

with expansion coefficients $a_l = \dfrac{2l+1}{2} \displaystyle\int_{-1}^{+1} f(x)P_l(x)dx.$

(11.132)

Let us expand the Heaviside function defined in Eq. (7.110)

$$H(x) = \begin{cases} 1 & \text{for } x \ge 0 \\ 0 & \text{for } x < 0 \end{cases}$$

(11.133)

in terms of Legendre polynomials.

We shall use the recursion formula $(2l+1)P_l = P'_{l+1} - P'_{l-1}$ and rewrite

$$
\begin{aligned}
a_l &= \frac{1}{2}\int_0^1 \left(P'_{l+1}(x) - P'_{l-1}(x)\right)dx = \frac{1}{2}\left(P_{l+1}(x) - P_{l-1}(x)\right)\Big|_{x=0}^1 = \\
&= \frac{1}{2}\underbrace{\left[P_{l+1}(1) - P_{l-1}(1)\right]}_{\substack{=0 \text{ because of}\\ \text{``normalization''}}} - \frac{1}{2}\left[P_{l+1}(0) - P_{l-1}(0)\right].
\end{aligned}
$$

Note that $P_n(0) = 0$ for *odd* n; hence $a_l = 0$ for *even* $l \ne 0$. We shall treat the case $l = 0$ with $P_0(x) = 1$ separately. Upon substituting $2l+1$ for l one obtains

$$a_{2l+1} = -\frac{1}{2}\left[P_{2l+2}(0) - P_{2l}(0)\right].$$

We shall next use the formula

$$P_l(0) = (-1)^{\frac{l}{2}} \frac{l!}{2^l\left(\left(\frac{l}{2}\right)!\right)^2},$$

and for *even* $l \ge 0$ one obtains

$$
\begin{aligned}
a_{2l+1} &= -\frac{1}{2}\left[\frac{(-1)^{l+1}(2l+2)!}{2^{2l+2}((l+1)!)^2} - \frac{(-1)^l(2l)!}{2^{2l}(l!)^2}\right] = \\
&= (-1)^l \frac{(2l)!}{2^{2l+1}(l!)^2}\left[\frac{(2l+1)(2l+2)}{2^2(l+1)^2} + 1\right] = \\
&= (-1)^l \frac{(2l)!}{2^{2l+1}(l!)^2}\left[\frac{2(2l+1)(l+1)}{2^2(l+1)^2} + 1\right] = \\
&= (-1)^l \frac{(2l)!}{2^{2l+1}(l!)^2}\left[\frac{2l+1+2l+2}{2(l+1)}\right] = \\
&= (-1)^l \frac{(2l)!}{2^{2l+1}(l!)^2}\left[\frac{4l+3}{2(l+1)}\right] = \\
&= (-1)^l \frac{(2l)!(4l+3)}{2^{2l+2}l!(l+1)!}
\end{aligned}
$$

$$a_0 = \frac{1}{2}\int_{-1}^{+1} H(x)\underbrace{P_0(x)}_{=1}\,dx = \frac{1}{2}\int_0^1 dx = \frac{1}{2};$$

and finally

$$H(x) = \frac{1}{2} + \sum_{l=0}^{\infty}(-1)^l \frac{(2l)!(4l+3)}{2^{2l+2}l!(l+1)!}P_{2l+1}(x).$$

11.7   Associated Legendre polynomial

Associated Legendre polynomials $P_l^m(x)$ are the solutions of the *general Legendre equation*

$$\left\{(1-x^2)\frac{d^2}{dx^2} - 2x\frac{d}{dx} + \left[l(l+1) - \frac{m^2}{1-x^2}\right]\right\} P_l^m(x) = 0,$$
$$\text{or } \left[\frac{d}{dx}\left((1-x^2)\frac{d}{dx}\right) + l(l+1) - \frac{m^2}{1-x^2}\right] P_l^m(x) = 0 \tag{11.134}$$

Eq. (11.134) reduces to the Legendre equation (11.122) on page 224 for $m = 0$; hence

$$P_l^0(x) = P_l(x). \tag{11.135}$$

More generally, by differentiating m times the Legendre equation (11.122) it can be shown that

$$P_l^m(x) = (-1)^m (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} P_l(x). \tag{11.136}$$

By inserting $P_l(x)$ from the *Rodrigues formula* for Legendre polynomials (11.123) we obtain

$$P_l^m(x) = (-1)^m (1-x^2)^{\frac{m}{2}} \frac{d^m}{dx^m} \frac{1}{2^l l!} \frac{d^l}{dx^l} (x^2-1)^l$$
$$= \frac{(-1)^m (1-x^2)^{\frac{m}{2}}}{2^l l!} \frac{d^{m+l}}{dx^{m+l}} (x^2-1)^l. \tag{11.137}$$

In terms of the Gauss hypergeometric function the associated Legendre polynomials can be generalied to arbitrary complex indices μ, λ and argument x by

$$P_\lambda^\mu(x) = \frac{1}{\Gamma(1-\mu)}\left(\frac{1+x}{1-x}\right)^{\frac{\mu}{2}} {}_2F_1\left(-\lambda, \lambda+1; 1-\mu; \frac{1-x}{2}\right). \tag{11.138}$$

No proof is given here.

11.8   Spherical harmonics

Let us define the *spherical harmonics* $Y_l^m(\theta, \varphi)$ by

$$Y_l^m(\theta, \varphi) = \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} P_l^m(\cos\theta) e^{im\varphi} \text{ for } -l \le m \le l.. \tag{11.139}$$

Spherical harmonics are solutions of the differential equation

$$\{\Delta + l(l+1)\} Y_l^m(\theta, \varphi) = 0. \tag{11.140}$$

This equation is what typically remains after separation and "removal" of the radial part of the Laplace equation $\Delta\psi(r, \theta, \varphi) = 0$ in three dimensions when the problem is invariant (symmetric) under rotations.

Twice continuously differentiable, complex-valued solutions u of the Laplace equation $\Delta u = 0$ are called *harmonic functions*

Sheldon Axler, Paul Bourdon, and Wade Ramey. *Harmonic Function Theory*, volume 137 of *Graduate texts in mathematics*. second edition, 1994. ISBN 0-387-97875-5

11.9   Solution of the Schrödinger equation for a hydrogen atom

Suppose Schrödinger, in his 1926 *annus mirabilis* – a year which seems to have been initiated by a trip to Arosa with 'an old girlfriend from Vienna'

(apparently it was neither his wife Anny who remained in Zurich, nor Lotte, nor Irene nor Felicie [12]), – came down from the mountains or from whatever realm he was in – and handed you over some partial differential equation for the hydrogen atom – an equation (note that the quantum mechanical "momentum operator" $\mathscr{P}$ is identified with $-i\text{—}h\nabla$)

$$\frac{1}{2\mu}\mathscr{P}^2\psi = \frac{1}{2\mu}\left(\mathscr{P}_x^2 + \mathscr{P}_y^2 + \mathscr{P}_z^2\right)\psi = (E - V)\,\psi,$$

$$\text{or, with } V = -\frac{e^2}{4\pi\epsilon_0 r},$$

$$-\left[\frac{-h^2}{2\mu}\Delta + \frac{e^2}{4\pi\epsilon_0 r}\right]\psi(\mathbf{x}) = E\psi, \qquad (11.141)$$

$$\text{or } \left[\Delta + \frac{2\mu}{-h^2}\left(\frac{e^2}{4\pi\epsilon_0 r} + E\right)\right]\psi(\mathbf{x}) = 0,$$

which would later bear his name – and asked you if you could be so kind to please solve it for him. Actually, by Schrödinger's own account [13] he handed over this eigenwert equation to Hermann Klaus Hugo Weyl; in this instance he was not dissimilar from Einstein, who seemed to have employed a (human) computator on a very regular basis. Schrödinger might also have hinted that μ, e, and ϵ_0 stand for some (reduced) mass, charge, and the permittivity of the vacuum, respectively, $\text{—}h$ is a constant of (the dimension of) action, and E is some eigenvalue which must be determined from the solution of (11.141).

So, what could you do? First, observe that the problem is spherical symmetric, as the potential just depends on the radius $r = \sqrt{\mathbf{x}\cdot\mathbf{x}}$, and also the Laplace operator $\Delta = \nabla\cdot\nabla$ allows spherical symmetry. Thus we could write the Schrödinger equation (11.141) in terms of spherical coordinates (r,θ,φ) with $x = r\sin\theta\cos\varphi$, $y = r\sin\theta\sin\varphi$, $z = r\cos\theta$, whereby θ is the polar angle in the x–z-plane measured from the z-axis, with $0 \le \theta \le \pi$, and φ is the azimuthal angle in the x–y-plane, measured from the x-axis with $0 \le \varphi < 2\pi$. In terms of spherical coordinates the Laplace operator essentially "decays into" (i.e. consists additively of) a radial part and an angular part

$$\Delta = \left(\frac{\partial}{\partial x}\right)^2 + \left(\frac{\partial}{\partial y}\right)^2 + \left(\frac{\partial}{\partial z}\right)^2$$

$$= \frac{1}{r^2}\left[\frac{\partial}{\partial r}\left(r^2\frac{\partial}{\partial r}\right)\right. \qquad (11.142)$$

$$\left. + \frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta}\frac{\partial^2}{\partial\varphi^2}\right].$$

11.9.1   Separation of variables Ansatz

This can be exploited for a *separation of variable Ansatz*, which, according to Schrödinger, should be well known (in German *sattsam bekannt*) by now (cf chapter 10). We thus write the solution ψ as a product of functions of separate variables

$$\psi(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi) = R(r)Y_l^m(\theta,\varphi) \qquad (11.143)$$

That the angular part $\Theta(\theta)\Phi(\varphi)$ of this product will turn out to be the spherical harmonics $Y_l^m(\theta,\varphi)$ introduced earlier on page 228 is nontrivial

– indeed, at this point it is an *ad hoc* assumption. We will come back to its derivation in fuller detail later.

11.9.2   *Separation of the radial part from the angular one*

For the time being, let us first concentrate on the radial part $R(r)$. Let us first totally separate the variables of the Schrödinger equation (11.141) in radial coordinates

$$
\left\{ \frac{1}{r^2} \left[\frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) \right. \right.
$$
$$
\left. + \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right] \tag{11.144}
$$
$$
\left. + \frac{2\mu}{-\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0 r} + E \right) \right\} \psi(r,\theta,\varphi) = 0,
$$

and multiplying it with r^2

$$
\left\{ \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{2\mu r^2}{-\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0 r} + E \right) \right.
$$
$$
\left. + \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right\} \psi(r,\theta,\varphi) = 0, \tag{11.145}
$$

so that, after division by $\psi(r,\theta,\varphi) = R(r)\Theta(\theta)\Phi(\varphi)$ and writing separate variables on separate sides of the equation,

$$
\frac{1}{R(r)} \left\{ \frac{\partial}{\partial r} \left(r^2 \frac{\partial}{\partial r} \right) + \frac{2\mu r^2}{-\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0 r} + E \right) \right\} R(r)
$$
$$
= -\frac{1}{\Theta(\theta)\Phi(\varphi)} \left\{ \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right\} \Theta(\theta)\Phi(\varphi) \tag{11.146}
$$

Because the left hand side of this equation is independent of the angular variables θ and φ, and its right hand side is independent of the radius r, both sides have to be constant; say λ. Thus we obtain two differential equations for the radial and the angular part, respectively

$$
\left\{ \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{2\mu r^2}{-\hbar^2} \left(\frac{e^2}{4\pi\epsilon_0 r} + E \right) \right\} R(r) = \lambda R(r), \tag{11.147}
$$

and

$$
\left\{ \frac{1}{\sin\theta} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial}{\partial\theta} + \frac{1}{\sin^2\theta} \frac{\partial^2}{\partial\varphi^2} \right\} \Theta(\theta)\Phi(\varphi) = -\lambda\Theta(\theta)\Phi(\varphi). \tag{11.148}
$$

11.9.3   *Separation of the polar angle θ from the azimuthal angle φ*

As already hinted in Eq. (11.143) The angular portion can still be separated into a polar and an azimuthal part because, when multiplied by $\sin^2\theta/[\Theta(\theta)\Phi(\varphi)]$, Eq. (11.148) can be rewritten as

$$
\left\{ \frac{\sin\theta}{\Theta(\theta)} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial\Theta(\theta)}{\partial\theta} + \lambda\sin^2\theta \right\} + \frac{1}{\Phi(\varphi)} \frac{\partial^2\Phi(\varphi)}{\partial\varphi^2} = 0, \tag{11.149}
$$

and hence

$$
\frac{\sin\theta}{\Theta(\theta)} \frac{\partial}{\partial\theta} \sin\theta \frac{\partial\Theta(\theta)}{\partial\theta} + \lambda\sin^2\theta = -\frac{1}{\Phi(\varphi)} \frac{\partial^2\Phi(\varphi)}{\partial\varphi^2} = m^2, \tag{11.150}
$$

where m is some constant.

11.9.4   Solution of the equation for the azimuthal angle factor $\Phi(\varphi)$

The resulting differential equation for $\Phi(\varphi)$

$$\frac{d^2\Phi(\varphi)}{d\varphi^2} = -m^2\Phi(\varphi), \qquad (11.151)$$

has the general solution

$$\Phi(\varphi) = Ae^{im\varphi} + Be^{-im\varphi}. \qquad (11.152)$$

As Φ must obey the periodic boundary conditions $\Phi(\varphi) = \Phi(\varphi + 2\pi)$, m must be an integer. The two constants A, B must be equal if we require the system of functions $\{e^{im\varphi} \mid m \in \mathbb{Z}\}$ to be orthonormalized. Indeed, if we define

$$\Phi_m(\varphi) = Ae^{im\varphi} \qquad (11.153)$$

and require that it is normalized, it follows that

$$\begin{aligned}
&\int_0^{2\pi} \overline{\Phi}_m(\varphi)\Phi_m(\varphi)\,d\varphi \\
&= \int_0^{2\pi} \overline{A}e^{-im\varphi} Ae^{im\varphi}\,d\varphi \\
&\qquad = \int_0^{2\pi} |A|^2\,d\varphi \\
&\qquad\qquad = 2\pi|A|^2 \\
&\qquad\qquad\quad = 1,
\end{aligned} \qquad (11.154)$$

it is consistent to set

$$A = \frac{1}{\sqrt{2\pi}}; \qquad (11.155)$$

and hence,

$$\Phi_m(\varphi) = \frac{e^{im\varphi}}{\sqrt{2\pi}} \qquad (11.156)$$

Note that, for different $m \neq n$,

$$\begin{aligned}
&\int_0^{2\pi} \overline{\Phi}_n(\varphi)\Phi_m(\varphi)\,d\varphi \\
&= \int_0^{2\pi} \frac{e^{-in\varphi}}{\sqrt{2\pi}} \frac{e^{im\varphi}}{\sqrt{2\pi}}\,d\varphi \\
&\quad = \int_0^{2\pi} \frac{e^{i(m-n)\varphi}}{2\pi}\,d\varphi \\
&\quad = -\left.\frac{ie^{i(m-n)\varphi}}{2(m-n)\pi}\right|_{\varphi=0}^{\varphi=2\pi} \\
&\qquad\qquad = 0,
\end{aligned} \qquad (11.157)$$

because $m - n \in \mathbb{Z}$.

11.9.5   Solution of the equation for the polar angle factor $\Theta(\theta)$

The left hand side of Eq. (11.150) contains only the polar coordinate. Upon division by $\sin^2\theta$ we obtain

$$\begin{aligned}
&\frac{1}{\Theta(\theta)\sin\theta}\frac{d}{d\theta}\sin\theta\frac{d\Theta(\theta)}{d\theta} + \lambda = \frac{m^2}{\sin^2\theta}, \text{ or} \\
&\frac{1}{\Theta(\theta)\sin\theta}\frac{d}{d\theta}\sin\theta\frac{d\Theta(\theta)}{d\theta} - \frac{m^2}{\sin^2\theta} = -\lambda,
\end{aligned} \qquad (11.158)$$

Now, first, let us consider the case $m = 0$. With the variable substitution $x = \cos\theta$, and thus $\frac{dx}{d\theta} = -\sin\theta$ and $dx = -\sin\theta\, d\theta$, we obtain from (11.158)

$$\frac{d}{dx}\sin^2\theta\frac{d\Theta(x)}{dx} = -\lambda\Theta(x),$$

$$\frac{d}{dx}(1-x^2)\frac{d\Theta(x)}{dx} + \lambda\Theta(x) = 0, \qquad (11.159)$$

$$\left(x^2-1\right)\frac{d^2\Theta(x)}{dx^2} + 2x\frac{d\Theta(x)}{dx} = \lambda\Theta(x),$$

which is of the same form as the *Legendre equation* (11.122) mentioned on page 224.

Consider the series *Ansatz*

$$\Theta(x) = a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k + \cdots \qquad (11.160)$$

for solving (11.159). Insertion into (11.159) and comparing the coefficients of x for equal degrees yields the recursion relation

This is actually a "shortcut" solution of the Fuchian Equation mentioned earlier.

$$\left(x^2-1\right)\frac{d^2}{dx^2}[a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k + \cdots]$$

$$+2x\frac{d}{dx}[a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k + \cdots]$$

$$= \lambda[a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k + \cdots],$$

$$\left(x^2-1\right)[2a_2 + \cdots + k(k-1)a_k x^{k-2} + \cdots]$$

$$+[2xa_1 + 2x2a_2 x + \cdots + 2xka_k x^{k-1} + \cdots]$$

$$= \lambda[a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k + \cdots],$$

$$\left(x^2-1\right)[2a_2 + \cdots + k(k-1)a_k x^{k-2} + \cdots]$$

$$+[2a_1 x + 4a_2 x^2 + \cdots + 2ka_k x^k + \cdots]$$

$$= \lambda[a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k + \cdots], \qquad (11.161)$$

$$[2a_2 x^2 + \cdots + k(k-1)a_k x^k + \cdots]$$

$$-[2a_2 + \cdots + k(k-1)a_k x^{k-2} + \cdots]$$

$$+[2a_1 x + 4a_2 x^2 + \cdots + 2ka_k x^k + \cdots]$$

$$= \lambda[a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k + \cdots],$$

$$[2a_2 x^2 + \cdots + k(k-1)a_k x^k + \cdots]$$

$$-[2a_2 + \cdots + k(k-1)a_k x^{k-2}$$

$$+(k+1)ka_{k+1}x^{k-1} + (k+2)(k+1)a_{k+2}x^k + \cdots]$$

$$+[2a_1 x + 4a_2 x^2 + \cdots + 2ka_k x^k + \cdots]$$

$$= \lambda[a_0 + a_1 x + a_2 x^2 + \cdots + a_k x^k + \cdots],$$

and thus, by taking all polynomials of the order of k and proportional to x^k, so that, for $x^k \neq 0$ (and thus excluding the trivial solution),

$$k(k-1)a_k x^k - (k+2)(k+1)a_{k+2}x^k + 2ka_k x^k - \lambda a_k x^k = 0,$$

$$k(k+1)a_k - (k+2)(k+1)a_{k+2} - \lambda a_k = 0, \qquad (11.162)$$

$$a_{k+2} = a_k\frac{k(k+1)-\lambda}{(k+2)(k+1)}.$$

In order to converge also for $x = \pm 1$, and hence for $\theta = 0$ and $\theta = \pi$, the sum in (11.160) has to have only *a finite number of terms*. Because if the

sum would be infinite, the terms a_k, for large k, would be dominated by $a_{k-2}O(k^2/k^2) = a_{k-2}O(1)$, and thus would converge to $a_k \xrightarrow{k \to \infty} a_\infty$ with constant $a_\infty \neq 0$, and thus Θ would diverge as $\Theta(1) \overset{k \to \infty}{\approx} ka_\infty \xrightarrow{k \to \infty} \infty$. That means that, in Eq. (11.162) for some $k = l \in \mathbb{N}$, the coefficient $a_{l+2} = 0$ has to vanish; thus

$$\lambda = l(l+1). \tag{11.163}$$

This results in *Legendre polynomials* $\Theta(x) \equiv P_l(x)$.

Let us now shortly mention the case $m \neq 0$. With the same variable substitution $x = \cos\theta$, and thus $\frac{dx}{d\theta} = -\sin\theta$ and $dx = -\sin\theta \, d\theta$ as before, the equation for the polar angle dependent factor (11.158) becomes

$$\left\{ \frac{d}{dx}(1-x^2)\frac{d}{dx} + l(l+1) - \frac{m^2}{1-x^2} \right\} \Theta(x) = 0, \tag{11.164}$$

This is exactly the form of the general Legendre equation (11.134), whose solution is a multiple of the associated Legendre polynomial $P_l^m(x)$, with $|m| \leq l$.

Note (without proof) that, for equal m, the $P_l^m(x)$ satisfy the orthogonality condition

$$\int_{-1}^{1} P_l^m(x) P_{l'}^m(x) dx = \frac{2(l+m)!}{(2l+1)(l-m)!} \delta_{ll'}. \tag{11.165}$$

Therefore we obtain a normalized polar solution by dividing $P_l^m(x)$ by $\{[2(l+m)!] / [(2l+1)(l-m)!]\}^{1/2}$.

In putting both normalized polar and azimuthal angle factors together we arrive at the spherical harmonics (11.139); that is,

$$\Theta(\theta)\Phi(\varphi) = \sqrt{\frac{(2l+1)(l-m)!}{2(l+m)!}} P_l^m(\cos\theta) \frac{e^{im\varphi}}{\sqrt{2\pi}} = Y_l^m(\theta, \varphi) \tag{11.166}$$

for $-l \leq m \leq l$, $l \in \mathbb{N}_0$. Note that the discreteness of these solutions follows from physical requirements about their finite existence.

11.9.6   Solution of the equation for radial factor $R(r)$

The solution of the equation (11.147)

$$\left\{ \frac{d}{dr} r^2 \frac{d}{dr} + \frac{2\mu r^2}{-\hbar^2}\left(\frac{e^2}{4\pi\epsilon_0 r} + E \right) \right\} R(r) = l(l+1)R(r) \text{ , or}$$
$$-\frac{1}{R(r)} \frac{d}{dr} r^2 \frac{d}{dr} R(r) + l(l+1) - 2\frac{\mu e^2}{4\pi\epsilon_0 -\hbar^2} r = \frac{2\mu}{-\hbar^2} r^2 E \tag{11.167}$$

for the radial factor $R(r)$ turned out to be the most difficult part for Schrödinger [14].

Note that, since the additive term $l(l+1)$ in (11.167) is non-dimensional, so must be the other terms. We can make this more explicit by the substitution of variables.

First, consider $y = \frac{r}{a_0}$ obtained by dividing r by the *Bohr radius*

$$a_0 = \frac{4\pi\epsilon_0 -\hbar^2}{m_e e^2} \approx 5 \, 10^{-11} m, \tag{11.168}$$

[14] Walter Moore. *Schrödinger: Life and Thought.* Cambridge University Press, Cambridge, UK, 1989

thereby assuming that the reduced mass is equal to the electron mass $\mu \approx m_e$. More explicitly, $r = y a_0 = y(4\pi\epsilon_0 \hbar^2)/(m_e e^2)$, or $y = r/a_0 = r(m_e e^2)/(4\pi\epsilon_0 \hbar^2)$. Furthermore, let us define $\varepsilon = E\frac{2\mu a_0^2}{\hbar^2}$.

These substitutions yield

$$-\frac{1}{R(y)}\frac{d}{dy}y^2\frac{d}{dy}R(y) + l(l+1) - 2y = y^2\varepsilon, \text{ or}$$

$$-y^2\frac{d^2}{dy^2}R(y) - 2y\frac{d}{dy}R(y) + \left[l(l+1) - 2y - \varepsilon y^2\right]R(y) = 0. \tag{11.169}$$

Now we introduce a new function $\hat{R}$ *via*

$$R(\xi) = \xi^l e^{-\frac{1}{2}\xi}\hat{R}(\xi), \tag{11.170}$$

with $\xi = \frac{2y}{n}$ and by replacing the energy variable with $\varepsilon = -\frac{1}{n^2}$. (It will later be argued that ε must be dicrete; with $n \in \mathbb{N} - 0$.) This yields

$$\xi\frac{d^2}{d\xi^2}\hat{R}(\xi) + [2(l+1) - \xi]\frac{d}{d\xi}\hat{R}(\xi) + (n - l - 1)\hat{R}(\xi) = 0. \tag{11.171}$$

The discretization of n can again be motivated by requiring physical properties from the solution; in particular, convergence. Consider again a series solution *Ansatz*

$$\hat{R}(\xi) = c_0 + c_1\xi + c_2\xi^2 + \cdots + c_k\xi^k + \cdots, \tag{11.172}$$

which, when inserted into (11.169), yields

$$\xi\frac{d^2}{d^2\xi}[c_0 + c_1\xi + c_2\xi^2 + \cdots + c_k\xi^k + \cdots]$$
$$+[2(l+1) - \xi]\frac{d}{dy}[c_0 + c_1\xi + c_2\xi^2 + \cdots + c_k\xi^k + \cdots]$$
$$+(n - l - 1)[c_0 + c_1\xi + c_2\xi^2 + \cdots + c_k\xi^k + \cdots]$$
$$= 0,$$
$$\xi[2c_2 + \cdots k(k-1)c_k\xi^{k-2} + \cdots]$$
$$+[2(l+1) - \xi][c_1 + 2c_2\xi + \cdots + kc_k\xi^{k-1} + \cdots]$$
$$+(n - l - 1)[c_0 + c_1\xi + c_2\xi^2 + \cdots + c_k\xi^k + \cdots]$$
$$= 0,$$
$$[2c_2\xi + \cdots + k(k-1)c_k\xi^{k-1} + \cdots] \tag{11.173}$$
$$+2(l+1)[c_1 + 2c_2\xi + \cdots + k + c_k\xi^{k-1} + \cdots]$$
$$-[c_1\xi + 2c_2\xi^2 + \cdots + kc_k\xi^k + \cdots]$$
$$+(n - l - 1)[c_0 + c_1\xi + c_2\xi^2 + \cdots c_k\xi^k + \cdots]$$
$$= 0,$$
$$[2c_2\xi + \cdots + k(k-1)c_k\xi^{k-1} + k(k+1)c_{k+1}\xi^k + \cdots]$$
$$+2(l+1)[c_1 + 2c_2\xi + \cdots + kc_k\xi^{k-1} + (k+1)c_{k+1}\xi^k + \cdots]$$
$$-[c_1\xi + 2c_2\xi^2 + \cdots + kc_k\xi^k + \cdots]$$
$$+(n - l - 1)[c_0 + c_1\xi + c_2\xi^2 + \cdots + c_k\xi^k + \cdots]$$
$$= 0,$$

so that, by comparing the coefficients of ξ^k, we obtain

$$k(k+1)c_{k+1}\xi^k + 2(l+1)(k+1)c_{k+1}\xi^k = kc_k\xi^k - (n-l-1)c_k\xi^k,$$

$$c_{k+1}[k(k+1) + 2(l+1)(k+1)] = c_k[k-(n-l-1)],$$

$$c_{k+1}(k+1)(k+2l+2) = c_k(k-n+l+1),$$ \quad (11.174)

$$c_{k+1} = c_k \frac{k-n+l+1}{(k+1)(k+2l+2)}.$$

Because of convergence of $\hat{R}$ and thus of R – note that, for large ξ and k, the k'th term in Eq. (11.172) determining $\hat{R}(\xi)$ would behave as $\xi^k/k!$ and thus $\hat{R}(\xi)$ would roughly behave as e^ξ – the series solution (11.172) should terminate at some $k = n - l - 1$, or $n = k + l + 1$. Since k, l, and 1 are all integers, n must be an integer as well. And since $k \geq 0$, n must be at least $l + 1$, or

$$l \leq n - 1. \quad (11.175)$$

Thus, we end up with an *associated Laguerre equation* of the form

$$\left\{ \xi \frac{d^2}{d\xi^2} + [2(l+1) - \xi]\frac{d}{d\xi} + (n-l-1) \right\} \hat{R}(\xi) = 0, \text{ with } n \geq l+1, \text{ and } n, l \in \mathbb{Z}.$$
$$(11.176)$$

Its solutions are the *associated Laguerre polynomials* L_{n+l}^{2l+1} which are the $(2l+1)$-th derivatives of the Laguerre's polynomials L_{n+l}; that is,

$$L_n(x) = e^x \frac{d^n}{dx^n}\left(x^n e^{-x}\right),$$

$$L_n^m(x) = \frac{d^m}{dx^m} L_n(x). \quad (11.177)$$

This yields a normalized wave function

$$R_n(r) = \mathcal{N}\left(\frac{2r}{na_0}\right)^l e^{-\frac{r}{a_0 n}} L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right), \text{ with}$$

$$\mathcal{N} = -\frac{2}{n^2}\sqrt{\frac{(n-l-1)!}{[(n+l)!a_0]^3}}, \quad (11.178)$$

where $\mathcal{N}$ stands for the normalization factor.

11.9.7   Composition of the general solution of the Schrödinger Equation

Now we shall coagulate and combine the factorized solutions (11.143) into a complete solution of the Schrödinger equation for $n+1$, l, $|m| \in \mathbb{N}_0$, $0 \leq l \leq n-1$, and $|m| \leq l$,

Always remember the alchemic principle of *solve et coagula*!

$$\psi_{n,l,m}(r,\theta,\varphi)$$

$$= R_n(r) Y_l^m(\theta,\varphi)$$

$$= -\frac{2}{n^2}\sqrt{\frac{(n-l-1)!}{[(n+l)!a_0]^3}}\left(\frac{2r}{na_0}\right)^l e^{-\frac{r}{a_0 n}} L_{n+l}^{2l+1}\left(\frac{2r}{na_0}\right)\sqrt{\frac{(2l+1)(l-m)!}{2(l+m)!}}P_l^m(\cos\theta)\frac{e^{im\varphi}}{\sqrt{2\pi}}.$$
$$(11.179)$$

12

Divergent series

In this final chapter we will consider *divergent series*, which, as has already been mentioned earlier, seem to have been "invented by the devil"[1]. Unfortunately such series occur very often in physical situations; for instance in celestial mechanics or in quantum field theory[2], and one may wonder with Abel why, *"for the most part, it is true that the results are correct, which is very strange"*[3]. On the other hand, there appears to be another view on diverging series, a view that has been expressed by Berry as follows[4]: *"… an asymptotic series … is a compact encoding of a function, and its divergence should be regarded not as a deficiency but as a source of information about the function."*

12.1 Convergence and divergence

Let us first define *convergence* in the context of series. A series

$$\sum_{j=0}^{\infty} a_j = a_0 + a_1 + a_2 + \cdots \tag{12.1}$$

is said to converge to the *sum s*, if the *partial sum*

$$s_n = \sum_{j=0}^{n} a_j = a_0 + a_1 + a_2 + \cdots + a_n \tag{12.2}$$

tends to a finite limit s when $n \to \infty$; otherwise it is said to be divergent.

One of the most prominent series is the Grandi's series, sometimes also referred to as Leibniz series[5]

$$s = \sum_{j=0}^{\infty} (-1)^j = 1 - 1 + 1 - 1 + 1 - \cdots, \tag{12.3}$$

whose summands may be – inconsistently – "rearranged," yielding

$$\text{either } 1 - 1 + 1 - 1 + 1 - 1 + \cdots = (1-1) + (1-1) + (1-1) - \cdots = 0$$
$$\text{or } 1 - 1 + 1 - 1 + 1 - 1 + \cdots = 1 + (-1+1) + (-1+1) + \cdots = 1.$$

One could tentatively associate the arithmetical average $\frac{1}{2}$ to represent "the sum of Grandi's series."

Note that, by *Riemann's rearrangement theorem*, even convergent series which do not absolutely converge (i.e., $\sum_{j=0}^{n} a_j$ converges but

[1] Godfrey Harold Hardy. *Divergent Series*. Oxford University Press, 1949

[2] John P. Boyd. The devil's invention: Asymptotic, superasymptotic and hyperasymptotic series. *Acta Applicandae Mathematica*, 56:1–98, 1999. ISSN 0167-8019. DOI: 10.1023/A:1006145903624. URL http://doi.org/10.1023/A:1006145903624; Freeman J. Dyson. Divergence of perturbation theory in quantum electrodynamics. *Phys. Rev.*, 85(4):631–632, Feb 1952. DOI: 10.1103/PhysRev.85.631. URL http://doi.org/10.1103/PhysRev.85.631; Sergio A. Pernice and Gerardo Oleaga. Divergence of perturbation theory: Steps towards a convergent series. *Physical Review D*, 57:1144–1158, Jan 1998. DOI: 10.1103/PhysRevD.57.1144. URL http://doi.org/10.1103/PhysRevD.57.1144; and Ulrich D. Jentschura. Resummation of nonalternating divergent perturbative expansions. *Physical Review D*, 62:076001, Aug 2000. DOI: 10.1103/PhysRevD.62.076001. URL http://doi.org/10.1103/PhysRevD.62.076001

[3] Christiane Rousseau. Divergent series: Past, present, future …. preprint, 2004. URL http://www.dms.umontreal.ca/~rousseac/divergent.pdf

[4] Michael Berry. Asymptotics, superasymptotics, hyperasymptotics… In Harvey Segur, Saleh Tanveer, and Herbert Levine, editors, *Asymptotics beyond All Orders*, volume 284 of *NATO ASI Series*, pages 1–14. Springer, 1992. ISBN 978-1-4757-0437-2. DOI: 10.1007/978-1-4757-0435-8. URL http://doi.org/10.1007/978-1-4757-0435-8

[5] Gottfried Wilhelm Leibniz. Letters LXX, LXXI. In Carl Immanuel Gerhardt, editor, *Briefwechsel zwischen Leibniz und Christian Wolf. Handschriften der Königlichen Bibliothek zu Hannover,*. H. W. Schmidt, Halle, 1860. URL http://books.google.de/books?id=TUkJAAAAQAAJ; Charles N. Moore. *Summable Series and Convergence Factors*. American Mathematical Society, New York, NY, 1938; Godfrey Harold Hardy. *Divergent Series*. Oxford University Press, 1949; and Graham Everest, Alf van der Poorten, Igor Shparlinski, and Thomas Ward. *Recurrence sequences. Volume 104 in the AMS Surveys and Mono-*

$\sum_{j=0}^{n} |a_j|$ diverges) can converge to any arbitrary (even infinite) value by permuting (rearranging) the (ratio of) positive and negative terms (the series of which must both be divergent).

The Grandi series is a particular case $q = -1$ of a *geometric series*

$$s(q) = \sum_{j=0}^{\infty} q^j = 1 + q + q^2 + q^3 + \cdots = 1 + qs \qquad (12.4)$$

which, since $s(q) = 1 + qs(q)$, for $|q| < 1$, converges to

$$s(q) = \sum_{j=0}^{\infty} q^j = \frac{1}{1-q}. \qquad (12.5)$$

One way to sum the Grandi series is by "continuing" Eq. (12.5) for arbitrary $q \neq 1$, thereby defining the *Abel sum*

$$\sum_{j=0}^{\infty} (-1)^j \overset{\text{A}}{=} \frac{1}{1-(-1)} = \frac{1}{2}. \qquad (12.6)$$

Another divergent series, which can be obtained by formally expanding the square of the Abel sum of the Grandi series $s^2 \overset{\text{A}}{=} (1 + x)^{-2}$ around 0 and inserting $x = 1$ [6] is

$$s^2 = \left(\sum_{j=0}^{\infty} (-1)^j \right) \left(\sum_{k=0}^{\infty} (-1)^k \right) = \sum_{j=0}^{\infty} (-1)^{j+1} j = 0 + 1 - 2 + 3 - 4 + 5 - \cdots. \qquad (12.7)$$

In the same sense as the Grandi series, this yields the Abel sum $s^2 \overset{\text{A}}{=} \frac{1}{4}$.

Note that, once established this identifican, all of Abel's hell breaks loose: One could, for instance, "compute the finite sum of all natural numbers"

$$S = \sum_{j=0}^{\infty} j = 1 + 2 + 3 + 4 + 5 + \cdots \overset{\text{A}}{=} -\frac{1}{12} \qquad (12.8)$$

by sorting out

$$S - \frac{1}{4} \overset{\text{A}}{=} S - s^2 = 1 + 2 + 3 + 4 + 5 + \cdots - (1 - 2 + 3 - 4 + 5 - \cdots)$$
$$= 4 + 8 + 12 + \cdots = 4S, \qquad (12.9)$$

so that $3S \overset{\text{A}}{=} -\frac{1}{4}$, and, finally, $S \overset{\text{A}}{=} -\frac{1}{12}$.

The Riemann zeta function (sometimes also referred to as the Euler-Riemann zeta function), defined for $\Re t > 1$ by

$$\zeta(t) \overset{\text{def}}{=} \sum_{n=1}^{\infty} \frac{1}{n^t} = \prod_{p \text{ prime}} \left(\sum_{n=1}^{\infty} p^{-nt} \right) = \prod_{p \text{ prime}} \frac{1}{1 - \frac{1}{p^t}}, \qquad (12.10)$$

can be continued analytically to all complex values $t \neq 1$, so that, for $t = -1$, $\zeta(-1) = S$.

Note that the sequence of its partial sums $s_n^2 = \sum_{j=0}^{n} (-1)^{j+1} j$ yield every integer once; that is, $s_0^2 = 0$, $s_1^2 = 0 + 1 = 1$, $s_2^2 = 0 + 1 - 2 = -1$, $s_3^2 = 0 + 1 - 2 + 3 = 2$, $s_4^2 = 0 + 1 - 2 + 3 - 4 = -2$, ..., $s_n^2 = -\frac{n}{2}$ for even n, and $s_n^2 = -\frac{n+1}{2}$ for odd n. It thus establishes a strict one-to-one mapping $s^2 : \mathbb{N} \mapsto \mathbb{Z}$ of the natural numbers onto the integers.

12.2   Euler differential equation

In what follows we demonstrate that divergent series may make sense, in the way Abel wondered. That is, we shall show that the first partial sums of divergent series may yield "good" approximations of the exact result; and that, from a certain point onward, more terms contributing to the sum might worsen the approximation rather an make it better – a situation totally different from convergent series, where more terms always result in better approximations.

Let us, with Rousseau, for the sake of demonstration of the former situation, consider the *Euler differential equation*

$$\left(x^2 \frac{d}{dx} + 1\right) y(x) = x, \text{ or } \left(\frac{d}{dx} + \frac{1}{x^2}\right) y(x) = \frac{1}{x}. \tag{12.11}$$

We shall solve this equation by two methods: we shall, on the one hand, present a divergent series solution, and on the other hand, an exact solution. Then we shall compare the series approximation to the exact solution by considering the difference.

A series solution of the Euler differential equation can be given by

$$y_s(x) = \sum_{j=0}^{\infty} (-1)^j j! x^{j+1}. \tag{12.12}$$

That (12.12) solves (12.11) can be seen by inserting the former into the latter; that is,

$$\left(x^2 \frac{d}{dx} + 1\right) \sum_{j=0}^{\infty} (-1)^j j! x^{j+1} = x,$$

$$\sum_{j=0}^{\infty} (-1)^j (j+1)! x^{j+2} + \sum_{j=0}^{\infty} (-1)^j j! x^{j+1} = x,$$

[change of variable in the first sum: $j \to j - 1$]

$$\sum_{j=1}^{\infty} (-1)^{j-1} (j+1-1)! x^{j+2-1} + \sum_{j=0}^{\infty} (-1)^j j! x^{j+1} = x,$$

$$\sum_{j=1}^{\infty} (-1)^{j-1} j! x^{j+1} + x + \sum_{j=1}^{\infty} (-1)^j j! x^{j+1} = x, \tag{12.13}$$

$$x + \sum_{j=1}^{\infty} (-1)^j \left[(-1)^{-1} + 1\right] j! x^{j+1} = x,$$

$$x + \sum_{j=1}^{\infty} (-1)^j \underbrace{[-1+1]}_{=0} j! x^{j+1} = x,$$

$$x = x.$$

On the other hand, an exact solution can be found by *quadrature;* that is, by explicit integration (see, for instance, Chapter one of Ref. [7]). Consider the homogenuous first order differential equation

$$\left(\frac{d}{dx} + p(x)\right) y(x) = 0,$$

$$\text{or } \frac{dy(x)}{dx} = -p(x) y(x), \tag{12.14}$$

$$\text{or } \frac{dy(x)}{y(x)} = -p(x) dx.$$

[7] Garrett Birkhoff and Gian-Carlo Rota. *Ordinary Differential Equations.* John Wiley & Sons, New York, Chichester, Brisbane, Toronto, fourth edition, 1959, 1960, 1962, 1969, 1978, and 1989

Integrating both sides yields

$$\log|y(x)| = -\int p(x)dx + C \text{ , or } |y(x)| = Ke^{-\int p(x)dx}, \qquad (12.15)$$

where C is some constant, and $K = e^C$. Let $P(x) = \int p(x)dx$. Hence, heuristically, $y(x)e^{P(x)}$ is constant, as can also be seen by explicit differentiation of $y(x)e^{P(x)}$; that is,

$$\begin{aligned}
\frac{d}{dx}y(x)e^{P(x)} &= e^{P(x)}\frac{dy(x)}{dx} + y(x)\frac{d}{dx}e^{P(x)} \\
&= e^{P(x)}\frac{dy(x)}{dx} + y(x)p(x)e^{P(x)} \\
&= e^{P(x)}\left(\frac{d}{dx} + p(x)\right)y(x) \\
&= 0
\end{aligned} \qquad (12.16)$$

if and, since $e^{P(x)} \neq 0$, only if $y(x)$ satisfies the homogenuous equation (12.14). Hence,

$$y(x) = ce^{-\int p(x)dx} \text{ is the solution of}$$
$$\left(\frac{d}{dx} + p(x)\right)y(x) = 0 \qquad (12.17)$$

for some constant c.

Similarly, we can again find a solution to the inhomogenuos first order differential equation

$$\begin{aligned}
\left(\frac{d}{dx} + p(x)\right)y(x) + q(x) &= 0, \\
\text{or } \left(\frac{d}{dx} + p(x)\right)y(x) &= -q(x)
\end{aligned} \qquad (12.18)$$

by differentiating the function $y(x)e^{P(x)} = y(x)e^{\int p(x)dx}$; that is,

$$\begin{aligned}
\frac{d}{dx}y(x)e^{\int p(x)dx} &= e^{\int p(x)dx}\frac{d}{dx}y(x) + p(x)e^{\int p(x)dx}y(x) \\
&= e^{\int p(x)dx}\underbrace{\left(\frac{d}{dx} + p(x)\right)y(x)}_{=q(x)} \\
&= -e^{\int p(x)dx}q(x).
\end{aligned} \qquad (12.19)$$

Hence, for some constant y_0 and some a, b, we must have, by integration,

$$\begin{aligned}
\int_b^x \frac{d}{dt}y(t)e^{\int_a^t p(s)ds}dt &= y(x)e^{\int_a^x p(t)dt} \\
&= y_0 - \int_b^x e^{\int_a^t p(s)ds}q(t)dt, \quad (12.20)
\end{aligned}$$

and hence $y(x) = y_0 e^{-\int_a^x p(t)dt} - e^{-\int_a^x p(t)dt}\int_b^x e^{\int_a^t p(s)ds}q(t)dt$.

If $a = b$, then $y(b) = y_0$.

Coming back to the Euler differential equation and identifying $p(x) = 1/x^2$ and $q(x) = -1/x$ we obtain, up to a constant, with $b = 0$

and arbitrary constant $a \neq 0$,

$$
\begin{aligned}
y(x) &= -e^{-\int_a^x \frac{dt}{t^2}} \int_0^x e^{\int_a^t \frac{ds}{s^2}} \left(-\frac{1}{t} \right) dt \\
&= e^{-\left(-\frac{1}{t} \right)|_a^x} \int_0^x e^{-\frac{1}{s}|_a^t} \left(\frac{1}{t} \right) dt \\
&= e^{\frac{1}{x} - \frac{1}{a}} \int_0^x e^{-\frac{1}{t} + \frac{1}{a}} \left(\frac{1}{t} \right) dt \\
&= e^{\frac{1}{x}} \underbrace{e^{-\frac{1}{a} + \frac{1}{a}}}_{=e^0 = 1} \int_0^x e^{-\frac{1}{t}} \left(\frac{1}{t} \right) dt \\
&= e^{\frac{1}{x}} e^{-\frac{1}{a}} \int_0^x e^{-\frac{1}{t}} e^{\frac{1}{a}} \left(\frac{1}{t} \right) dt \\
&= e^{\frac{1}{x}} \int_0^x \frac{e^{-\frac{1}{t}}}{t} dt \\
&= \int_0^x \frac{e^{\frac{1}{x} - \frac{1}{t}}}{t} dt.
\end{aligned}
\tag{12.21}
$$

With a change of the integration variable

$$
\begin{aligned}
\frac{\xi}{x} &= \frac{1}{t} - \frac{1}{x}, \text{ and thus } \xi = \frac{x}{t} - 1, \text{ and } t = \frac{x}{1+\xi}, \\
\frac{dt}{d\xi} &= -\frac{x}{(1+\xi)^2}, \text{ and thus } dt = -\frac{x}{(1+\xi)^2} d\xi, \\
\text{and thus } \frac{dt}{t} &= \frac{-\frac{x}{(1+\xi)^2}}{\frac{x}{1+\xi}} d\xi = -\frac{d\xi}{1+\xi},
\end{aligned}
\tag{12.22}
$$

the integral (12.21) can be rewritten as

$$
y(x) = \int_\infty^0 \left(-\frac{e^{-\frac{\xi}{x}}}{1+\xi} \right) d\xi = \int_0^\infty \frac{e^{-\frac{\xi}{x}}}{1+\xi} d\xi.
\tag{12.23}
$$

It is proportional to the *Stieltjes Integral* [8]

$$
S(x) = \int_0^\infty \frac{e^{-\xi}}{1 + x\xi} d\xi.
\tag{12.24}
$$

Note that whereas the series solution $y_s(x)$ diverges for all nonzero x, the solution $y(x)$ in (12.23) converges and is well defined for all $x \geq 0$.

Let us now estimate the absolute difference between $y_{s_k}(x)$ which represents the partial sum "$y_s(x)$ truncated after the kth term" and $y(x)$; that is, let us consider

$$
|y(x) - y_{s_k}(x)| = \left| \int_0^\infty \frac{e^{-\frac{\xi}{x}}}{1+\xi} d\xi - \sum_{j=0}^k (-1)^j j! x^{j+1} \right|.
\tag{12.25}
$$

For any $x \geq 0$ this difference can be estimated [9] by a bound from above

$$
|R_k(x)| \stackrel{\text{def}}{=} |y(x) - y_{s_k}(x)| \leq k! x^{k+1},
\tag{12.26}
$$

that is, this difference between the exact solution $y(x)$ and the diverging partial series $y_{s_k}(x)$ is smaller than the first neglected term; and all subsequent ones.

For a proof, observe that, since a partial *geometric series* is the sum of all the numbers in a geometric progression up to a certain power; that is,

$$
\sum_{k=0}^n r^k = 1 + r + r^2 + \cdots + r^k + \cdots + r^n.
\tag{12.27}
$$

[8] Carl M. Bender Steven A. Orszag. *Advanced Mathematical Methods for Scientists and Enineers*. McGraw-Hill, New York, NY, 1978; and John P. Boyd. The devil's invention: Asymptotic, superasymptotic and hyperasymptotic series. *Acta Applicandae Mathematica*, 56:1–98, 1999. ISSN 0167-8019. DOI: 10.1023/A:1006145903624. URL http://doi.org/10.1023/A:1006145903624

[9] Christiane Rousseau. Divergent series: Past, present, future …. preprint, 2004. URL http://www.dms.umontreal.ca/~rousseac/divergent.pdf

By multiplying both sides with $1 - r$, the sum (12.27) can be rewritten as

$$(1 - r) \sum_{k=0}^{n} r^k = (1 - k)(1 + r + r^2 + \cdots + r^k + \cdots + r^n)$$

$$= 1 + r + r^2 + \cdots + r^k + \cdots + r^n - r(1 + r + r^2 + \cdots + r^k + \cdots + r^n + r^n)$$

$$= 1 + r + r^2 + \cdots + r^k + \cdots + r^n - (r + r^2 + \cdots + r^k + \cdots + r^n + r^{n+1})$$

$$= 1 - r^{n+1},$$

$$(12.28)$$

and, since the middle terms all cancel out,

$$\sum_{k=0}^{n} r^k = \frac{1 - r^{n+1}}{1 - r}, \quad \text{or} \quad \sum_{k=0}^{n-1} r^k = \frac{1 - r^n}{1 - r} = \frac{1}{1 - r} - \frac{r^n}{1 - r}. \qquad (12.29)$$

Thus, for $r = -\zeta$, it is true that

$$\frac{1}{1 + \zeta} = \sum_{k=0}^{n-1} (-1)^k \zeta^k + (-1)^n \frac{\zeta^n}{1 + \zeta}. \qquad (12.30)$$

Thus

$$f(x) = \int_0^\infty \frac{e^{-\frac{\zeta}{x}}}{1 + \zeta} d\zeta$$

$$= \int_0^\infty e^{-\frac{\zeta}{x}} \left(\sum_{k=0}^{n-1} (-1)^k \zeta^k + (-1)^n \frac{\zeta^n}{1 + \zeta} \right) d\zeta \qquad (12.31)$$

$$= \sum_{k=0}^{n-1} \int_0^\infty (-1)^k \zeta^k e^{-\frac{\zeta}{x}} d\zeta + \int_0^\infty (-1)^n \frac{\zeta^n e^{-\frac{\zeta}{x}}}{1 + \zeta} d\zeta.$$

Since

$$k! = \Gamma(k + 1) = \int_0^\infty z^k e^{-z} dz, \qquad (12.32)$$

one obtains

$$\int_0^\infty \zeta^k e^{-\frac{\zeta}{x}} d\zeta$$

$$[\text{substitution: } z = \frac{\zeta}{x}, d\zeta = x dz] \qquad (12.33)$$

$$= \int_0^\infty x^{k+1} z^k e^{-z} dz$$

$$= x^{k+1} k!,$$

and hence

$$f(x) = \sum_{k=0}^{n-1} \int_0^\infty (-1)^k \zeta^k e^{-\frac{\zeta}{x}} d\zeta + \int_0^\infty (-1)^n \frac{\zeta^n e^{-\frac{\zeta}{x}}}{1 + \zeta} d\zeta$$

$$= \sum_{k=0}^{n-1} (-1)^k x^{k+1} k! + \int_0^\infty (-1)^n \frac{\zeta^n e^{-\frac{\zeta}{x}}}{1 + \zeta} d\zeta \qquad (12.34)$$

$$= f_n(x) + R_n(x),$$

where $f_n(x)$ represents the partial sum of the power series, and $R_n(x)$ stands for the remainder, the difference between $f(x)$ and $f_n(x)$. The absolute of the remainder can be estimated by

$$|R_n(x)| = \int_0^\infty \frac{\zeta^n e^{-\frac{\zeta}{x}}}{1 + \zeta} d\zeta$$

$$\leq \int_0^\infty \zeta^n e^{-\frac{\zeta}{x}} d\zeta \qquad (12.35)$$

$$= n! x^{n+1}.$$

12.2.1 Borel's resummation method – "The Master forbids it"

In what follows we shall again follow Christiane Rousseau's treatment [10] and use a resummation method invented by Borel [11] to obtain the exact convergent solution (12.23) of the Euler differential equation (12.11) from the divergent series solution (12.12). First we can rewrite a suitable infinite series by an integral representation, thereby using the integral representation of the factorial (12.32) as follows:

$$\sum_{j=0}^{\infty} a_j = \sum_{j=0}^{\infty} a_j \frac{j!}{j!} = \sum_{j=0}^{\infty} \frac{a_j}{j!} j!$$

$$= \sum_{j=0}^{\infty} \frac{a_j}{j!} \int_0^{\infty} t^j e^{-t} dt \overset{\text{B}}{=} \int_0^{\infty} \left(\sum_{j=0}^{\infty} \frac{a_j t^j}{j!} \right) e^{-t} dt. \qquad (12.36)$$

A series $\sum_{j=0}^{\infty} a_j$ is *Borel summable* if $\sum_{j=0}^{\infty} \frac{a_j t^j}{j!}$ has a non-zero radius of convergence, if it can be extended along the positive real axis, and if the integral (12.36) is convergent. This integral is called the *Borel sum* of the series.

In the case of the series solution of the Euler differential equation, $a_j = (-1)^j j! x^{j+1}$ [cf. Eq. (12.12)]. Thus,

$$\sum_{j=0}^{\infty} \frac{a_j t^j}{j!} = \sum_{j=0}^{\infty} \frac{(-1)^j j! x^{j+1} t^j}{j!} = x \sum_{j=0}^{\infty} (-xt)^j = \frac{x}{1+xt}, \qquad (12.37)$$

and therefore, with the substitionion $xt = \zeta$, $dt = \frac{d\zeta}{x}$

$$\sum_{j=0}^{\infty} (-1)^j j! x^{j+1} \overset{\text{B}}{=} \int_0^{\infty} \sum_{j=0}^{\infty} \frac{a_j t^j}{j!} e^{-t} dt = \int_0^{\infty} \frac{x}{1+xt} e^{-t} dt = \int_0^{\infty} \frac{e^{-\frac{\zeta}{x}}}{1+\zeta} d\zeta, \qquad (12.38)$$

which is the exact solution (12.23) of the Euler differential equation (12.11).

We can also find the Borel sum (which in this case is equal to the Abel sum) of the Grandi series (12.3) by

$$s = \sum_{j=0}^{\infty} (-1)^j \overset{\text{B}}{=} \int_0^{\infty} \left(\sum_{j=0}^{\infty} \frac{(-1)^j t^j}{j!} \right) e^{-t} dt$$

$$= \int_0^{\infty} \left(\sum_{j=0}^{\infty} \frac{(-t)^j}{j!} \right) e^{-t} dt = \int_0^{\infty} e^{-2t} dt$$

$$[\text{variable substitution } 2t = \zeta, dt = \frac{1}{2} d\zeta] \qquad (12.39)$$

$$= \frac{1}{2} \int_0^{\infty} e^{-\zeta} d\zeta$$

$$= \frac{1}{2} \left(-e^{-\zeta} \right) \Big|_{\zeta=0}^{\infty} = \frac{1}{2} \left(-e^{-\infty} + e^{-0} \right) = \frac{1}{2}.$$

[10] Christiane Rousseau. Divergent series: Past, present, future …. preprint, 2004. URL http://www.dms.umontreal.ca/~rousseac/divergent.pdf

For more resummation techniques, please see Chapter 16 of

Hagen Kleinert and Verena Schulte-Frohlinde. *Critical Properties of ϕ^4-Theories*. World scientific, Singapore, 2001. ISBN 9810246595

[11] Émile Borel. Mémoire sur les séries divergentes. *Annales scientifiques de l'École Normale Supérieure*, 16:9–131, 1899. URL http://eudml.org/doc/81143

"The idea that a function could be determined by a divergent asymptotic series was a foreign one to the nineteenth century mind. Borel, then an unknown young man, discovered that his summation method gave the "right" answer for many classical divergent series. He decided to make a pilgrimage to Stockholm to see Mittag-Leffler, who was the recognized lord of complex analysis. Mittag-Leffler listened politely to what Borel had to say and then, placing his hand upon the complete works by Weierstrass, his teacher, he said in Latin, "The Master forbids it." A tale of Mark Kac," quoted (on page 38) by

Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics IV: Analysis of Operators*, volume 4 of *Methods of Modern Mathematical Physics Volume*. Academic Press, New York, 1978. ISBN 0125850042,9780125850049. URL https://www.elsevier.com/books/iv-analysis-of-operators/reed/978-0-08-057045-7

A similar calculation for s^2 defined in Eq. (12.7) yields

$$s^2 = \sum_{j=0}^{\infty} (-1)^{j+1} j = (-1) \sum_{j=1}^{\infty} (-1)^j j \overset{B}{=} -\int_0^{\infty} \left(\sum_{j=1}^{\infty} \frac{(-1)^j j t^j}{j!} \right) e^{-t} dt$$

$$= -\int_0^{\infty} \left(\sum_{j=1}^{\infty} \frac{(-t)^j}{(j-1)!} \right) e^{-t} dt$$

$$= -\int_0^{\infty} \left(\sum_{j=0}^{\infty} \frac{(-t)^{j+1}}{j!} \right) e^{-t} dt$$

$$= -\int_0^{\infty} (-t) \left(\sum_{j=0}^{\infty} \frac{(-t)^j}{j!} \right) e^{-t} dt \quad (12.40)$$

$$= -\int_0^{\infty} (-t) e^{-2t} dt$$

[variable substitution $2t = \zeta, dt = \frac{1}{2} d\zeta$]

$$= \frac{1}{4} \int_0^{\infty} \zeta e^{-\zeta} d\zeta$$

$$= \frac{1}{4} \Gamma(2) = \frac{1}{4} 1! = \frac{1}{4},$$

which is again equal to the Abel sum.

Bibliography

Milton Abramowitz and Irene A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Number 55 in National Bureau of Standards Applied Mathematics Series. U.S. Government Printing Office, Washington, D.C., 1964. URL http://www.math.sfu.ca/~cbm/aands/.

Lars V. Ahlfors. *Complex Analysis: An Introduction of the Theory of Analytic Functions of One Complex Variable.* McGraw-Hill Book Co., New York, third edition, 1978.

M. A. Al-Gwaiz. *Sturm-Liouville Theory and its Applications.* Springer, London, 2008.

A. D. Alexandrov. On Lorentz transformations. *Uspehi Mat. Nauk.*, 5(3): 187, 1950.

A. D. Alexandrov. A contribution to chronogeometry. *Canadian Journal of Math.*, 19:1119–1128, 1967.

A. D. Alexandrov. Mappings of spaces with families of cones and spacetime transformations. *Annali di Matematica Pura ed Applicata*, 103: 229–257, 1975. ISSN 0373-3114. DOI: 10.1007/BF02414157. URL http://doi.org/10.1007/BF02414157.

A. D. Alexandrov. On the principles of relativity theory. In *Classics of Soviet Mathematics. Volume 4. A. D. Alexandrov. Selected Works*, pages 289–318. 1996.

Philip W. Anderson. More is different. Broken symmetry and the nature of the hierarchical structure of science. *Science*, 177(4047):393–396, August 1972. DOI: 10.1126/science.177.4047.393. URL http://doi.org/10.1126/science.177.4047.393.

George E. Andrews, Richard Askey, and Ranjan Roy. *Special Functions*, volume 71 of *Encyclopedia of Mathematics and its Applications.* Cambridge University Press, Cambridge, 1999. ISBN 0-521-62321-9.

Tom M. Apostol. *Mathematical Analysis: A Modern Approach to Advanced Calculus.* Addison-Wesley Series in Mathematics. Addison-Wesley, Reading, MA, second edition, 1974. ISBN 0-201-00288-4.

Thomas Aquinas. *Summa Theologica. Translated by Fathers of the English Dominican Province.* Christian Classics Ethereal Library, Grand Rapids, MI, 1981. URL http://www.ccel.org/ccel/aquinas/summa.html.

George B. Arfken and Hans J. Weber. *Mathematical Methods for Physicists.* Elsevier, Oxford, sixth edition, 2005. ISBN 0-12-059876-0;0-12-088584-0.

Sheldon Axler, Paul Bourdon, and Wade Ramey. *Harmonic Function Theory*, volume 137 of *Graduate texts in mathematics.* second edition, 1994. ISBN 0-387-97875-5.

L. E. Ballentine. *Quantum Mechanics.* Prentice Hall, Englewood Cliffs, NJ, 1989.

Asim O. Barut. $e = -\hbar\omega$. *Physics Letters A*, 143(8):349–352, 1990. ISSN 0375-9601. DOI: 10.1016/0375-9601(90)90369-Y. URL http://doi.org/10.1016/0375-9601(90)90369-Y.

W. W. Bell. *Special Functions for Scientists and Engineers.* D. Van Nostrand Company Ltd, London, 1968.

Paul Benacerraf. Tasks and supertasks, and the modern Eleatics. *Journal of Philosophy*, LIX(24):765–784, 1962. URL http://www.jstor.org/stable/2023500.

Walter Benz. *Geometrische Transformationen.* BI Wissenschaftsverlag, Mannheim, 1992.

Michael Berry. Asymptotics, superasymptotics, hyperasymptotics... In Harvey Segur, Saleh Tanveer, and Herbert Levine, editors, *Asymptotics beyond All Orders*, volume 284 of *NATO ASI Series*, pages 1–14. Springer, 1992. ISBN 978-1-4757-0437-2. DOI: 10.1007/978-1-4757-0435-8. URL http://doi.org/10.1007/978-1-4757-0435-8.

Garrett Birkhoff and Gian-Carlo Rota. *Ordinary Differential Equations.* John Wiley & Sons, New York, Chichester, Brisbane, Toronto, fourth edition, 1959, 1960, 1962, 1969, 1978, and 1989.

Garrett Birkhoff and John von Neumann. The logic of quantum mechanics. *Annals of Mathematics*, 37(4):823–843, 1936. DOI: 10.2307/1968621. URL http://doi.org/10.2307/1968621.

Guy Bonneau, Jacques Faraut, and Galliano Valent. Self-adjoint extensions of operators and the teaching of quantum mechanics. *American Journal of Physics*, 69(3):322–331, 2001. DOI: 10.1119/1.1328351. URL http://doi.org/10.1119/1.1328351.

H. J. Borchers and G. C. Hegerfeldt. The structure of space-time transformations. *Communications in Mathematical Physics*, 28(3):259–266, 1972. URL http://projecteuclid.org/euclid.cmp/1103858408.

Émile Borel. Mémoire sur les séries divergentes. *Annales scientifiques de l'École Normale Supérieure*, 16:9–131, 1899. URL http://eudml.org/doc/81143.

John P. Boyd. The devil's invention: Asymptotic, superasymptotic and hyperasymptotic series. *Acta Applicandae Mathematica*, 56:1–98, 1999. ISSN 0167-8019. DOI: 10.1023/A:1006145903624. URL http://doi.org/10.1023/A:1006145903624.

Percy W. Bridgman. A physicist's second reaction to Mengenlehre. *Scripta Mathematica*, 2:101–117, 224–234, 1934.

Yuri Alexandrovich Brychkov and Anatolii Platonovich Prudnikov. *Handbook of special functions: derivatives, integrals, series and other formulas*. CRC/Chapman & Hall Press, Boca Raton, London, New York, 2008.

B.L. Burrows and D.J. Colwell. The Fourier transform of the unit step function. *International Journal of Mathematical Education in Science and Technology*, 21(4):629–635, 1990. DOI: 10.1080/0020739900210418. URL http://doi.org/10.1080/0020739900210418.

Adán Cabello. Kochen-Specker theorem and experimental test on hidden variables. *International Journal of Modern Physics A*, 15(18):2813–2820, 2000. DOI: 10.1142/S0217751X00002020. URL http://doi.org/10.1142/S0217751X00002020.

Adán Cabello, José M. Estebaranz, and G. García-Alcaine. Bell-Kochen-Specker theorem: A proof with 18 vectors. *Physics Letters A*, 212(4): 183–187, 1996. DOI: 10.1016/0375-9601(96)00134-X. URL http://doi.org/10.1016/0375-9601(96)00134-X.

J. B. Conway. *Functions of Complex Variables. Volume I*. Springer, New York, NY, 1973.

Rene Descartes. *Discours de la méthode pour bien conduire sa raison et chercher la verité dans les sciences (Discourse on the Method of Rightly Conducting One's Reason and of Seeking Truth)*. 1637. URL http://www.gutenberg.org/etext/59.

Rene Descartes. *The Philosophical Writings of Descartes. Volume 1*. Cambridge University Press, Cambridge, 1985. translated by John Cottingham, Robert Stoothoff and Dugald Murdoch.

Hermann Diels. *Die Fragmente der Vorsokratiker, griechisch und deutsch*. Weidmannsche Buchhandlung, Berlin, 1906. URL http://www.archive.org/details/diefragmentederv01dieluoft.

Paul Adrien Maurice Dirac. *The Principles of Quantum Mechanics*. Oxford University Press, Oxford, fourth edition, 1930, 1958. ISBN 9780198520115.

Hans Jörg Dirschmid. *Tensoren und Felder*. Springer, Vienna, 1996.

Dean G. Duffy. *Green's Functions with Applications*. Chapman and Hall/CRC, Boca Raton, 2001.

Thomas Durt, Berthold-Georg Englert, Ingemar Bengtsson, and Karol Zyczkowski. On mutually unbiased bases. *International Journal of Quantum Information*, 8:535–640, 2010. DOI: 10.1142/S0219749910006502. URL http://doi.org/10.1142/S0219749910006502.

Anatolij Dvurečenskij. *Gleason's Theorem and Its Applications*, volume 60 of *Mathematics and its Applications*. Kluwer Academic Publishers, Springer, Dordrecht, 1993. ISBN 9048142091,978-90-481-4209-5,978-94-015-8222-3. DOI: 10.1007/978-94-015-8222-3. URL http://doi.org/10.1007/978-94-015-8222-3.

Freeman J. Dyson. Divergence of perturbation theory in quantum electrodynamics. *Phys. Rev.*, 85(4):631–632, Feb 1952. DOI: 10.1103/PhysRev.85.631. URL http://doi.org/10.1103/PhysRev.85.631.

Artur Ekert and Peter L. Knight. Entangled quantum systems and the Schmidt decomposition. *American Journal of Physics*, 63(5):415–423, 1995. DOI: 10.1119/1.17904. URL http://doi.org/10.1119/1.17904.

Lawrence C. Evans. *Partial differential equations. Graduate Studies in Mathematics, volume 19*. American Mathematical Society, Providence, Rhode Island, 1998.

Graham Everest, Alf van der Poorten, Igor Shparlinski, and Thomas Ward. *Recurrence sequences. Volume 104 in the AMS Surveys and Monographs series*. American mathematical Society, Providence, RI, 2003.

Hugh Everett III. The Everett interpretation of quantum mechanics: Collected works 1955-1980 with commentary. Princeton University Press, Princeton, NJ, 2012. ISBN 9780691145075. URL http://press.princeton.edu/titles/9770.html.

William Norrie Everitt. A catalogue of Sturm-Liouville differential equations. In Werner O. Amrein, Andreas M. Hinz, and David B. Pearson, editors, *Sturm-Liouville Theory, Past and Present*, pages 271–331. Birkhäuser Verlag, Basel, 2005. URL http://www.math.niu.edu/SL2/papers/birk0.pdf.

Richard Phillips Feynman. *The Feynman lectures on computation*. Addison-Wesley Publishing Company, Reading, MA, 1996. edited by A.J.G. Hey and R. W. Allen.

Stefan Filipp and Karl Svozil. Generalizing Tsirelson's bound on Bell inequalities using a min-max principle. *Physical Review Letters*, 93: 130407, 2004. DOI: 10.1103/PhysRevLett.93.130407. URL http://doi.org/10.1103/PhysRevLett.93.130407.

Edward Fredkin. Digital mechanics. an informational process based on reversible universal cellular automata. *Physica*, D45:254–270, 1990. DOI: 10.1016/0167-2789(90)90186-S. URL http://doi.org/10.1016/0167-2789(90)90186-S.

Eberhard Freitag and Rolf Busam. *Funktionentheorie 1*. Springer, Berlin, Heidelberg, fourth edition, 1993,1995,2000,2006. English translation in [12].

Eberhard Freitag and Rolf Busam. *Complex Analysis*. Springer, Berlin, Heidelberg, 2005.

Sigmund Freud. Ratschläge für den Arzt bei der psychoanalytischen Behandlung. In Anna Freud, E. Bibring, W. Hoffer, E. Kris, and O. Isakower, editors, *Gesammelte Werke. Chronologisch geordnet. Achter Band. Werke aus den Jahren 1909–1913*, pages 376–387. Fischer, Frankfurt am Main, 1912, 1999. URL http://gutenberg.spiegel.de/buch/kleine-schriften-ii-7122/15.

Theodore W. Gamelin. *Complex Analysis*. Springer, New York, NY, 2001.

Elisabeth Garber, Stephen G. Brush, and C. W. Francis Everitt. *Maxwell on Heat and Statistical Mechanics: On "Avoiding All Personal Enquiries" of Molecules*. Associated University Press, Cranbury, NJ, 1995. ISBN 0934223343.

I. M. Gel'fand and G. E. Shilov. *Generalized Functions. Vol. 1: Properties and Operations*. Academic Press, New York, 1964. Translated from the Russian by Eugene Saletan.

François Gieres. Mathematical surprises and Dirac's formalism in quantum mechanics. *Reports on Progress in Physics*, 63(12):1893–1931, 2000. DOI: http://doi.org/10.1088/0034-4885/63/12/201. URL 10.1088/0034-4885/63/12/201.

Andrew M. Gleason. Measures on the closed subspaces of a Hilbert space. *Journal of Mathematics and Mechanics (now Indiana University Mathematics Journal)*, 6(4):885–893, 1957. ISSN 0022-2518. DOI: 10.1512/iumj.1957.6.56050". URL http://doi.org/10.1512/iumj.1957.6.56050.

Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. November 2016. ISBN 9780262035613, 9780262337434. URL https://mitpress.mit.edu/books/deep-learning.

I. S. Gradshteyn and I. M. Ryzhik. *Tables of Integrals, Series, and Products, 6th ed.* Academic Press, San Diego, CA, 2000.

Dietrich Grau. *Übungsaufgaben zur Quantentheorie*. Karl Thiemig, Karl Hanser, München, 1975, 1993, 2005. URL http://www.dietrich-grau.at.

Richard J. Greechie. Orthomodular lattices admitting no states. *Journal of Combinatorial Theory. Series A*, 10:119–132, 1971. DOI: 10.1016/0097-3165(71)90015-X. URL http://doi.org/10.1016/0097-3165(71)90015-X.

Robert E. Greene and Stephen G. Krantz. *Function theory of one complex variable*, volume 40 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, Rhode Island, third edition, 2006.

[12] Eberhard Freitag and Rolf Busam. *Complex Analysis*. Springer, Berlin, Heidelberg, 2005

Werner Greub. *Linear Algebra*, volume 23 of *Graduate Texts in Mathematics*. Springer, New York, Heidelberg, fourth edition, 1975.

A. Grünbaum. *Modern Science and Zeno's paradoxes*. Allen and Unwin, London, second edition, 1968.

Paul Richard Halmos. *Finite-dimensional Vector Spaces*. Undergraduate Texts in Mathematics. Springer, New York, 1958. ISBN 978-1-4612-6387-6,978-0-387-90093-3. DOI: 10.1007/978-1-4612-6387-6. URL http://doi.org/10.1007/978-1-4612-6387-6.

Jan Hamhalter. *Quantum Measure Theory. Fundamental Theories of Physics, Vol. 134*. Kluwer Academic Publishers, Dordrecht, Boston, London, 2003. ISBN 1-4020-1714-6.

Godfrey Harold Hardy. *Divergent Series*. Oxford University Press, 1949.

Hans Havlicek. *Lineare Algebra für Technische Mathematiker*. Heldermann Verlag, Lemgo, second edition, 2008.

Hans Havlicek. private communication, 2016.

Oliver Heaviside. *Electromagnetic theory*. "The Electrician" Printing and Publishing Corporation, London, 1894-1912. URL http://archive.org/details/electromagnetict02heavrich.

Jim Hefferon. Linear algebra. 320-375, 2011. URL http://joshua.smcvt.edu/linalg.html/book.pdf.

Russell Herman. *A Second Course in Ordinary Differential Equations: Dynamical Systems and Boundary Value Problems*. University of North Carolina Wilmington, Wilmington, NC, 2008. URL http://people.uncw.edu/hermanr/mat463/ODEBook/Book/ODE_LargeFont.pdf. Creative Commons Attribution-NoncommercialShare Alike 3.0 United States License.

Russell Herman. *Introduction to Fourier and Complex Analysis with Applications to the Spectral Analysis of Signals*. University of North Carolina Wilmington, Wilmington, NC, 2010. URL http://people.uncw.edu/hermanr/mat367/FCABook/Book2010/FTCA-book.pdf. Creative Commons Attribution-NoncommercialShare Alike 3.0 United States License.

Einar Hille. *Analytic Function Theory*. Ginn, New York, 1962. 2 Volumes.

Einar Hille. *Lectures on ordinary differential equations*. Addison-Wesley, Reading, Mass., 1969.

Edmund Hlawka. Zum Zahlbegriff. *Philosophia Naturalis*, 19:413–470, 1982.

Howard Homes and Chris Rorres. *Elementary Linear Algebra: Applications Version*. Wiley, New York, tenth edition, 2010.

Kenneth B. Howell. *Principles of Fourier analysis*. Chapman & Hall/CRC, Boca Raton, London, New York, Washington, D.C., 2001.

Klaus Jänich. *Analysis für Physiker und Ingenieure. Funktionentheorie, Differentialgleichungen, Spezielle Funktionen.* Springer, Berlin, Heidelberg, fourth edition, 2001. URL http://www.springer.com/mathematics/analysis/book/978-3-540-41985-3.

Klaus Jänich. *Funktionentheorie. Eine Einführung.* Springer, Berlin, Heidelberg, sixth edition, 2008. DOI: 10.1007/978-3-540-35015-6. URL 10.1007/978-3-540-35015-6.

Edwin Thompson Jaynes. Clearing up mysteries - the original goal. In John Skilling, editor, *Maximum-Entropy and Bayesian Methods: Proceedings of the 8th Maximum Entropy Workshop, held on August 1-5, 1988, in St. John's College, Cambridge, England*, pages 1–28. Kluwer, Dordrecht, 1989. URL http://bayes.wustl.edu/etj/articles/cmystery.pdf.

Edwin Thompson Jaynes. Probability in quantum theory. In Wojciech Hubert Zurek, editor, *Complexity, Entropy, and the Physics of Information: Proceedings of the 1988 Workshop on Complexity, Entropy, and the Physics of Information, held May - June, 1989, in Santa Fe, New Mexico*, pages 381–404. Addison-Wesley, Reading, MA, 1990. ISBN 9780201515091. URL http://bayes.wustl.edu/etj/articles/prob.in.qm.pdf.

Ulrich D. Jentschura. Resummation of nonalternating divergent perturbative expansions. *Physical Review D*, 62:076001, Aug 2000. DOI: 10.1103/PhysRevD.62.076001. URL http://doi.org/10.1103/PhysRevD.62.076001.

Satish D. Joglekar. *Mathematical Physics: The Basics.* CRC Press, Boca Raton, Florida, 2007.

Ernst Jünger. *Heliopolis. Rückblick auf eine Stadt.* Heliopolis-Verlag Ewald Katzmann, Tübingen, 1949.

Vladimir Kisil. Special functions and their symmetries. Part II: Algebraic and symmetry methods. Postgraduate Course in Applied Analysis, May 2003. URL http://www1.maths.leeds.ac.uk/~kisilv/courses/sp-repr.pdf.

Hagen Kleinert and Verena Schulte-Frohlinde. *Critical Properties of ϕ^4-Theories.* World scientific, Singapore, 2001. ISBN 9810246595.

Morris Kline. Euler and infinite series. *Mathematics Magazine*, 56(5): 307–314, 1983. ISSN 0025570X. DOI: 10.2307/2690371. URL http://doi.org/10.2307/2690371.

Ebergard Klingbeil. *Tensorrechnung für Ingenieure.* Bibliographisches Institut, Mannheim, 1966.

Simon Kochen and Ernst P. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics (now Indiana University Mathematics Journal)*, 17(1):59–87, 1967. ISSN 0022-2518. DOI: 10.1512/iumj.1968.17.17004. URL http://doi.org/10.1512/iumj.1968.17.17004.

T. W. Körner. *Fourier Analysis*. Cambridge University Press, Cambridge, UK, 1988.

Gerhard Kristensson. Equations of Fuchsian type. In *Second Order Differential Equations*, pages 29–42. Springer, New York, 2010. ISBN 978-1-4419-7019-0. DOI: 10.1007/978-1-4419-7020-6. URL http://doi.org/10.1007/978-1-4419-7020-6.

Dietrich Küchemann. *The Aerodynamic Design of Aircraft*. Pergamon Press, Oxford, 1978.

Vadim Kuznetsov. Special functions and their symmetries. Part I: Algebraic and analytic methods. Postgraduate Course in Applied Analysis, May 2003. URL http://www1.maths.leeds.ac.uk/~kisilv/courses/sp-funct.pdf.

Imre Lakatos. *Philosophical Papers. 1. The Methodology of Scientific Research Programmes*. Cambridge University Press, Cambridge, 1978. ISBN 9781316038765, 9780521280310, 0521280311.

Rolf Landauer. Information is physical. *Physics Today*, 44(5):23–29, May 1991. DOI: 10.1063/1.881299. URL http://doi.org/10.1063/1.881299.

Ron Larson and Bruce H. Edwards. *Calculus*. Brooks/Cole Cengage Learning, Belmont, CA, nineth edition, 2010. ISBN 978-0-547-16702-2.

N. N. Lebedev. *Special Functions and Their Applications*. Prentice-Hall Inc., Englewood Cliffs, N.J., 1965. R. A. Silverman, translator and editor; reprinted by Dover, New York, 1972.

H. D. P. Lee. *Zeno of Elea*. Cambridge University Press, Cambridge, 1936.

Gottfried Wilhelm Leibniz. Letters LXX, LXXI. In Carl Immanuel Gerhardt, editor, *Briefwechsel zwischen Leibniz und Christian Wolf. Handschriften der Königlichen Bibliothek zu Hannover,*. H. W. Schmidt, Halle, 1860. URL http://books.google.de/books?id=TUkJAAAAQAAJ.

Steven J. Leon, Åke Björck, and Walter Gander. Gram-Schmidt orthogonalization: 100 years and more. *Numerical Linear Algebra with Applications*, 20(3):492–532, 2013. ISSN 1070-5325. DOI: 10.1002/nla.1839. URL http://doi.org/10.1002/nla.1839.

June A. Lester. Distance preserving transformations. In Francis Buekenhout, editor, *Handbook of Incidence Geometry*, pages 921–944. Elsevier, Amsterdam, 1995.

M. J. Lighthill. *Introduction to Fourier Analysis and Generalized Functions*. Cambridge University Press, Cambridge, 1958.

Ismo V. Lindell. Delta function expansions, complex delta functions and the steepest descent method. *American Journal of Physics*, 61(5):438–442, 1993. DOI: 10.1119/1.17238. URL http://doi.org/10.1119/1.17238.

Seymour Lipschutz and Marc Lipson. *Linear algebra.* Schaum's outline series. McGraw-Hill, fourth edition, 2009.

George Mackiw. A note on the equality of the column and row rank of a matrix. *Mathematics Magazine*, 68(4):pp. 285–286, 1995. ISSN 0025570X. URL http://www.jstor.org/stable/2690576.

T. M. MacRobert. *Spherical Harmonics. An Elementary Treatise on Harmonic Functions with Applications*, volume 98 of *International Series of Monographs in Pure and Applied Mathematics.* Pergamon Press, Oxford, third edition, 1967.

Eli Maor. *Trigonometric Delights.* Princeton University Press, Princeton, 1998. URL http://press.princeton.edu/books/maor/.

Francisco Marcellán and Walter Van Assche. *Orthogonal Polynomials and Special Functions*, volume 1883 of *Lecture Notes in Mathematics.* Springer, Berlin, 2006. ISBN 3-540-31062-2.

David N. Mermin. Lecture notes on quantum computation. accessed on Jan 2nd, 2017, 2002-2008. URL http://www.lassp.cornell.edu/mermin/qcomp/CS483.html.

David N. Mermin. *Quantum Computer Science.* Cambridge University Press, Cambridge, 2007. ISBN 9780521876582. URL http://people.ccmr.cornell.edu/~mermin/qcomp/CS483.html.

A. Messiah. *Quantum Mechanics*, volume I. North-Holland, Amsterdam, 1962.

Charles N. Moore. *Summable Series and Convergence Factors.* American Mathematical Society, New York, NY, 1938.

Walter Moore. *Schrödinger: Life and Thought.* Cambridge University Press, Cambridge, UK, 1989.

Francis D. Murnaghan. *The Unitary and Rotation Groups*, volume 3 of *Lectures on Applied Mathematics.* Spartan Books, Washington, D.C., 1962.

Michael A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2000.

Michael A. Nielsen and I. L. Chuang. *Quantum Computation and Quantum Information.* Cambridge University Press, Cambridge, 2010. 10th Anniversary Edition.

Carl M. Bender Steven A. Orszag. *Andvanced Mathematical Methods for Scientists and Enineers.* McGraw-Hill, New York, NY, 1978.

Asher Peres. Defining length. *Nature*, 312:10, 1984. DOI: 10.1038/312010b0. URL http://doi.org/10.1038/312010b0.

Asher Peres. *Quantum Theory: Concepts and Methods.* Kluwer Academic Publishers, Dordrecht, 1993.

Sergio A. Pernice and Gerardo Oleaga. Divergence of perturbation theory: Steps towards a convergent series. *Physical Review D*, 57:1144–1158, Jan 1998. DOI: 10.1103/PhysRevD.57.1144. URL http://doi.org/10.1103/PhysRevD.57.1144.

Itamar Pitowsky. Infinite and finite Gleason's theorems and the logic of indeterminacy. *Journal of Mathematical Physics*, 39(1):218–228, 1998. DOI: 10.1063/1.532334. URL http://doi.org/10.1063/1.532334.

Michael Reck and Anton Zeilinger. Quantum phase tracing of correlated photons in optical multiports. In F. De Martini, G. Denardo, and Anton Zeilinger, editors, *Quantum Interferometry*, pages 170–177, Singapore, 1994. World Scientific.

Michael Reck, Anton Zeilinger, Herbert J. Bernstein, and Philip Bertani. Experimental realization of any discrete unitary operator. *Physical Review Letters*, 73:58–61, 1994. DOI: 10.1103/PhysRevLett.73.58. URL http://doi.org/10.1103/PhysRevLett.73.58.

Michael Reed and Barry Simon. *Methods of Mathematical Physics I: Functional Analysis.* Academic Press, New York, 1972.

Michael Reed and Barry Simon. *Methods of Mathematical Physics II: Fourier Analysis, Self-Adjointness.* Academic Press, New York, 1975.

Michael Reed and Barry Simon. *Methods of Modern Mathematical Physics IV: Analysis of Operators*, volume 4 of *Methods of Modern Mathematical Physics Volume.* Academic Press, New York, 1978. ISBN 0125850042,9780125850049. URL https://www.elsevier.com/books/iv-analysis-of-operators/reed/978-0-08-057045-7.

Fred Richman and Douglas Bridges. A constructive proof of Gleason's theorem. *Journal of Functional Analysis*, 162:287–312, 1999. DOI: 10.1006/jfan.1998.3372. URL http://doi.org/10.1006/jfan.1998.3372.

Joseph J. Rotman. *An Introduction to the Theory of Groups*, volume 148 of *Graduate texts in mathematics.* Springer, New York, fourth edition, 1995. ISBN 0387942858.

Christiane Rousseau. Divergent series: Past, present, future .... preprint, 2004. URL http://www.dms.umontreal.ca/~rousseac/divergent.pdf.

Richard Mark Sainsbury. *Paradoxes.* Cambridge University Press, Cambridge, United Kingdom, third edition, 2009. ISBN 0521720796.

Dietmar A. Salamon. *Funktionentheorie.* Birkhäuser, Basel, 2012. DOI: 10.1007/978-3-0348-0169-0. URL http://doi.org/10.1007/978-3-0348-0169-0. see also URL http://www.math.ethz.ch/ salamon/PREPRINTS/cxana.pdf.

Leonard I. Schiff. *Quantum Mechanics.* McGraw-Hill, New York, 1955.

Erwin Schrödinger.   Quantisierung als Eigenwertproblem.   *Annalen der Physik*, 384(4):361–376, 1926. ISSN 1521-3889. DOI: 10.1002/andp.19263840404. URL http://doi.org/10.1002/andp.19263840404.

Erwin Schrödinger. Discussion of probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 31(04):555–563, 1935a. DOI: 10.1017/S0305004100013554. URL http://doi.org/10.1017/S0305004100013554.

Erwin Schrödinger. Die gegenwärtige Situation in der Quantenmechanik. *Naturwissenschaften*, 23:807–812, 823–828, 844–849, 1935b. DOI: 10.1007/BF01491891, 10.1007/BF01491914, 10.1007/BF01491987. URL http://doi.org/10.1007/BF01491891,http://doi.org/10.1007/BF01491914,http://doi.org/10.1007/BF01491987.

Erwin Schrödinger. Probability relations between separated systems. *Mathematical Proceedings of the Cambridge Philosophical Society*, 32 (03):446–452, 1936. DOI: 10.1017/S0305004100019137. URL http://doi.org/10.1017/S0305004100019137.

Erwin Schrödinger. *Nature and the Greeks.* Cambridge University Press, Cambridge, 1954, 2014. ISBN 9781107431836. URL http://www.cambridge.org/9781107431836.

Laurent Schwartz. *Introduction to the Theory of Distributions.* University of Toronto Press, Toronto, 1952. collected and written by Israel Halperin.

Julian Schwinger.   Unitary operators bases.   *Proceedings of the National Academy of Sciences (PNAS)*, 46:570–579, 1960. DOI: 10.1073/pnas.46.4.570. URL http://doi.org/10.1073/pnas.46.4.570.

R. Sherr, K. T. Bainbridge, and H. H. Anderson. Transmutation of mercury by fast neutrons. *Physical Review*, 60(7):473–479, Oct 1941. DOI: 10.1103/PhysRev.60.473. URL http://doi.org/10.1103/PhysRev.60.473.

Ernst Snapper and Robert J. Troyer. *Metric Affine Geometry.* Academic Press, New York, 1971.

Thomas Sommer.   Verallgemeinerte Funktionen.   unpublished manuscript, 2012.

Ernst Specker.   Die Logik nicht gleichzeitig entscheidbarer Aussagen. *Dialectica*, 14(2-3):239–246, 1960. DOI: 10.1111/j.1746-8361.1960.tb00422.x. URL http://doi.org/10.1111/j.1746-8361.1960.tb00422.x.

Gilbert Strang. *Introduction to linear algebra.* Wellesley-Cambridge Press, Wellesley, MA, USA, fourth edition, 2009. ISBN 0-9802327-1-6. URL http://math.mit.edu/linearalgebra/.

Robert Strichartz. *A Guide to Distribution Theory and Fourier Transforms.* CRC Press, Boca Roton, Florida, USA, 1994. ISBN 0849382734.

Karl Svozil. Conventions in relativity theory and quantum mechanics. *Foundations of Physics*, 32:479–502, 2002. DOI: 10.1023/A:1015017831247. URL http://doi.org/10.1023/A:1015017831247.

Karl Svozil. Computational universes. *Chaos, Solitons & Fractals*, 25 (4):845–859, 2006. DOI: 10.1016/j.chaos.2004.11.055. URL http://doi.org/10.1016/j.chaos.2004.11.055.

Karl Svozil. Versklavung durch Verbuchhalterung. *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare*, 60(1): 101–111, 2013. URL http://eprints.rclis.org/19560/.

Frederick Winslow Taylor. *The Principles of Scientific Management.* Harper Bros., New York, NY, 1911. URL https://archive.org/details/principlesofscie00taylrich.

Nico M. Temme. *Special functions: an introduction to the classical functions of mathematical physics.* John Wiley & Sons, Inc., New York, 1996. ISBN 0-471-11313-1.

Nico M. Temme. Numerical aspects of special functions. *Acta Numerica*, 16:379–478, 2007. ISSN 0962-4929. DOI: 10.1017/S0962492904000077. URL http://doi.org/10.1017/S0962492904000077.

Gerald Teschl. *Ordinary Differential Equations and Dynamical Systems. Graduate Studies in Mathematics, volume 140.* American Mathematical Society, Providence, Rhode Island, 2012. ISBN ISBN-10: 0-8218-8328-3 / ISBN-13: 978-0-8218-8328-0. URL http://www.mat.univie.ac.at/~gerald/ftp/book-ode/ode.pdf.

Tommaso Toffoli. The role of the observer in uniform systems. In George J. Klir, editor, *Applied General Systems Research: Recent Developments and Trends*, pages 395–400. Plenum Press, Springer US, New York, London, and Boston, MA, 1978. ISBN 978-1-4757-0555-3. DOI: 10.1007/978-1-4757-0555-3_29. URL http://doi.org/10.1007/978-1-4757-0555-3_29.

William F. Trench. Introduction to real analysis. Free Hyperlinked Edition 2.01, 2012. URL http://ramanujan.math.trinity.edu/wtrench/texts/TRENCH_REAL_ANALYSIS.PDF.

John von Neumann. Über Funktionen von Funktionaloperatoren. *Annalen der Mathematik (Annals of Mathematics)*, 32, 04 1931. DOI: 10.2307/1968185. URL http://doi.org/10.2307/1968185.

John von Neumann. *Mathematische Grundlagen der Quantenmechanik.* Springer, Berlin, Heidelberg, second edition, 1932, 1996. ISBN 978-3-642-61409-5,978-3-540-59207-5,978-3-642-64828-1. DOI: 10.1007/978-3-642-61409-5. URL http://doi.org/10.1007/978-3-642-61409-5. English translation in Ref. [13].

[13] John von Neumann. *Mathematical Foundations of Quantum Mechanics.* Princeton University Press, Princeton, NJ, 1955. ISBN 9780691028934. URL http://press.princeton.edu/titles/2113.html. German original in Ref.

John von Neumann. *Mathematische Grundlagen der Quantenmechanik.* Springer, Berlin, Heidelberg, second

John von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, NJ, 1955. ISBN 9780691028934. URL http://press.princeton.edu/titles/2113.html. German original in Ref. [14].

Gabriel Weinreich. *Geometrical Vectors (Chicago Lectures in Physics)*. The University of Chicago Press, Chicago, IL, 1998.

David Wells. Which is the most beautiful? *The Mathematical Intelligencer*, 10:30–31, 1988. ISSN 0343-6993. DOI: 10.1007/BF03023741. URL http://doi.org/10.1007/BF03023741.

E. T. Whittaker and G. N. Watson. *A Course of Modern Analysis*. Cambridge University Press, Cambridge, fourth edition, 1927. URL http://archive.org/details/ACourseOfModernAnalysis. Reprinted in 1996. Table errata: Math. Comp. v. 36 (1981), no. 153, p. 319.

Eugene P. Wigner. The unreasonable effectiveness of mathematics in the natural sciences. Richard Courant Lecture delivered at New York University, May 11, 1959. *Communications on Pure and Applied Mathematics*, 13:1–14, 1960. DOI: 10.1002/cpa.3160130102. URL http://doi.org/10.1002/cpa.3160130102.

Herbert S. Wilf. *Mathematics for the physical sciences*. Dover, New York, 1962. URL http://www.math.upenn.edu/~wilf/website/Mathematics_for_the_Physical_Sciences.html.

William K. Wootters and B. D. Fields. Optimal state-determination by mutually unbiased measurements. *Annals of Physics*, 191:363–381, 1989. DOI: 10.1016/0003-4916(89)90322-9. URL http://doi.org/10.1016/0003-4916(89)90322-9.

Anton Zeilinger. A foundational principle for quantum mechanics. *Foundations of Physics*, 29(4):631–643, 1999. DOI: 10.1023/A:1018820410908. URL http://doi.org/10.1023/A:1018820410908.

Konrad Zuse. *Rechnender Raum*. Friedrich Vieweg & Sohn, Braunschweig, 1969.

Konrad Zuse. Discrete mathematics and Rechnender Raum. 1994. URL http://www.zib.de/PaperWeb/abstracts/TR-94-10/.

[14] John von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer, Berlin, Heidelberg, second edition, 1932, 1996. ISBN 978-3-642-61409-5,978-3-540-59207-5,978-3-642-64828-1. DOI: 10.1007/978-3-642-61409-5. URL http://doi.org/10.1007/978-3-642-61409-5. English translation in Ref.

John von Neumann. *Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, NJ, 1955. ISBN 9780691028934. URL http://press.princeton.edu/titles/2113.html. German original in Ref.

Index