Fields

In the physical sciences, positions in space or in space-time are important. These can be denoted e.g. $\boldsymbol{r}$, or $(\boldsymbol{r}, t)$. In mathematical physics, a field is a differentiable function of such positions.

Let $\mathbb{S}$ be the set of positions. Typical fields are:

- Scalar fields $\mathbb{S} \mapsto \mathbb{R}$, e.g. potential fields, such as the electrostatic potential.
- Scalar fields describing some density, such as the charge density, mass density, or energy density.
- Vector fields $\mathbb{S} \mapsto \mathbb{R}^3$ describing velocities, flux density, electric or magnetic field strength, etc.
- Spinor fields, represented by embedding into $\mathbb{S} \mapsto (\mathbb{C}^2)$ or $\mathbb{S} \mapsto (\mathbb{C}^4)$, which may describe the amplitudes of relativistic wave functions used in some equations,
- Tensor fields $\mathbb{S} \mapsto (\mathbb{R}^4)^N$, used a lot in relativistic theories and in general relativity.

The values of a field is in principle any differentiable vector space. If we stick to scalar and vector fields in a 3-dimensional Euclidean space, then there is a much-used calculus, mostly due to Gibbs, which may be called vector calculus or vector analysis.

Vector multiplication

Let $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ be vectors in $\mathbb{R}^3$. The vector product $\mathbf{C} = \mathbf{A} \times \mathbf{B}$ is computed as

$$C_x = A_y B_z - A_z B_y \qquad C_y = A_z B_x - A_x B_z \qquad C_z = A_x B_y - A_y B_x$$

Some vector product rules are

$$\begin{aligned}
\mathbf{A} \times \mathbf{B} &= -\mathbf{B} \times \mathbf{A} \\
\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) &= \det \begin{pmatrix} A_x & B_x & C_x \\ A_y & B_y & C_y \\ A_z & B_z & C_z \end{pmatrix} \\
\mathbf{A} \cdot (\mathbf{B} \times \mathbf{C}) &= \mathbf{B} \cdot (\mathbf{C} \times \mathbf{A}) = \mathbf{C} \cdot (\mathbf{A} \times \mathbf{B}) = (\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C} = \cdots \\
\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) &= (\mathbf{A} \cdot \mathbf{C})\,\mathbf{B} - (\mathbf{A} \cdot \mathbf{B})\,\mathbf{C} \\
(\text{but } \textbf{note} \text{ that} \quad & \mathbf{A} \times (\mathbf{B} \times \mathbf{C}) \neq (\mathbf{A} \times \mathbf{B}) \times \mathbf{C} \text{ !})
\end{aligned}$$

These (and many other) rules can be found in any textbook or handbook which covers vector algebra, or check wikipedia under 'vector product'.

To work it out yourself, use the cyclic permutation rule $x \to y \to z \to x \ldots$. Also convenient: the scalar triple product is a determinant.

Nabla, or del, operator

Let $u(\mathbf{r})$ and $v(\mathbf{r})$ be two scalar fields, while $\mathbf{a}(\mathbf{r})$ and $\mathbf{b}(\mathbf{r})$ are vector fields.

The vector differential ∇, called *nabla* or *del*, can be written as

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$$

The following first derivatives can be formed:

$$\nabla u = \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}\right) \qquad \nabla \cdot \mathbf{a} = \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z}$$

$$\nabla \times \mathbf{a} = \left(\frac{\partial a_y}{\partial z} - \frac{\partial a_z}{\partial y}, \frac{\partial a_z}{\partial x} - \frac{\partial a_x}{\partial z}, \frac{\partial a_x}{\partial y} - \frac{\partial a_y}{\partial x}\right)$$

Rules:

$$\nabla uv = u\nabla v + v\nabla u \qquad \nabla \cdot u\mathbf{a} = u\nabla \mathbf{a} + \mathbf{a} \cdot \nabla u \qquad \nabla \cdot \mathbf{a} \times \mathbf{b} = \mathbf{b} \cdot \nabla \times \mathbf{a} - \mathbf{a} \cdot \nabla \times \mathbf{b}$$

$$\nabla \cdot (\nabla \times \mathbf{a}) = \mathbf{0} \qquad \nabla \times (\nabla u) = 0 \qquad \nabla \cdot (\nabla u) = \nabla^2 u$$

Integrals in vector calculus

The volume integral is simply a triple integral, where the integration element depends on the coordinate system used, e.g. $d\Omega = dx\,dy\,dz$ (Cartesian) or $d\Omega = r^2 \sin\theta\, dr\, d\theta\, d\phi$ (Spher. polar).

The surface integral is a vector. Assume the surface is defined, at least locally, by coordinates (u, v), i.e. S is the set of points $\mathbf{r}$ with cartesian coordinates (x, y, z) that each depend on (u, v),

$$\begin{cases} x = x(u,v) \\ y = y(u,v) \\ z = z(u,v) \end{cases} \qquad d\mathbf{S} = \left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v} \right) du\, dv$$

so for e.g. a spherical polar system, suppose we integrate over a spherical shell with radius R,

$$d\mathbf{S} = \begin{pmatrix} R\cos\theta\cos\phi \\ R\cos\theta\sin\phi \\ -R\sin\theta \end{pmatrix} \times \begin{pmatrix} R\sin\theta\sin\phi \\ -R\sin\theta\cos\phi \\ 0 \end{pmatrix} d\theta\, d\phi = \begin{pmatrix} \sin\theta\cos\phi \\ \sin\theta\sin\phi \\ \cos\theta \end{pmatrix} R^2 \sin\theta\, d\theta\, d\phi$$

Integration by parts

The formula for the derivative of a product can be rearranged as

$$\frac{df(x)}{dx}g(x) = \frac{d}{dx}(f(x)g(x)) - f(x)\frac{dg(x)}{dx}$$

Integrating both sides gives the formula for partial integration,

$$\begin{aligned}\int_{x=a}^{b} \frac{df(x)}{dx}g(x)\,dx &= f(b)g(b) - f(a)g(a) - \int_{x=a}^{b} f(x)\frac{dg(x)}{dx}\,dx \\ &= [f(x)g(x)]_a^b - \int_{x=a}^{b} f(x)\frac{dg(x)}{dx}\,dx\end{aligned}$$

In multiple integrals, the same can be done with each integral, and results in the various formulae for relating surface/curve or volume/surface integrals (at least for rectangular or cuboid regions),

Relations between integral and differential formulae

$$\iint_{\partial\Omega} \mathbf{a}\cdot d\boldsymbol{\Sigma} = \iiint_{\Omega} \nabla\cdot\mathbf{a} d\Omega \quad \text{(Gauss' theorem)} \quad \text{and} \quad \iint_{\partial\Omega} \mathbf{a}\times d\boldsymbol{\Sigma} = -\iiint_{\Omega} \nabla\times\mathbf{a} d\Omega$$

$$\int_{\partial S} \mathbf{a}\cdot d\mathbf{r} = \iint_{S} \nabla\times\mathbf{a}\cdot d\mathbf{S} \quad \text{(Stokes' theorem)} \quad \text{and} \quad \int_{\partial S} \mathbf{a}\times d\mathbf{r} = -\iint_{S} (d\mathbf{S}\times\nabla)\times\mathbf{a}$$

Green's theorem

$$\iint_{\partial\Omega} \left(u\nabla v - v\nabla u\right) d\boldsymbol{\Sigma} = \iiint_{\Omega} \left(u\nabla^2 v - v\nabla^2 u\right) d\Omega$$

Other theorems of this kind:

$$\iint_{\partial\Omega} u\, d\boldsymbol{\Sigma} = \iiint_{\Omega} \nabla u\, d\Omega$$

$$\int_{\partial S} u\, d\boldsymbol{r} = \int_{S} (d\,\boldsymbol{S}\times\nabla u)$$

Example of Gauss' and Stokes' theorems: Maxwell's equations

$$\iint_{\partial\Omega} \boldsymbol{E}\cdot d\boldsymbol{\Sigma} = \frac{1}{\varepsilon_0}\iiint_{\Omega}\rho d\Omega \quad\Leftrightarrow\quad \nabla\cdot\boldsymbol{E} = \frac{\rho}{\varepsilon_0}$$

$$\iint_{\partial\Omega} \boldsymbol{B}\cdot d\boldsymbol{\Sigma} = 0 \quad\Leftrightarrow\quad \nabla\cdot\boldsymbol{B} = 0$$

$$\int_{C} \boldsymbol{E}\cdot d\boldsymbol{r} = -\int_{S}\frac{\partial\boldsymbol{B}}{\partial t}\cdot d\boldsymbol{S} \quad\Leftrightarrow\quad \nabla\times\boldsymbol{E} = -\frac{\partial\boldsymbol{B}}{\partial t}$$

$$\int_{C} \boldsymbol{B}\cdot d\boldsymbol{r} = \mu_0\int_{S}\left(\boldsymbol{j}+\varepsilon_0\frac{\partial\boldsymbol{E}}{\partial t}\right)\cdot d\boldsymbol{S} \quad\Leftrightarrow\quad \nabla\times\boldsymbol{B} = \mu_0\boldsymbol{j}+\mu_0\varepsilon_0\frac{\partial\boldsymbol{E}}{\partial t}$$

The first equation relates a surface integral of the electric field to the enclosed charge. Using Gauss' theorem, this means that for any choice of Ω,

$$\iiint_{\Omega}\left(\nabla\cdot\boldsymbol{E}-\frac{1}{\varepsilon_0}\rho\right)d\Omega = 0$$

which implies the statement on the right.

Harmonic functions in 3D

Harmonic functions are used in all areas of applied mathematics; this are functions $\psi(\mathbf{r})$ for which $\nabla^2\psi(\mathbf{r} = 0$ in some region of space. Together, they form a linear space.

- In simply connected, compact regions, this space is spanned by a set of polynomials in the Cartesian coordinates.
- The standard choice of these polynomials is as the Solid Harmonics,

$$\mathcal{Y}_l^m \stackrel{\text{def}}{=} r^l Y_l^m(\theta, \phi)$$

where Y_l^m is a standard spherical harmonic, $l \in \{0, 1, 2, \dots\}$ and $m \in \{l, l-1, l-2, \dots, -l\}$, and (r, θ, ϕ) are spherical polar coordinates.

- These are in fact polynomials, with total degree l.
- There are other systems of harmonic functions, usually tied to special boundary conditions or symmetry demands.
- In charge-free regions of space, the electrostatic potential is a harmonic function.

Using gradients for optimization

In optimization problems, one tries to find maximum or minimum of a scalar function Q by varying a number of parameters – potentially an infinite number, but let's assume a finite number N.

If parameters can be varied freely, we look for stationary 'points' in parameter space: Treat parameters as a vector $\mathbf{p} = (p_1, \ldots, p_N)$, and try to find $\mathbf{p}$ such that $\nabla Q(\mathbf{p}) = \mathbf{0}$.

But parameters cannot usually be varied freely. There are side relations, and the simplest are like $R(\mathbf{p}) = 0$. The restriction $Q(\mathbf{p})|_{R(\mathbf{p})=0}$ is stationary when $\nabla Q(\mathbf{p})$ has no component along the (hyper-)surface $R(\mathbf{p}) = 0$. This means that $\nabla Q(\mathbf{p})$ is orthogonal to that surface.

But at every point, $\nabla R(\mathbf{p})$ is orthogonal to the surface. The criterion is thus that these two gradients are parallel:

$$\nabla Q(\mathbf{p}) = \lambda \nabla R(\mathbf{p})$$

for some unknown Lagrange multiplier λ.

Lagrange multipliers

There are the usual caveats: Q should be differentiable, $R(\mathbf{p}) = 0$ should define a differentiable manifold, and so on. Leaving the validity of this approach to be determined in individual applications, the general rules are:

- Look for optima among the stationary points.
- Stationary points of Q, when there is also a number of restrictions $R_1(\mathbf{p}) = 0$, $R_2(\mathbf{p}) = 0$, etc. are found from the equation

$$\nabla Q(\mathbf{p}) = \lambda_1 \nabla R_1(\mathbf{p}) + \lambda_2 \nabla R_2(\mathbf{p}) \ldots$$

- Any restriction of the type $S(\mathbf{p}) \leq 0$ is treated by checking the case of equality first; then, if a stationary point is found on the surface, check to see if ∇Q points out from the region allowed by $S(\mathbf{p}) < 0$. If it does, and this is a minimization, this is a valid stationary point; else go for interior points.

Series, in general

A series is a sum, with any number of terms – usually infinitely many. An infinite series is also called a formal series $\sum_{n=0}^{\infty} t_n$, when it is not intended to be evaluated.

But usually it is, and then one requires that the limit $= \lim_{N\to\infty} \sum_{n=0}^{N} t_n$ must exist. This is what defines the sum - it should be summed in that order, and if then the limit exists, it is a **convergent** series.

Example: Often used is Newtons binomial series, which is a **power series**,

$$(1+z)^\alpha = \sum_{n=0}^{\infty} \frac{\Gamma(\alpha+1)}{n!\,\Gamma(\alpha+1-n)} z^n, \quad \text{, e.g. with} \alpha = -1/2, \quad \sum_{n=0}^{\infty} \frac{(2n-1)!!}{(2n)!!}(-z)^n = \frac{1}{\sqrt{1+z}}$$

This series is a finite sum if α is a non-negative integer. Else, it is an infinite series, which converges if $|z| < 1$. In both cases, the sum is $(1+z)^\alpha$. Numerically, the sum is conveniently computed **recursively**, as

```
n:=0; t:=1; S:=1;
while (|t|>eps) do
  n:=n+1; t:=t*z*(alpha+1-n)/n; S:=S+t;
end do
```

Convergence of series

Some good rules for series:

- If the terms have alternating sign and are decreasing in size towards zero, the series converges.
- If the terms decrease at least as fast as the terms of a known convergent series, then the series converges. Comparing with the geometric series, we find: $S = \sum_n t_n$ is convergent, if $|t_n| < cr^n$ for all large enough n, where $0 < r < 1$.

A number of other useful convergence rules can be found in most math textbooks.

For a series with real terms, we can collect the positive terms, in order, into one 'subseries', and the negative terms in another. If both subseries converge, then so does the original series. If both diverge, then you can exhaust all terms by picking terms sometimes from one, sometimes from the other subseries. This way it is possible to arrive at any choosen value at the limit.

The first case is called 'absolute' or 'unconditional' convergence. In the other case, the series is conditionally convergent.

And of course, if only one of the subseries converges, then the series is divergent.

Series of functions

Criterion: An infinite series $\sum_n t_n$ converges absolutely, if the series $\sum_n |t_n|$ is convergent (hence that name).

Let $A = \sum_n a_n$ and $B = \sum_n b_n$ be two absolutely convergent series.

- They can be summed in any order, with the same sum.
- The series can be termwise added, and $\sum_n (a_n + b_n) = A + B$ converges absolutely.

The terms of the series can be functions of one or more variables. In that case, the absolute convergence is an important quality. Another is uniform convergence:

- Suppose the terms are functions $f_n(z)$, and that the series converges to $F(z)$ for all z in some set Q. If the supremum, $\sup_{z\in Q} |F(z) - \sum_{n=0}^{N} f_n(z)|$ tends to 0 when $N \to \infty$, the convergence is uniform.
- Particularly nice is the combined property, absolute uniform convergence, since then if $f_n(z)$ are continuous, then so is $F(z)$, for $z \in Q$.
- Weierstrass M-test: Suppose that $|f_n(z)| < M_n$ for $z \in Q$, were $\sum_{n=0}^{\infty} M_n$ is convergent. Then the function series converges absolutely and uniformly.

Taylor series of functions

A series of functions, $F(z) = \sum_{n=0}^{\infty} f_n(z)$, is in general convergent only for some particular set of z values. This is the convergence region of the series. Usually, the functions are analytic, and the region is some region in the complex plane.

Most common: Power series. The convergence region is the inside of a circle, $|z| < \rho$. The parameter ρ is the convergence radius, and it can be anything from 0 to ∞, depending on the series. Examples:

- For any given function $F(z)$, if all its derivatives $F^{(n)(a)}$ are known, the series

$$F(z) = \sum_{n=0}^{\infty} \frac{F^{(n)}(a)}{n!}(z-a)^n$$

is called a Taylor series.

- $\exp(z)$, $\cos(z)$, $\sin(z)$, $\sinh(z)$, etc; exponential-related functions have a Taylor series with $\rho = \infty$.
- $\tan(z)$ has $\rho = \pi/2$; $\arctan(z)$, and $(1+z)^\alpha$ all have $\rho = 1$
- $\exp(-1/|x|)$, for real x, has a well-defined Taylor series around $z = 0$, but this series is the constant 0. The function is not analytic at $z = 0$, and the convergence radius is $\rho = 0$.

The binomial series for $\sqrt{1+z}$

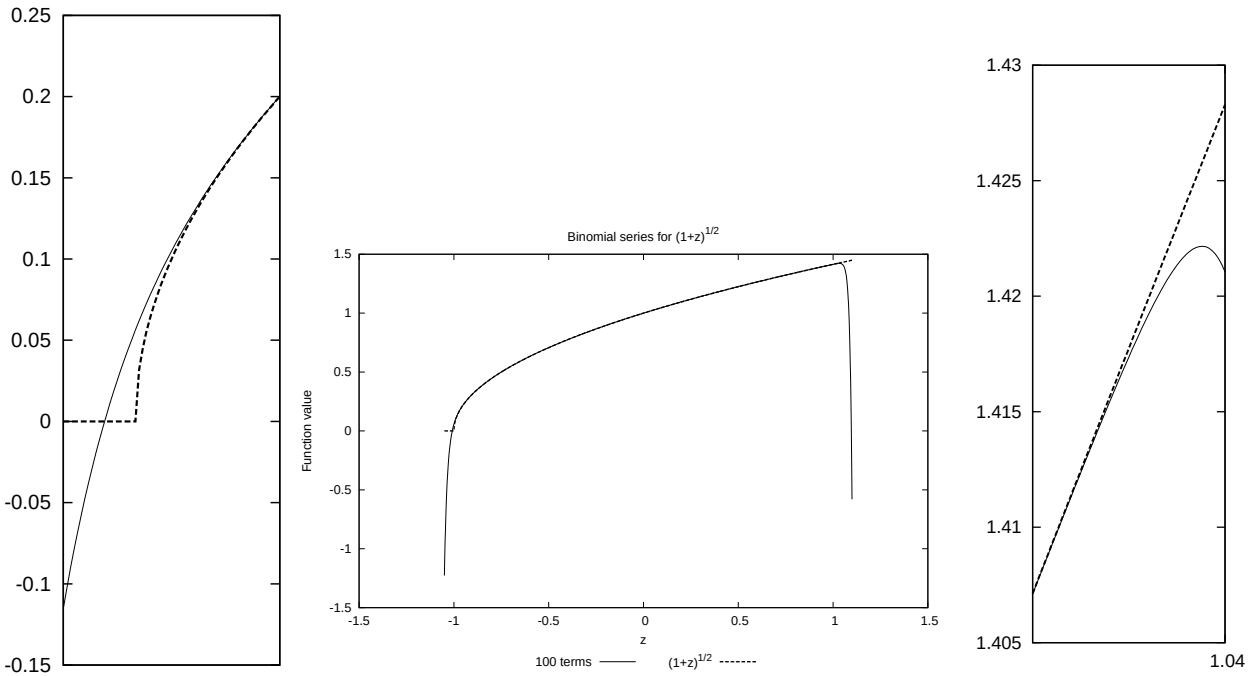


Figure 1: Newtons binomial series with exponent 1/2 truncated to 100 terms. The series converges in $z \in]-1, 1]$.

The series is clearly seen to diverge, not only for $z \leq -1$, which is due to the singularity, but also for $z > 1$, illustrating the concept of a convergence circle.

Series of other functions

A power series converges absolutely inside the convergence circle, and conditionally on its boundary except at the actual singularities. Inside the convergence circle, a power series can be differentiated or integrated by terms, which is a way of obtaining new series, and which also offers a powerful method for solving differential equations.

Analogous rules hold for other function series, e.g.

- Laurent series, which converge in an annulus of convergence,

$$F(z) = \sum_{n=-\infty}^{\infty} a_n z^n, \rho_1 < |z| < \rho_2$$

- Fourier series, which in general converges inside a band around the real axis, to a periodic function,

$$F(z) = \sum_{n=-\infty}^{\infty} c_n \exp(inx),$$

which is seen to be identical to a Laurent series with $z = \exp(ix)$ (so the above is true when the Laurent series has convergence annulus $\rho_1 < 1 < \rho_2$).

Series of polynomials

The series $F(z) = \sum_{n=0}^{\infty} a_n p_n(z)$, where $p_n(z)$ are polynomials of degree n, are called polynomial expansions. Typical examples are Chebyshev $(p_n(z) = T_n(z))$ or Legendre $(p_n(z) = P_n(z))$ polynomials.

Other examples are the Hermite and Laguerre polynomials, in the forms $\sum_{n=0}^{\infty} a_n H_n(z) \exp(-z^2/2)$ and $\sum_{n=0}^{\infty} a_n L_n(z) \exp(-z/2)$.

Fitting to an analytic function converges within a convergence region, which for Chebyshev and Legendre is the largest ellipse with foci at $z = -1$ and 1 within which the function is analytic. Similarly for Hermite and Laguerre, where the regions contain the real axis (Hermite) and the positive real axis (Laguerre).

A rather special, but very nice, polynomial series arises from interpolation polynomials. In analogy to a power series, this series is designed to give values and derivatives at more than one point, and the convergence region is then the largest region bounded by a so-called lemniscate curve surrounding the selected points and within which the function is analytic.

Chebyshev series

$$T_0 = 1; \quad T_1 = x; \quad T_{n+1} = 2xT_n - T_{n-1}$$
$$T_n(\cos\theta) = \cos n\theta; \quad T_n\left(T_m(x)\right) = T_{nm}(x)$$

Zeroes: $T_n(x_k^{(n)}) = 0$ where $x_k^{(n)} = \cos\left(\dfrac{2k-1}{2n}\pi\right), \quad k = 1, 2, \ldots, n$

Orthogonality: $\displaystyle\int_{-1}^{1} \frac{T_n(x)T_m(x)}{\sqrt{1-x^2}}\,dx = \begin{cases} \pi, & \text{if } n = m = 0, \\ \frac{\pi}{2}\delta_{nm} & \text{else} \end{cases}$

And also: $\displaystyle\sum_{k=1}^{n} T_i(x_k^{(n)})T_j(x_k^{(n)}) = \begin{cases} n, & \text{if } i = j = 0, \\ \frac{n}{2}\delta_{ij} & \text{else} \end{cases}$

Chebyshev series: $\displaystyle f(x) = {\sum_{n=0}^{\infty}}' a_n T_n(x), \quad \text{where } a_n = \frac{2}{\pi}\int_{-1}^{1} \frac{f(x)T_n(x)}{\sqrt{1-x^2}}\,dx$

Chebyshev series, usage

• The Chebyshev series converges absolutely and uniformly inside the largest ellipse, with foci $x = -1$ and $x = 1$, within which $f(x)$ is analytic.

• It is identical to the Fourier series of $f(\cos\theta)$, which converges e.g. for a continuous functions, and usually also if there are discontinuities (convergence in mean).

• It can be computed numerically and evaluated easily, also for expansions with a million terms.

• In the so-called pseudo-spectral methods, it is used together with the Fast Fourier Transform e.g. to simulate large-scale time evolution and spectral properties.

• If the means are at hand to compute $\mathbf{u} = \mathbf{H}\mathbf{v}$ many times with any vector $\mathbf{v}$, then also the set of vectors $T_n(\mathbf{H})\mathbf{v}$ can be computed and used with, e.g., Chebyshev expansions of $\exp(i\mathbf{H}t)$ to produce time evolution of various differential equations, e.g. the Schrödinger equation:

$$\mathbf{v}_{n+1} = 2\mathbf{H}\mathbf{v}_n - \mathbf{v}_{n-1} \quad \Leftrightarrow \quad \mathbf{v}_n = T_n(\mathbf{H})\mathbf{v}_0$$

Example: artificial spectrum for H atom

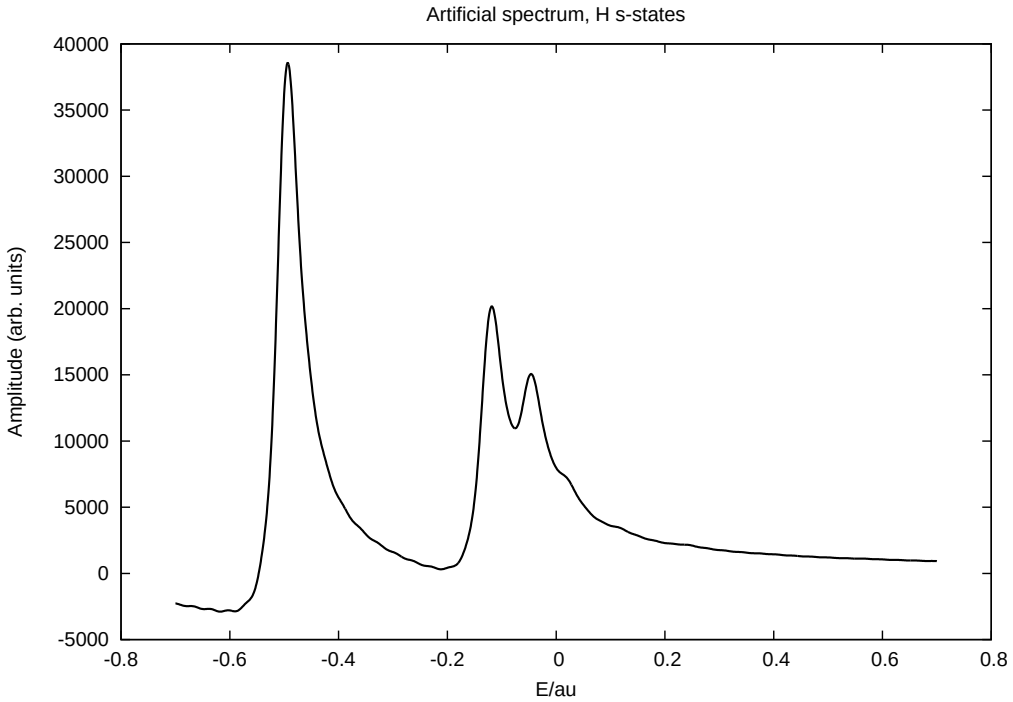


Figure 2: The spectrum of H (s states only) computed from matrix elements of $\left\langle \psi(0)|T_n(\hat{H})\psi(0)\right\rangle$, used in a Chebyshev series.

The wave function $\psi(0) = \psi_0$ at time $t = 0$ is a "thermalized" wave functions, with non-zero amplitudes for energies up to ≈ 5000 a.u. (!). Vectors $\psi_n = T_n(\hat{H})\psi_0$ with n up to 10^6 (!) were combined to yield the spectrum shown.

Perturbation series

One of the many uses of series expansions is in perturbation theory. For example, using the so-called Møller-Plesset perturbation theory, the electronic energy of a molecular system is obtained as a power series,

$$\varepsilon(\lambda) = \sum_{n=0}^{\infty} \varepsilon^{(n)} \lambda^n \tag{1}$$

The values $\lambda = 0$ and $\lambda = 1$ give the Hartree-Fock energy, and the exact energy, respectively. The terms beyond the zeroth order give the correlation energy. In practical calculations, this is rarely taken beyond second order, which is denoted MP2.

Using a basis set approximation, for small molecules, the MPn series can be computed to very high orders. This has been done, e.g. for water, the Ne atom, etc.

Such calculations show that the series, while still useful, is actually divergent in a number of simple cases.

The MPn series of Neon

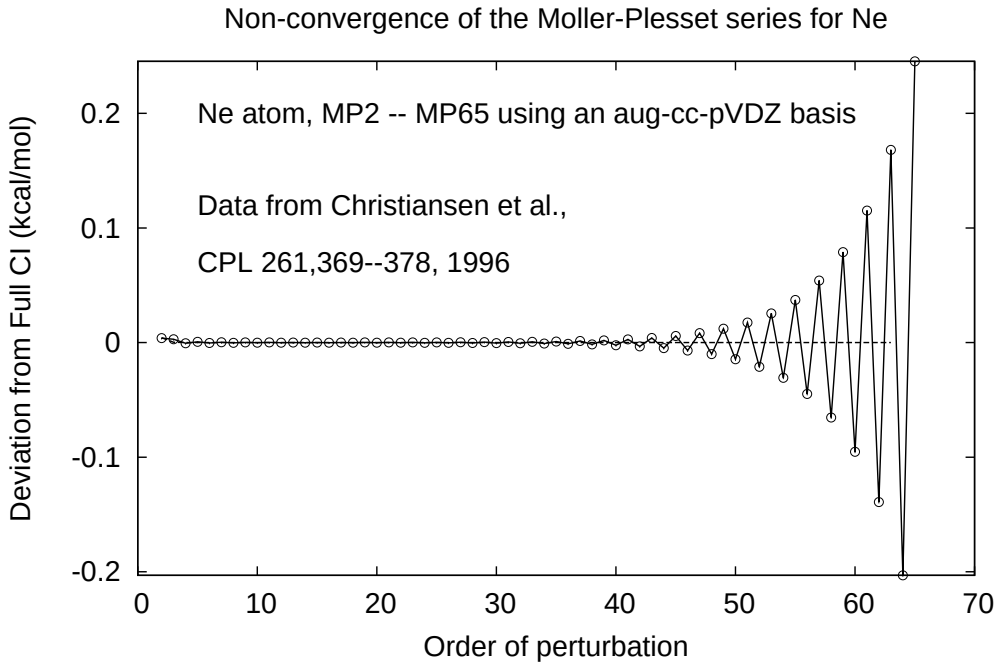


Figure 3: The errors of the Möller-Plesset series using an aug-cc-pVDZ basis. The dashed line in the middle is obtained by applying Shank's sequence transformation formula to the data.

Sequence transformations belong to the very large number of convergence acceleration tools that can dramatically improve convergence or give sensible results out of diverging series.

The Fourier transform

Function spaces characterized by integrability: Let $f(x)$ be defined on $x \in \mathbb{R}$. Then if $|f|$ is Lebesgue integrable, it is called an "L1 function", $f \in L^1(\mathbb{R})$. In that case, we are guaranteed that the Fourier transform $\tilde{f}$ exists. It can be back transformed, if also a similar integral for $\tilde{f}$ converges:

$$\tilde{f}(k) \stackrel{\text{def}}{=} \sqrt{\frac{1}{2\pi}} \int_{-\infty}^{\infty} f(x) \exp(-ikx)\, dx \qquad f(x) = \sqrt{\frac{1}{2\pi}} \int_{-\infty}^{\infty} \tilde{f}(k) \exp(ikx)\, dx \text{ a.e.}$$

(The Lebesgue integral is evaluated using rules that give values also to integrals with 'nasty' discontinuities, etc.)

"a.e."="almost everywhere" is a code phrase that means that the result is true except perhaps at isolated points. These are the points where the function f has a discontinuity. At any point x, the reconstructed function has the value $\lim_{\epsilon\to\infty}(f(x+\epsilon)+f(x-\epsilon)/2)$. In the Banach space $L^1(\mathbb{R})$, the reconstructed function, and the function $f(x)$, represent identically the same vector.

The Fourier transform

If we define "Lp functions" as those for which $|f(x)|^p$ are integrable, then evaluation rules have been designed such that they usually have a Fourier transform. In fact, if $f \in L^p(\mathbb{R})$, then $\tilde{f} \in L^q(\mathbb{R})$, where $1/p + 1/q = 1$, under fairly general circumstances.

There is the special case $p = q = 2$: then $f \in L^2(\mathbb{R})$ and $\tilde{f} \in L^2(\mathbb{R})$, they are both in the same space, which is a Hilbert space, and the Fourier transform is a unitary mapping. Quantum-mechanically, this implies that the state vector representing a particle can be transformed back and forth between position and momentum representation.

The Fourier gives back the original function if applied four times; in $L^2(\mathbb{R})$, this means that it is a unitary operator with eigenvalues $1, i, -1, -i$. Eigenfunctions are known (Harmonic Oscillator wave functions), which can be used to show that Heisenbergs Uncertainty Relation is an identity when the wave function is a Gaussian: The product of variances of the position and momentum, Var(x) and Var(k), is equal to $1/2$ if the wave function is Gaussian, else $> 1/2$ (assuming electron mass, atomic units, and in 1D only).

The Fourier transform

With $p = \infty$, the formula $1/p + 1/q = 1$ suggests that the Fourier transform should map $L^\infty(\mathbb{R})$ onto $L^1(\mathbb{R})$. $L^\infty(\mathbb{R})$ denotes functions with the maximum norm, a.k.a the uniform norm, i.e. functions such that $|f(x)|$ is bounded. The constant function $f(x) = 1$ is such a function. But the transform is not integrable.

The Dirac distribution is not an $L^1(\mathbb{R})$ function, although it is integrable by definition. Used in the inverse Fourier transform

$$\frac{1}{\sqrt{2\pi}} \int_\infty^\infty \delta(k) \mathrm{e}^{ikx}\, dk = \frac{1}{\sqrt{2\pi}}$$

so we can define the transform pair $f(x) = 1 \quad \tilde{f}(k) = \sqrt{2\pi}\delta(k)$. This (and other) 'tricks' make it possible to define a large set of transform pairs without contradictions. This has allowed the Fourier transform to become a workhorse in most areas of applied mathematics.

Fourier transform: some facts

$$g(k) = \tilde{f}(k) \quad \Leftrightarrow \quad \tilde{g}(x) = f(-x)$$

Linearity: $\widetilde{\alpha f + \beta g} = \alpha \tilde{f} + \beta \tilde{g}$

Scaling: $f(x) = g(\alpha x) \quad \Leftrightarrow \quad \tilde{f}(k) = \tilde{g}(k/\alpha)/|\alpha|$

Translation: $f(x) = g(x + t) \quad \Leftrightarrow \quad \tilde{f}(k) = \mathrm{e}^{ikt}\tilde{g}(k)$

Plancherel's theorem: $\int f(x)^* g(x)\, dx = \int \tilde{f}(x)^* \tilde{g}(x)\, dx$

Derivatives: $\widetilde{f'}(k) = ik\tilde{f}(k) \qquad \tilde{f}'(k) = -i\tilde{g}(k)$ where $g(x) = xf(x)$

Poisson sum: $\displaystyle\sum_{n=-\infty}^{\infty} f(\alpha n) = \frac{\sqrt{2\pi}}{\alpha} \sum_{n=-\infty}^{\infty} \tilde{f}(2\pi n/\alpha)$

Convolution: $\widetilde{f * g}(k) = \sqrt{2\pi}\tilde{f}(k)\tilde{g}(k)$

In higher dimensions than 1:

$$f(\boldsymbol{r}) = g(\boldsymbol{U}\boldsymbol{r}) \quad \Leftrightarrow \quad \tilde{f}(\boldsymbol{k}) = \tilde{g}(\boldsymbol{U}\boldsymbol{k})$$

if $\boldsymbol{U}$ is an orthogonal matrix, e.g. a 3D rotation matrix.

Sample Fourier pairs

$$\text{Heaviside:} \quad \tilde{\Theta}(k) = \frac{1}{\sqrt{2\pi}}\left(\frac{1}{ik} + \pi\delta(k)\right)$$

$$f(x) = \exp(-\alpha|x|) \quad \Leftrightarrow \quad \tilde{f}(k) = \sqrt{\frac{2}{\pi}}\frac{2\alpha}{\alpha^2 + k^2} \quad (\text{Re}(\alpha) > 0)$$

$$f(x) = \exp(-\alpha x^2) \quad \Leftrightarrow \quad \tilde{f}(k) = \frac{1}{\sqrt{2\alpha}}\exp(-k^2/4\alpha) \quad (\text{Re}(\alpha) > 0)$$

$$f(x) = \sin(\alpha x^2) \quad \Leftrightarrow \quad \tilde{f}(k) = \frac{1}{\sqrt{2\alpha}}\cos(k^2/4\alpha + \pi/4)$$

$$f(x) = \cos(\alpha x^2) \quad \Leftrightarrow \quad \tilde{f}(k) = \frac{1}{\sqrt{2\alpha}}\cos(k^2/4\alpha - \pi/4)$$

$$f(x) = \frac{1}{\cosh(x)} \quad \Leftrightarrow \quad \tilde{f}(k) = \sqrt{\frac{\pi}{2}}\frac{1}{\cosh(\pi k/2)}$$

(The Heaviside function: $\Theta(x) = 1$, if $x > 0$; else it is $= 0$.)

Fourier transforms in $\mathbb{R}^n$

Cartesian 3D: $f(\boldsymbol{r}) = f(x,y,z) \Leftrightarrow \tilde{f}((\boldsymbol{k}) = (2\pi)^{-3/2} \int\int\int f(\boldsymbol{r}) \mathrm{e}^{-i\boldsymbol{k}\boldsymbol{r}}\, d^3\boldsymbol{r}$

Polar 2D: $f(r,\theta) = f(r)\mathrm{e}^{im\theta} \Leftrightarrow \tilde{f}(k,\theta) = i^m \int_0^\infty f(r) J_m(kr)\, r\, dr\, \mathrm{e}^{im\theta}$

3D: $f(r,\theta,\phi) = f(r) Y_{lm}(\theta,\phi) \Leftrightarrow \tilde{f}(k,\theta,\phi) = \sqrt{\frac{2}{\pi}}(-i)^l \int_0^\infty f(r) j_l(kr)\, r^2\, dr\, Y_{lm}(\theta,\phi)$

The Fourier transform can be applied component-wise to vector-valued functions in a fixed ON basis. **Examples:** electrodynamics; relativity; plane-wave expansions in solid-state theory; particle physics. The mapping of operators is logical and intuitive:

$\nabla\phi(\boldsymbol{r})$	corresponds to	$i\boldsymbol{k}\tilde{\phi}(\boldsymbol{k})$
$\nabla \cdot \boldsymbol{v}(\boldsymbol{r})$	corresponds to	$i\boldsymbol{k} \cdot \tilde{\boldsymbol{v}}(\boldsymbol{k})$
$\nabla \times \boldsymbol{A}(\boldsymbol{r})$	corresponds to	$i\boldsymbol{k} \times \tilde{\boldsymbol{A}}(\boldsymbol{k})$

One nice use of eigenvectors: Natural orbitals.

Suppose that the exact wave function Ψ were known, and contains n electrons. We want to use at most N (where $N \geq n$) one-electron spin-orbitals $\{|\psi_p\rangle\}$, with $p \in \{1, 2, 3 \ldots, N\}$. One criterion on a good orbital set is that they allow a Full-CI wave function with large overlap with Ψ. There exists such an ordered set of such orbitals, that selecting the first N solves the N-orbital maximum-overlap problem, for any N.

This set of orbitals are the eigenfunctions of the 1-particle spin density of the wave function, ordered by decreasing eigenvalue. They are called natural (spin-)orbitals, and the eigenvalues are called natural occupation numbers. These are between 0 and 1, and add up to slightly less than n.

This application to Quantum Machanical systems was done by P. O. Löwdin, who showed a number of interesting properties.

In general, the use of eigenvectors to find an optimal approximation to e.g. matrices is old, and much used in numerical algebra, with great practical and economic value.

'Condensed' data: The Singular Value Decomposition.

Similarly to the natural orbitals: Suppose that a large $n \times m$ (or even infinite) general $n \times m$ matrix $\mathbf{A}$ is to be approximated as well as possible (in a least-square sense) by factorizing into smaller matrices:

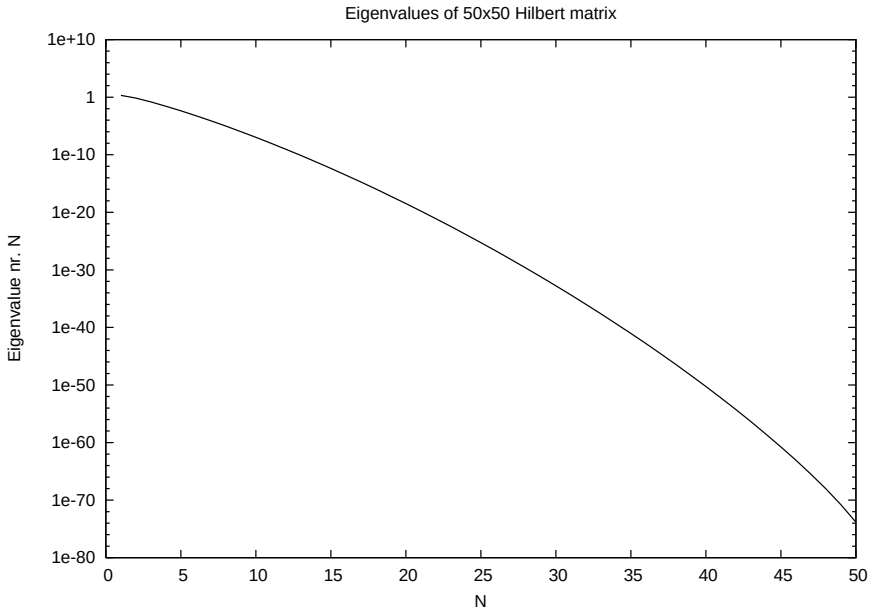
$$\mathbf{A} \approx \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad \text{or} \quad A_{ij} = \sum_{k=1}^{N} \sigma_k U_{ik} V_{jk} + R_{ij}$$

where N is assumed to be much smaller than n and m, and where the weights w_k and the arrays $\mathbf{U}$ and $\mathbf{V}$ are determined such as to minimize $||\mathbf{R}||^2$. Then this is a well-known minimization problem, yielding the weights as so-called 'Singular Values' and the columns of $\mathbf{U}$ and $\mathbf{V}$ as corresponding 'Singular Vectors' in a Singular Value Decomposition, SVD.

Just as for natural orbitals, the optimal choice for any N is to choose the first N singular vectors (ordered by decreasing singular value).

Example of an SVD.

The 50×50 Hilbert matrix $H_{kl} := 1/(k + l - 1)$ has 2500 elements, ranging in value from $1/99$ up to 1. Only a handful of eigenvalues are appreciably different from 0:



Since it is a square and symmetric matrix, its SVD is actually an ordinary spectral decomposition. Truncated to 9 vectors, the matrix is represented with a precision a precision of about 10^{-8}; 16 vectors give precision 10^{-16}.

Typical savings using matrix decomposition (SVD, Cholesky. . .) of two-electron integral data sets: 90% – 99.9%.

How an SVD is done.

Decomposing a **small** $n \times m$ matrix $\mathbf{A}$ is usually done by full diagonalization. Assume $n \geq m$, so rank($\mathbf{A}$) is at most m. Then compute the square positive (semi-)definite matrix $\mathbf{A}^\dagger\mathbf{A}$ and proceed by diagonalizing it:

$$\mathbf{A}^\dagger\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\dagger, \text{where} \quad \mathbf{V}\mathbf{V}^\dagger = \mathbf{V}^\dagger\mathbf{V} = \mathbf{1},$$

Λ is diagonal, and has m real non-negative diagonal elements.

Then let $\mathbf{\Sigma} = \mathbf{\Lambda}^{1/2}$, and compute

$$\mathbf{U} = \mathbf{A}\mathbf{V}\mathbf{\Sigma}^{-1}$$

.

Note that $\mathbf{\Sigma}$ is a diagonal matrix. We assume all elements are positive; if not, some vectors can be thrown away.

If neither n nor m is small, the SVD can be computed numerically using specialized linear algebra library calls.