## Example: Complete ON basis set in $\mathbf{L}^{2}([0, \infty[)$ :

All finite scalar product spaces can thus be represented faithfully by $\mathrm{R}^{n}$, but cn em infinite-dimensional spaces be represented in a similar manner? It turns out that all (separable) Hilbert spaces are equivalent to the space of infinite square-summable sequences, which can be regarded as infinite-dimensional arrays. The space of such sequences is called $l^{2}$.

The question remains whether some particular sequence of orthonormal functions constitutes a basis or not.

For the $L^{2}(\mathbb{R})$ space of function spaces, a standard test is to check if the completeness relation is fulfilled or not.

A certain orthonormal basis set $\left\{\phi_{n}(r)\right\}_{n=0}^{\infty}$ is complete if

$$
\delta(r-a)=\sum_{n=0}^{\infty} \phi_{n}(a) \phi_{n}(r)
$$

Here tested numerically for $a=2$. The sum is truncated at $n=399$ and $n=999$ :

Numerical test for completeness in $\mathbf{L}^{2}([0, \infty[)$ :
A certain set of orthonormal functions (Associated Laguerre functions of order 2) were tested by checking if they gave a reasonable approximation to a Dirac distribution.


It does not, it seems. Yet it is known to be complete (in a subspace of $L^{2}\left(\mathbb{R}_{+}\right)$, in this particular case), and in many cases, this would be an eminent basis set to use!

Conclusion: It is usually a waste of time to test for formal completeness numerically. A basis must be checked for completeness by other means, and anyway, its suitability for truncated expasions must be judged on other grounds.

## Completeness of other Hilbert function spaces

It was stated that a complete ON basis in $L^{2}(R)$ has the property

$$
\sum\left|\phi_{n}><\phi_{n}\right|=\hat{1} \quad \Leftrightarrow \quad \delta(r-a)=\sum_{n=0}^{\infty} \phi_{n}(a) \phi_{n}(r)
$$

since one may assume for a function space that the unit operator corresponds to the Dirac distribution.

But other function spaces may give other results. Example: The space of band-limited functions are defined by having a Fourier expansion

$$
f(x)=\frac{1}{2 \pi} \int_{-K}^{K} \tilde{f}(k) \exp (i k x) d k
$$

and has an ON basis with the "resolution of unity"

$$
\sum_{n=0}^{\infty} \phi_{n}(a) \phi_{n}(r)=\frac{\sin (K(a-x)}{\pi(a-x)} \Longrightarrow f(a)=\int_{-\infty}^{\infty} \frac{\sin (K(a-x)}{\pi(a-x)} f(x) d x
$$

The sum on the left is not a delta distribution, but it defines a projector onto the space of
band-limited functions, and therefore it acts as a unit operator within that space.

## Lecture 1B

## Pauli's spin matrices, and the quaternions

The three matrices

$$
\sigma_{x}=\left(\begin{array}{cc}
0 & 1 \\
1 & 0
\end{array}\right), \quad \sigma_{y}=\left(\begin{array}{cc}
0 & -i \\
i & 0
\end{array}\right), \quad \sigma_{z}=\left(\begin{array}{cc}
1 & 0 \\
0 & -1
\end{array}\right)
$$

are called Pauli spin matrices, and are treated as a single vector in 3D space, except having $2 \times 2$ matrices as components. $\sigma=\left(\sigma_{x}, \sigma_{y}, \sigma_{z}\right)$. $\boldsymbol{\sigma}$ is used in relativistic or two-component non-relativistic calculations (with a factor $\hbar / 2$ ) as the spin of electrons or other fermions. They have eigenvalues 1 and -1 , and $\left[\sigma_{x}, \sigma_{y}\right]=i \sigma_{z}$, and so on, in cyclic permutation of $x, y$, and $z$.

They are also used to represent the four basic quaternions,

$$
1=\left(\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right), i=i \sigma_{z}, j=i \sigma_{y}, k=i \sigma_{x}
$$

with the properties

$$
1 i=i 1=i, \quad i^{2}=-1, \quad i j=k=-j i
$$

and additional properties obtained by cyclically permuting $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$. The quaternions are used e.g. to represent rotations.

## Pauli's matrices

The matrices are used in combination with vectors as

$$
\boldsymbol{a} \cdot \boldsymbol{\sigma} \stackrel{\text { def }}{=} a_{x} \sigma_{x}+a_{y} \sigma_{y}+a_{z} \sigma_{z}=\left(\begin{array}{cc}
a_{z} & a_{x}-i a_{y} \\
a_{x}+i a_{y} & -a_{z}
\end{array}\right)
$$

A useful expression is frequently used:

$$
(\boldsymbol{a} \cdot \boldsymbol{\sigma})(\boldsymbol{b} \cdot \boldsymbol{\sigma})=(\boldsymbol{a} \cdot \boldsymbol{b}) \mathbf{1}+i(\boldsymbol{a} \times \boldsymbol{b}) \cdot \boldsymbol{\sigma}
$$

When rearranging formulas, it is a bother to have to keep $\boldsymbol{a}$ adjacent to $\sigma$ in order to use the scalar product; it is then common to use cartesian tensor notation and summation rules, e.g. as

$$
\sigma_{\mu} \cdots \sigma_{\nu} \cdots a_{\mu} b_{\nu} \cdots
$$

which is the same contraction as before, regardless of their position in a larger expression. The relative order of the $\sigma$ matrices must be kept, since they do not commute. (Components are then indexed 123 instead of $x y z$ ).

Complex hermitian $2 \times 2$ matrices can be conveniently parametrized as

$$
\left(c_{0}, c_{1}, c_{2}, c_{3}\right) \mapsto c_{0} \sigma_{0}+\cdots+c_{3} \sigma_{3}=\left(\begin{array}{cc}
c_{0}+c_{3} & c_{1}-i c_{2} \\
c_{1}+i c_{2} & c_{0}-c_{3}
\end{array}\right)
$$

by adding the symbol $\sigma_{0} \stackrel{\text { def }}{=} 1$ for the unit matrix.

## Lecture 1B

## Quaternions

Quaternions are intimately connected to rotations and to spin matrices. They are formally defined as a division algebra with a unit 1 and three equivalent imaginary units $i, j, k$ with properties $\boldsymbol{i}^{2}=\boldsymbol{j}^{2}=\boldsymbol{k}^{2}=\boldsymbol{i j k}=-1$.

Used for rotations in 3-space, they are usually represented as

$$
\boldsymbol{q}=(s, \boldsymbol{v})=s+v_{x} \boldsymbol{i}+v_{y} \boldsymbol{j}+v_{z} k
$$

and the basic computation rules

- Quaternion conjugation $\boldsymbol{q}^{*}=(s,-\boldsymbol{v})$,
- Product $\left(s_{1}, \boldsymbol{v}_{1}\right)\left(s_{2}, \boldsymbol{v}_{2}\right)=\left(s_{1} s_{2}-\boldsymbol{v}_{1} \cdot \boldsymbol{v}_{2}, \quad s_{1} \boldsymbol{v}_{2}+s_{2} \boldsymbol{v}_{1}+\boldsymbol{v}_{1} \times \boldsymbol{v}_{2}\right)$.
- Squared norm $\|(s, \boldsymbol{v})\|^{2}=\left(s^{2}+|\boldsymbol{v}|^{2}, \mathbf{0}\right)=s^{2}+|\boldsymbol{v}|^{2}$.
- Inverse $q^{-1}=q^{*} /\left\|q q^{*}\right\|$.

Note: quaternions $(0, \boldsymbol{v})$ are formally identified with ordinary vectors $\boldsymbol{v}$, and $(s, \mathbf{0})$ with scalars $s$.
A vector $\boldsymbol{v}$ in 3-space is rotated an angle $\theta$ around an axis with direction $\boldsymbol{u}$ by

$$
\boldsymbol{v}^{\prime}=\boldsymbol{q} \boldsymbol{v} \boldsymbol{q}^{*}, \quad \text { where } \boldsymbol{q}=(\cos \theta / 2, \boldsymbol{u} \sin \theta / 2)
$$

## Lecture 1B

## The Dirac $\langle\mathrm{bra}| \mathrm{c} \mid$ ket $\rangle$ notation

Since expressions such as $\left\langle b_{i} \mid \hat{A} b_{j}\right\rangle$ are so common, it has turned out to be practical to use a notation introduced by Dirac:

- A vector is denoted e.g. $\left|b_{j}\right\rangle$. Written like that, they are called "ket vectors".
- $\left\langle b_{i}\right|$ is a linear functional, which operates on any vector $|x\rangle$ to produce the scalar number $\left\langle b_{i} \mid x\right\rangle$.
(In this context, a functional is a mapping from functions to e.g. scalar values.)
- The functionals $\left\langle b_{i}\right|$ do themselves form a linear space, and the vectors are called "bra functionals" or "bra vectors". This vector space is called "the dual space".
- The matrix element $\left\langle b_{i} \mid \hat{A} b_{j}\right\rangle$ will now be written as $\left\langle b_{i}\right| \hat{A}\left|b_{j}\right\rangle$, a "Dirac bracket".


## The Dirac notation, Resolution of identity

An important point with Dirac notation is that, since it has a notation for the bra vectors as individual vectors rather than as typographic part of a scalar product, operators of the form $|x\rangle\langle y|$ can be inserted into formulas.

Ex: assume a complete and ON basis, $\left\langle b_{i} \mid b_{j}\right\rangle=\delta_{i j}$. Then the basis set representation of $x$ is

$$
\begin{gathered}
|x\rangle=\sum_{i}\left|b_{i}\right\rangle\left\langle b_{i} \mid x\right\rangle \\
\text { true for any } x
\end{gathered}
$$

Using Diracs notation we get

$$
\sum_{i}\left|b_{i}\right\rangle\left\langle b_{i}\right|=\hat{1}
$$

This formula (called the "resolution of identity") can now be inserted into operator or operator-vector products in many different ways, producing a number of useful formulas. But of course, the sum is not really a unit operator generally,
but a projector of functions in the space spanned by the basis vectors.

## Eigenvalues of Hermitian matrices

For an $n \times n$ hermitian matrix $\mathbf{A}$, there exists a set of $n$ orthonormal eigenvectors

$$
\mathbf{v}_{k}, k=1, \ldots n \text {, with real eigenvalues: }
$$

$$
\begin{aligned}
\mathbf{A} \mathbf{v}_{k} & =\lambda_{k} \mathbf{v}_{k} \\
\lambda_{1} \leq \lambda_{2} & \leq \cdots \leq \lambda_{n}
\end{aligned}
$$

If the eigenvalues are different, these eigenvectors are unique except that they can be scaled by a phase factor of magnitude 1 , (i.e., $\pm 1$ in the real case, $\exp (i \theta)$ in the complex case).

If any eigenvalues are equal, the corresponding eigenvectors are unique except that they can be transformed into an equivalent set of eigenvectors by multiplication with a unitary matrix (or real orthogonal, in the real case).

Similar rules apply to Hermitian operators in a Hilbert space, if the eigenvalues are distict. For the now, let's treat the finite case (matrices) and the infinite case (operators) together.

## Normal operators and matrices

For any linear op $\hat{A}$ in a Hilbert space, its Hermitian conjugate or adjoint, $\hat{A}^{\dagger}$ is defined such that

$$
\left\langle\hat{A}^{\dagger} x \mid y\right\rangle \equiv\langle x \mid \hat{A} y\rangle
$$

Ex: in $L^{2}(\mathbb{R})$, with $\hat{A}=x+\partial / \partial x$, we have that $\hat{A}^{\dagger}=x-\partial / \partial x$, since

$$
\int_{-\infty}^{\infty}\left((x-\partial / \partial x) f^{*}(x)\right) g(x) d x=\int_{-\infty}^{\infty} f^{*}(x)((x+\partial / \partial x) g(x)) d x
$$

from integration by parts, if both integrals exist.
Ex: For matrices, $\left(\mathbf{A}^{\dagger}\right)_{i j}=A_{j i}^{*}$, e.g., transpose + complex conjugation.
If $\hat{A} \hat{A}^{\dagger}=\hat{A}^{\dagger} \hat{A}$, then the operator is called normal.
(N.B. If some nonstandard definition is used for the scalar product, then the adjoint has to change accordingly!)

## Important types of normal ops and matrices

- Hermitian operators, $\hat{X}^{\dagger}=\hat{X}$ (Matrix: $\mathbf{X}^{\dagger}=\mathbf{X}$.)
- Antihermitian ops, $\hat{X}^{\dagger}=-\hat{X}$ (Matrix: $\mathbf{X}^{\dagger}=-\mathbf{X}$.)
- Unitary ops, $\hat{X}^{\dagger} \hat{X}=\hat{1}$ (Matrix: $\mathbf{X}^{\dagger} \mathbf{X}=\mathbf{1}$.)
- Idempotent ops, $\hat{X}^{2}=\hat{X}$ (Matrix: $\mathbf{X}^{2}=\mathbf{X}$.)

RULE: In an ON basis, a Hermitian operator is represented by a Hermitian matrix, \& cetera.

## Typical usage of different kinds of matrices or operators

Hermitian operators: e.g. multiplication with a local potential function. All operators that represent possible 'physical observables' are Hermitian. All their eigenvalues are real.

Antihermitian: Either an imaginary number times a hermitian operator, or perhaps a differentiation operator $(\partial / \partial x, \nabla, \ldots)$. All their eigenvalues are imaginary.

Unitary operators can be used to change from one basis set to another. They preserve orthonormality, scalar products (overlaps), and distances (norms).
NOTE: A unitary matrix or operator is the exponential of an antihermitian one,

$$
\mathbf{U}=\exp (\mathbf{A}), \quad \text { or } \quad \hat{U}=\exp (\hat{A})
$$

Idempotent operators are used as projection operators, or projectors.
Example: $\hat{a}^{\dagger} \hat{a}$ (for a fermion).

## Eigensystems: Diagonalization vs. Spectral resolution

Apart from being orthogonal, the eigenvectors can also always be scaled so they are orthonormal. Arrange them as columns in a matrix
U. Orthonormality gives

$$
\mathbf{U}^{\dagger} \mathbf{U}=\left(\begin{array}{c}
\mathbf{v}_{1}^{\dagger} \\
\vdots \\
\mathbf{v}_{n}^{\dagger}
\end{array}\right)\left(\mathbf{v}_{1}, \ldots, \mathbf{v}_{n}\right)=\mathbf{1}
$$

Also $\mathrm{UU}^{\dagger}=1$. Put eigenvalues in diagonal matrix $\mathbf{D}$. We get

$$
\mathbf{A U}=\mathbf{U D} \text { (Eigensystem equation) or } \mathbf{U}^{\dagger} \mathbf{A} \mathbf{U}=\mathbf{D}
$$

(Diagonalisation) or $\mathbf{A}=\mathrm{UDU}^{\dagger}$ (Spectral resolution).
(Note: For real matrices, replace $\dagger$ with the transpose).

## Matrix and operator functions

Normal matrices or operators have a spectral resolution of the form $\mathbf{A}=\mathrm{UDU}^{\dagger}$, where the matrix $\mathbf{U}$ is unitary.

This implies that

$$
\begin{gathered}
\mathbf{A}^{2}=\mathbf{U D U}^{\dagger} \mathbf{U D U}^{\dagger}=\mathbf{U D}^{2} \mathbf{U}^{\dagger} \\
\text { since } \mathbf{U}^{\dagger} \mathbf{U}=1 .
\end{gathered}
$$

A moments reflection tells us that similar rules apply to any non-negative power $\mathbf{A}^{n}$.

If the matrix is invertible, the same applies to negative powers.
Forming linear combinations, we can get any polynomial; by taking limits, any smooth function can be approximated arbitrarily well by a sequence of polynomials.

## Matrix and operator functions

Conclusion: For a matrix, we can define unambigously any smooth function in analogy to the functions of numbers.

Caveat: For operators, this way of defining functions works strictly as implied above, if the eigenvalues are contained in a bounded set, i.e., the operator is bounded.

The point: The functions get to be defined as

$$
f(\mathbf{A})=\mathbf{U} \boldsymbol{\Lambda} \mathbf{U}^{\dagger} \quad \text { where } \boldsymbol{\Lambda}=\operatorname{diag}\left(f\left(\lambda_{1}\right), \ldots, f\left(\lambda_{n}\right)\right)
$$

and for operators with spectral resolution $\hat{A}=\sum \lambda_{n}\left|\psi_{n}\right\rangle\left\langle\psi_{n}\right|$ in Hilbert spaces,

$$
f(\hat{A})=\sum f\left(\lambda_{n}\right)\left|\psi_{n}\right\rangle\left\langle\psi_{n}\right|
$$

## Often used matrix/operator functions

$$
\text { Let } \mathbf{A}=\sum_{k} \lambda_{k} \mathbf{v}_{k} \mathbf{v}_{k}^{\dagger} \text {, a normal matrix. }
$$

The resolvent, $(\mathbf{A}-z \mathbf{1})^{-1}$, is defined for any $z$ that is not in the spectrum of $\mathbf{A}$.
The spectral resolution of the resolvent is

$$
(\mathbf{A}-z \mathbf{1})^{-1}=\sum_{k}\left(\lambda_{k}-z\right)^{-1} \mathbf{v}_{k} \mathbf{v}_{k}^{\dagger}
$$

The exponential function, is often defined as the solution of a differential equation,

$$
\exp (\mathbf{t} \mathbf{A})=\sum_{k} \exp \left(t \lambda_{k}\right) \mathbf{v}_{k} \mathbf{v}_{k}^{\dagger} \quad \Leftrightarrow \quad \frac{d}{d t} \exp (t \mathbf{A})=\mathbf{A} \exp (t \mathbf{A})
$$

The square root is unambigously defined if $\mathbf{A}$ is a positive semidefinite matrix, i.e., all eigenvalues are non-negative. If there are negative or complex eigenvalues, one has to decide which branch of the square root they should belong to.

The logarithm need a similar consideration of branches. In addition, A must be non-singular.

## Derivatives, Functionals, and Functional Derivatives

For ordinary functions $f(x)$, the derivative

$$
\frac{\partial f}{\partial x} \text { or } f_{x}^{\prime} \text { or just } f^{\prime}
$$

is familiar to all, and very useful, e.g.:

- to express an optimization condition
- to solve equations efficiently (Newton-Raphson)
- for approximation by Taylor expansion

Questions: Can this concept be extended? Is differentiation possible if $x$ is e.g. a vector, a matrix, a function, an operator... ?

Is such an extension useful?
Answers: Yes, and Yes!

## The general derivative concept

If, for a given $x$ and a variable $t$,

$$
f(x+t)-f(x)=D(t)+e(t)
$$

where

$$
D(t) \quad \text { is linear in } t
$$

$$
\lim _{|t| \rightarrow 0} \frac{e(t)}{|t|}=0
$$

then the derivative of $f$ at $x$ can be defined as the linear mapping $D$.
This definition requires sums, limits etc. can be formed. $x$ and $t$ must belong to some normed linear space. $f(x), f(x+t)$ must also belong to some normed linear space.

All combinations, where $f$ values are scalars, vectors, functions, ..., and where $x$ values are scalars, vectors, functions, ..., are POSSIBLE and VERY USEFUL.

## Derivative of a scalar wrt a vector variable

In any particular application, we need to represent the linear mapping $D(t)$ in some practical form, usually as a product.

Instead of $D(t)$, we prefer to write

$$
f(x+t)-f(x)=\left(\frac{\partial f}{\partial x}\right) t+\text { rest term }
$$

How this product is to be defined depends on the application.
Example: $f(\mathbf{x})$ is a scalar function of a vector $\mathbf{x}$ and a scalar product is defined.
Then $\left(\frac{\partial f}{\partial \mathrm{x}}\right)$ is defined as the vector g with property

$$
f(\mathbf{x}+\mathbf{t})-f(\mathbf{x})=\mathbf{g} \cdot \mathbf{t}+\text { rest term }
$$

and such a vector is also called the gradient of $f$ and is written as $\nabla f$ ('nabla $\mathrm{f}^{\prime}$ ).

## Derivative of a vector wrt a vector variable

If neither the function value nor the variable is a simple scalar, the derivative becomes a composite object of 'higher rank' than either of them.

Example: $\mathbf{f}(\mathbf{x})$ is a vector-valued function of a vector $\mathbf{x}$.
Then $\left(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right)$ can be defined as the matrix G with property

$$
\mathbf{f}(\mathbf{x}+\mathbf{t})-\mathbf{f}(\mathbf{x})=\mathbf{G} \mathbf{t}+\text { rest term }
$$

where the product is a matrix times vector product. In some contexts, such a matrix is called a Jacobian.
(The derivative of a gradient is the second derivative of a scalar field. It is a symmetric matrix, called the Hessian).

## Derivative of a scalar wrt a matrix variable

Derivatives with respect to matrices is analogous to derivatives with respect to vectors, except that one must decide how the product is to be defined.

Example : $f(\mathbf{A})$ is a scalar function of a real $n \times m$ matrix $\mathbf{A}$.
Then $\left(\frac{\partial f}{\partial \mathbf{A}}\right)$ can be defined as the $n \times m$ matrix $\mathbf{G}$ with property

$$
f(\mathbf{A}+\mathbf{T})-f(\mathbf{A})=\sum_{i j} G_{i j} T_{i j}+\text { rest term }
$$

There is no standard notation for derivatives $\mathrm{w} r \mathrm{t}$ matrices. Another common definition of the scalar product is as $\sum_{i j} G_{j i}^{*} T_{i j}$ which can also be written as $\operatorname{tr}\left(\mathbf{G}^{\dagger} \mathbf{T}\right)$. Then the derivative with respect to an $n \times m$ matrix is an $m \times n$ matrix.

## Derivative of a scalar wrt a function-valued variable

The case where $x$ is a function and $f(x)$ is a scalar is analogous to the first example, where $x$ was a vector. Such a function $f$, that takes a function as its argument, is called a functional.

Example: The perturbed ground state energy of a molecule in a perturbing electrostatic potential field $V(\mathbf{r})$ can be regarded as a function $E(V)$

However, in this case it is customary to write $E[V]$ in order to show clearly that this is a functional.

The derivative of $E[V]$ is written as $\frac{\delta E}{\delta V}$ or sometimes $\frac{\delta E[V]}{\delta V}$. It is a function of $\mathbf{r}$, such that

$$
E[V+W]-E[V]=\iiint \frac{\delta E}{\delta V}(\mathbf{r}) \cdot W(\mathbf{r}) d^{3} \mathbf{r}+\text { rest term }
$$

(Actually, this function happens to be identical to the charge density function.)

## A simple optimization problem

It is a very common problem to find that function which minimizes or maximizes some given scalar functional. When the function is a curve or a surface, the solution is called an extremal.

Other problems may involve finding the set of orbital functions that minimize a given energy functional, etc.

Consider the simple problem of minimizing

$$
Q[\chi]=\int_{0}^{1}\left(\frac{d \chi(s)}{d s}\right)^{2} d s
$$

with side conditions $\chi(0)=0, \chi(1)=1$. To find the functional derivative, we make a small perturbation in $\chi$ and evaluates the result. Integration by parts is used to get an expression containing just $\delta \chi$ :

$$
\begin{aligned}
\delta Q[\chi] & =\int_{0}^{1} 2\left(\frac{d}{d s} \chi(s)\right)\left(\frac{d}{d s} \delta \chi(s)\right) d s \\
& =2\left[\left(\frac{d}{d s} \chi(s)\right)(\delta \chi(s))\right]_{s=0}^{1}-2 \int_{0}^{1}\left(\frac{d^{2}}{d s^{2}} \chi(s)\right)(\delta \chi(s)) d s
\end{aligned}
$$

A simple optimization problem... (continued).
From the boundary conditions, we must require $\delta \chi(0)=0$ at both end points, so the first term is 0 :

$$
\delta Q[\chi]=-2 \int_{0}^{1}\left(\frac{d^{2}}{d s^{2}} \chi(s)\right)(\delta \chi(s)) d s
$$

We can conclude that

$$
\frac{\delta Q[\chi]}{\delta \chi}=-2\left(\frac{d^{2}}{d s^{2}} \chi(s)\right)
$$

Minimization condition is $\frac{\delta Q[\chi]}{\delta \chi}=0$, which together with the boundary conditions gives the

$$
\text { solution } \chi(s)=s
$$

## 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, which may describe the amplitudes of relativistic wave functions used in some equations,
- Tensor fields $\mathbb{S} \mapsto\left(\mathbb{R}^{4}\right)^{N}$, 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

$$
C_{x}=A_{y} B_{z}-A_{z} B_{y} \quad C_{y}=A_{z} B_{x}-A_{x} B_{z} \quad C_{z}=A_{x} B_{y}-A_{y} B_{x}
$$

Some vector product rules are

$$
\begin{aligned}
\mathbf{A} \times \mathbf{B}= & -\mathbf{B} \times \mathbf{A} \\
\mathbf{A} \cdot(\mathbf{B} \times \mathbf{C})= & \operatorname{det}\left(\begin{array}{lll}
A_{x} & B_{x} & C_{x} \\
A_{y} & B_{y} & C_{y} \\
A_{z} & B_{z} & C_{z}
\end{array}\right) \\
\mathbf{A} \cdot(\mathbf{B} \times \mathbf{C})= & \mathbf{B} \cdot(\mathbf{C} \times \mathbf{A})=\mathbf{C} \cdot(\mathbf{A} \times \mathbf{B})=(\mathbf{A} \times \mathbf{B}) \cdot \mathbf{C}=\cdots \\
\mathbf{A} \times(\mathbf{B} \times \mathbf{C})= & (\mathbf{A} \cdot \mathbf{C}) \mathbf{B}-(\mathbf{A} \cdot \mathbf{B}) \mathbf{C} \\
\text { (but note that } & \mathbf{A} \times(\mathbf{B} \times \mathbf{C}) \neq(\mathbf{A} \times \mathbf{B}) \times \mathbf{C}!)
\end{aligned}
$$

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 $x \rightarrow y \rightarrow z \rightarrow x \ldots$. 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 $\nabla$, 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:

$$
\begin{aligned}
\nabla u & =\left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}, \frac{\partial u}{\partial z}\right) \quad \nabla \cdot \mathbf{a}=\frac{\partial a_{x}}{\partial x}+\frac{\partial a_{y}}{\partial y}+\frac{\partial a_{z}}{\partial z} \\
\nabla \times \mathbf{a} & =\left(\frac{\partial a_{y}}{\partial z}-\frac{\partial a_{z}}{\partial y}, \frac{\partial a_{z}}{\partial x}-\frac{\partial a_{x}}{\partial z}, \frac{\partial a_{x}}{\partial y}-\frac{\partial a_{y}}{\partial x}\right)
\end{aligned}
$$

Rules:

$$
\begin{gathered}
\nabla u v=u \nabla v+v \nabla u \quad \nabla \cdot u \mathbf{a}=u \nabla \mathbf{a}+\mathbf{a} \cdot \nabla u \quad \nabla \cdot \mathbf{a} \times \mathbf{b}=\mathbf{b} \cdot \nabla \times \mathbf{a}-\mathbf{a} \cdot \nabla \times \mathbf{b} \\
\nabla \cdot(\nabla \times \mathbf{a})=\mathbf{0} \quad \nabla \times(\nabla u)=0 \\
\nabla \cdot(\nabla u)=\nabla^{2} u
\end{gathered}
$$

## Integrals in vector calculus

The volume integral is simply a triple integral, where the integration element depends on the coordinate system used, e.g. $d \Omega=d x d y d z$ (Cartesian) or $d \Omega=r^{2} \sin \theta d r d \theta d \phi$ (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 $(u, v)$,

$$
\left\{\begin{array}{l}
x=x(u, v) \\
y=y(u, v) \\
z=z(u, v)
\end{array} \quad d \mathbf{S}=\left(\frac{\partial \mathbf{r}}{\partial u} \times \frac{\partial \mathbf{r}}{\partial v}\right) d u d v\right.
$$

so for e.g. a spherical polar system, suppose we integrate over a spherical shell with radius R ,
$d \mathbf{S}=\left(\begin{array}{c}R \cos \theta \cos \phi \\ R \cos \theta \sin \phi \\ -R \sin \theta\end{array}\right) \times\left(\begin{array}{c}R \sin \theta \sin \phi \\ -R \sin \theta \cos \phi \\ 0\end{array}\right) d \theta d \phi=\left(\begin{array}{c}\sin \theta \cos \phi \\ \sin \theta \sin \phi \\ \cos \theta\end{array}\right) R^{2} \sin \theta d \theta d \phi$

## Integration by parts

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

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

Integrating both sides gives the formula for partial integration,

$$
\begin{aligned}
\int_{x=a}^{b} \frac{d f(x)}{d x} g(x) d x & =f(b) g(b)-f(a) g(a)-\int_{x=a}^{b} f(x) \frac{d g(x)}{d x} d x \\
& =[f(x) g(x)]_{a}^{b}-\int_{x=a}^{b} f(x) \frac{d g(x)}{d x} d x
\end{aligned}
$$

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

$$
\begin{aligned}
& \iint_{\partial \Omega} \mathbf{a} \cdot d \boldsymbol{\Sigma}=\iiint_{\Omega} \nabla \cdot \mathbf{a} d \Omega \quad \text { (Gauss' theorem) } \quad \text { and } \iint_{\partial \Omega} \mathbf{a} \times d \boldsymbol{\Sigma}=-\iiint_{\Omega} \nabla \times \mathbf{a} d \Omega \\
& \int_{\partial S} \mathbf{a} \cdot d \mathbf{r}=\iint_{S} \nabla \times \mathbf{a} \cdot d \mathbf{S} \quad \text { (Stokes' theorem) } \quad \text { and } \int_{\partial S} \mathbf{a} \times d \mathbf{r}=-\iint_{S}(d \mathbf{S} \times \nabla) \times \mathbf{a}
\end{aligned}
$$

Green's theorem

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

Other theorems of this kind:

$$
\begin{gathered}
\iint_{\partial \Omega} u d \boldsymbol{\Sigma}=\iiint_{\Omega} \nabla u d \Omega \\
\int_{\partial S} u d \boldsymbol{r}=\int_{S}(d \boldsymbol{S} \times \nabla u)
\end{gathered}
$$

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

$$
\begin{aligned}
\iint_{\partial \Omega} \boldsymbol{E} \cdot d \boldsymbol{\Sigma}=\frac{1}{\varepsilon_{0}} \iiint_{\Omega} \rho d \Omega & \Leftrightarrow \nabla \cdot \boldsymbol{E}=\frac{\rho}{\varepsilon_{0}} \\
\iint_{\partial \Omega} \boldsymbol{B} \cdot d \boldsymbol{\Sigma}=0 & \Leftrightarrow \nabla \cdot \boldsymbol{B}=0 \\
\int_{C} \boldsymbol{E} \cdot d \boldsymbol{r}=-\int_{S} \frac{\partial \boldsymbol{B}}{\partial t} \cdot d \boldsymbol{S} & \Leftrightarrow \nabla \times \boldsymbol{E}=-\frac{\partial \boldsymbol{B}}{\partial t} \\
\int_{C} \boldsymbol{B} \cdot d \boldsymbol{r}=\mu_{0} \int_{S}\left(\boldsymbol{j}+\varepsilon_{0} \frac{\partial \boldsymbol{E}}{\partial t}\right) \cdot d \boldsymbol{S} & \Leftrightarrow \nabla \times \boldsymbol{B}=\mu_{0} \boldsymbol{j}+\mu_{0} \varepsilon_{0} \frac{\partial \boldsymbol{E}}{\partial t}
\end{aligned}
$$

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 $\Omega$,

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

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, $l \in\{0,1,2, \ldots\}$ and $m \in\{l, l-1, l-2, \ldots,-l\}$, and $(r, \theta, \phi)$ 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.

