Section IV

Mathematical annex

Appendix A
Mathematical Annex

- *Numbers* used in the book are natural integers $0, 1, 2, \cdots \in \mathbb{N}$ or signed integers $\cdots, -1, 0, 1, \cdots \in \mathbb{Z}$ mainly as indices. Physical quantities take their values in (sets of) real numbers $\mathbb{R}$ or complex numbers $\mathbb{C}$. Real and complex numbers are called *scalars*. A complex number z can be viewed as a pair (a, b) of real numbers denoted by $z = a + ib$ where $i^2 = -1$. The real number $a = \Re\mathrm{e}(z)$ is called the real part, whereas the real number $b = \Im\mathrm{m}(z)$ is called the imaginary part. The *complex conjugate* of z is $\bar{z} = a - ib$ and its *modulus* or absolute value is $|z| = \sqrt{z\bar{z}} = \sqrt{a^2 + b^2}$. The exponential representation $z = |z|e^{i\theta(z)}$ involves the modulus and the *argument* $\theta(z)$. The function $\theta(z)$ is usually chosen as being a continuous function on an open set of $\mathbb{C}$, called a continuous determination of the argument. Two continuous determinations on the same set differ by an integer multiple of 2π and the principal determination is defined by $\theta(z) : \mathbb{C}\backslash\{\Re\mathrm{e}(z) \leq 0\} \to]-\pi, +\pi[$ and given by $\theta = 2\arctan\left(\dfrac{b}{a + \sqrt{a^2 + b^2}}\right)$.

- *Vector space*: Given scalars $\mathbb{R}$ or $\mathbb{C}$, a vector space V is a set of elements, called vectors, with two operations: the addition of two vectors and the product of a vector by a scalar, both defined axiomatically in a way quite obvious for the common intuition and giving a new vector as the result. Given a set of vectors $\mathbf{v}^i$ of a vector space V and a set of scalars a_i, the combination of products by a scalar and vector additions is a new vector called a *linear combination* and denoted $a_i\mathbf{v}^i$.

This expression uses the *Einstein summation convention* on repeated indices: $a_i \mathbf{v}^i$ denotes $\sum_i a_i \mathbf{v}^i$. If things are written with care, it always involves an upper and a lower index.
The null scalar 0 and the null vector $\mathbf{0}$ are defined so that $0\mathbf{v} = \mathbf{0}$ and $\mathbf{0} + \mathbf{v} = \mathbf{v}$ for any vector. Given vectors are said to be *linearly independent* if the only linear combination that gives $\mathbf{0}$ is $\sum_i 0\mathbf{v}^i$, else they are said to be linearly dependent. The maximum number of vectors that you can find in a set of linearly independent vectors of V is called the *dimension* $\dim(V)$ of the vector space. A maximum set of linearly independent vectors is called a *basis* and any vector can be expressed as a linear combination of the basis vectors.
The canonical examples of vector spaces are the sets of n-tuples of scalars: $(b^1, \cdots, b^n)$ or $(c^1, \cdots, c^n)$ with $b^1, \cdots, b^n, c^1, \cdots, c^n \in \mathbb{C}$. Explicitly, the addition is given by $(b^1, \cdots, b^n) + (c^1, \cdots, c^n) = (b^1 + c^1, \cdots, b^n + c^n)$ and the product by a scalar is given by $a(b^1, \cdots, b^n) = (ab^1, \cdots, ab^n)$ with $a \in \mathbb{C}$.
Given a set of vectors $\mathbf{v}^1, \cdots, \mathbf{v}^p$, they *span* a vector space $V = \text{span}(\mathbf{v}^1, \cdots, \mathbf{v}^p)$, which is the set of all the linear combinations of those vectors.

- *Multi-linear forms*: a p-multi-linear form is a map $f : \{\mathbf{v}^1, \mathbf{v}^2, \cdots, \mathbf{v}^p\} \in V \to f(\mathbf{v}^1, \mathbf{v}^2, \cdots, \mathbf{v}^p) \in \mathbb{C}$ from p vectors to a scalar (complex numbers $\mathbb{C}$ are taken as the default set of scalars) such that $f(\mathbf{v}^1, \cdots, a_i \mathbf{v}^i, \cdots, \mathbf{v}^p) = a_i f(\mathbf{v}^1, \cdots, \mathbf{v}^i, \cdots, \mathbf{v}^p)$, i.e., the map is linear in all entries. Particular cases are the *linear forms* that map linearly a single vector to a scalar and *bilinear forms* that map two vectors to a scalar.
  Note that the set of p-multilinear forms for a given p is obviously a vector space itself since a linear combination of m p-forms $f_1, \cdots, f_m$ is defined as a new m-multilinear form by $(a^i f_i)(\mathbf{v}^1, \cdots, \mathbf{v}^p) = a^i f_i(\mathbf{v}^1, \cdots, \mathbf{v}^p)$. In particular, the set of linear forms on a vector space V is called the *algebraic dual vector* space and is denoted by V^*. In the finite dimension case, the dual space V^* is *isomorphic* to V, i.e., there is a one-to-one correspondence between the spaces preserving the various operations. For instance, if a vector $\mathbf{v}$ of V is a set of n

numbers $(b^1, \cdots, b^n)$, a general linear form f is also given by a set of n numbers $(c_1, \cdots, c_n)$ so that the action of the linear form on the vector is $f(\mathbf{v}) = b^1 c_1 + \cdots + b^n c_n = b^i c_i$.
Note the lower indices for the coefficients of forms.
To stress on the duality, the action of f can be written as a *duality product* $f(\mathbf{v}) =< f, \mathbf{v} >$.
In the case of the canonical example, a basis is given by vectors $\mathbf{e}_i = (0, \cdots, 1, \cdots, 0)$ where 1 in the i^{th} position is the only nonzero component.
Given a basis $\mathbf{e}_i$ of a vector space V, the set of vectors $\mathbf{e}^j$ of V^* such that $< \mathbf{e}_i, \mathbf{e}^j >= \delta_i^j$, where δ_j^i is the *Kronecker delta symbol* equal to 1 when $i = j$ and else equal to 0, is a basis of V^*, called the *dual basis* and $< b^i \mathbf{e}_i, c_j \mathbf{e}^j >= b^i c_i$.

- Functional analysis and function spaces: Some sets of functions are vector spaces of particular interest, called function spaces or functional vector spaces. The main characteristic of such spaces is that they are usually infinite dimensional and not isomorphic to their dual spaces. For instance, the set of square integrable functions on the $[0, 2\pi]$ interval is a functional vector space. A basis is the infinite set of trigonometric functions $\{1, \cos(x), \sin(x), \sin(2x), \cos(2x), \cdots\}$.
- Important examples of function spaces: Let Ω be a bounded open set in $\mathbb{R}^n$ or $\mathbb{R}^n$ itself. We denote by $C^m(\Omega)$ the set of functions that are continuous in Ω together with all their first m derivatives.
  In a multi-variable context, the following notation is introduced: $\alpha = (\alpha_1 \cdots \alpha_n) \in \mathbb{N}^n$ is a multi-index with

$$|\alpha| = \sum_i \alpha_i \text{ and } D^\alpha = (-i)^{|\alpha|} \frac{\partial^{|\alpha|}}{\partial^{\alpha_1} x^1 \cdots \partial^{\alpha_n} x^n}.$$

  The first m derivatives of a function f are the $D^\alpha f$ with $|\alpha| \leq m$.
  The subset consisting of the functions that have a compact support (i.e., which vanish outside a compact subset of Ω) is denoted by $C_0^m(\Omega)$. The corresponding space of infinitely differentiable functions (also called smooth functions) is denoted by $C^\infty(\Omega)$ and $C_0^\infty(\Omega)$.

- *Tensor product*: The tensor product is a fundamental operation of linear algebra, which transforms multilinear forms to linear ones. Given two vector spaces V and W, there exists a unique vector space (up to an isomorphism) denoted by $V \otimes W$ and such that for any vector space U, the space of linear maps from $V \otimes W$ to U is isomorphic to the space of bilinear maps from $V \times W$ to U where $V \times W$ is the Cartesian product, i.e., the set of pairs $(v \in V, w \in W)$. The vector space $V \otimes W$ is called the tensor product.
  For example, in the case of functional spaces, the tensor product of two functions $f(x)$ and $g(x)$ of a single variable is the two-variable function given by $f \otimes g(x, y) = f(x)g(y)$. As another example, consider a vector space V of finite dimension n and its dual space V^*. The tensor product $V^* \otimes V^*$ associates to a pair of linear forms on V, i.e., an element of $V^* \times V^*$ given by the coefficients a_i and b_i (a single index), a bilinear form on V given by the coefficients $(a \otimes b)_{i,j} = a_i b_j$ (two indices) such that the action on two vectors given by the coefficients v^i and w^i is the scalar $(a \otimes b)_{i,j} v^i w^j = a_i b_j v^i w^j$. Tensor product is an important tool to manipulate multilinear maps as vectors spaces. For instance, if V is a vector space and V^* its dual, p-multilinear maps form the $\otimes^p V^*$ vector space. A basis of this space may be built by taking the tensor products of p basis vectors of V^* and the dimension of this space is $\dim(\otimes^p V^*) = \dim(V)^p$.
- Direct sum: A subspace U of V is a subset of V such that any linear combination of elements of U is in U, which is, therefore, itself a vector space. Suppose that a vector space V has two subspaces U and W such that any vector $\mathbf{v} \in V$ can be written in a unique way as the sum $\mathbf{u} = \mathbf{v} + \mathbf{w}$ where $\mathbf{u} \in U$ and $\mathbf{w} \in W$. The subspace V is, in this case, the *direct sum* of the vector spaces U and W, which is denoted by $V = U \oplus W$. One has $\dim(V) = \dim(U) + \dim(W)$.
- Scalar product and norm: New operations have often to be added to the bare structure of a vector space.
  The *scalar product* is a form $(\mathbf{v}, \mathbf{w}) \to \mathbb{C}$ with the following properties:

- Linear in the first variable, i.e., $(a\mathbf{u} + b\mathbf{v}, \mathbf{w}) = a(\mathbf{u}, \mathbf{w}) + b(\mathbf{v}, \mathbf{w})$.
- $(\mathbf{v}, \mathbf{w}) = \overline{(\mathbf{w}, \mathbf{v})}$
- positive, i.e., $(\mathbf{v}, \mathbf{v}) \geq 0$ and the equality arises only when $\mathbf{v} = \mathbf{0}$.

In the case of a real vector space, this form is simply bilinear, but in the complex case, it is *sesqui-linear*, i.e., semi-linear in the second argument:$(\mathbf{v}, a\mathbf{w}) = \bar{a}(\mathbf{v}, \mathbf{w})$. A vector space with a scalar product is called a *pre-Hilbert space*.

The *norm* $\|\mathbf{v}\|$ is the (positive) square root of the scalar product of a vector $\mathbf{v}$ with itself: $\|\mathbf{v}\|^2 = (\mathbf{v}, \mathbf{v})$.

In the case of the canonical example, the canonical scalar product is given by

$$((b^1, \cdots, b^n), (c^1, \cdots, c^n)) = b^1\overline{c^1} + \cdots + b^n\overline{c^n}.$$

Two norms $\|.\|_1$ and $\|.\|_2$ are *equivalent* if there exist two real strictly positive constants c and C such that $c\|\mathbf{v}\|_1 \leq \|\mathbf{v}\|_2 \leq C\|\mathbf{v}\|_1$.

A vector that has a norm equal to 1 is called a unit vector or a normalized vector. Two vectors are *orthogonal* if their scalar product is equal to 0.

If V is a vector space with a norm $\|.\|_V$, this defines a norm on the dual space V^* by $\|f\|_{V^*} = \frac{\sup_{\mathbf{v}\in V} f(\mathbf{v})}{\|\mathbf{v}\|_V} = \sup_{\|\mathbf{v}\|_V=1} f(\mathbf{v})$ for any linear form $f \in V^*$.

We use the name of the concerned vector space as a subscript to indicate the vector product or the norm when this is necessary, but it will be omitted if there is no ambiguity.

- The norm determines a *topology*, i.e., a system of subsets called open sets, on the vector space (which is, therefore, called a topological vector space), which is necessary to introduce the concepts of convergence and limit. Two equivalent norms determine the same topology.

  The distance between two vectors is defined by $\rho(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$ and the (open) ball of center $\mathbf{u}$ and radius $\epsilon \in \mathbb{R}$ is the set of vectors $\mathbf{v} \in B(\mathbf{u}, \epsilon)$ so that $\rho(\mathbf{u}, \mathbf{v}) < \epsilon$. A set Ω is called *open* if for any $\mathbf{u} \in \Omega$, there exists an ϵ such that the ball $B(\mathbf{u}, \epsilon)$ is contained in Ω. A set E is a *closed* if its complement (i.e., elements that are not in E) is open. The *closure* $\overline{\Omega}$ of a

set is the smallest closed set that contains it and the *interior* $Int(\Omega)$ is the largest open set contained in it.[a] The *boundary* $\partial\Omega$ is the set of elements that are in the closure and not in the interior.

A vector space with a norm that is complete (i.e., every *Cauchy sequence* of the space $\{\mathbf{u}_n\}$, i.e., such that $\rho(\mathbf{u}_n, \mathbf{u}_m) \to 0$ as $n, m \to +\infty$, has its limit in the space) is called a *Banach space*.

- Quotient: If W is a vector subspace of the vector space V, equivalence classes in V can be defined in the following way: Two elements of V are equivalent if they only differ by a vector of W. The set of equivalence classes is itself a vector space called the *quotient space* V/W and one has $V = W \oplus (V/W)$. If a norm $\|.\|_V$ is defined on the vector space V, a norm on the quotient space may be defined by $\|\mathbf{u}\|_{V/W} = \inf_{\mathbf{v}\in\mathbf{u}} \|\mathbf{v}\|_V$ for any equivalence class $\mathbf{u}$ of V/W.
- It is usual to build integration upon *Lebesgue measure theory*, which is quite technical and mathematically demanding. We just recall here that a fundamental concept is the one of *null measure set*. For all practical purpose, null measure sets on $\mathbb{R}$ are finite and denumerable sets of points (although there are non-denumerable null measure sets), e.g., the set of rational numbers, and null measure sets on $\mathbb{R}^n$ are denumerable sets of subsets of $\mathbb{R}^p$ with $p < n$. Two functions are equal *almost everywhere (a.e.)* if they only differ on a null measure set.
  For instance, $\mathcal{L}^p(\Omega)$ is the set of functions f such that $\int_\Omega |f(\mathbf{x})|^p d\mathbf{x} < \infty$ where $d\mathbf{x}$ denotes the Lebesgue measure on Ω. The sets of functions that are equal a.e. are equivalent classes. The set $\mathcal{N}(\Omega)$ of functions equal to zero a.e. is a subspace of $\mathcal{L}^p(\Omega)$ and one defines $L^p(\Omega) = \mathcal{L}^p(\Omega)/\mathcal{N}(\Omega)$.
  L^p spaces are Banach spaces with the norm

$$\|f\|_{L^p} = \left(\int_\Omega |f(\mathbf{x})|^p d\mathbf{x}\right)^{1/p}.$$

  $L^\infty(\Omega)$ is the Banach space of functions bounded a.e. on Ω, which has been given the norm $\|f\|_{L^\infty} = \text{ess sup}_\Omega |f(\mathbf{x})|$

[a] Note that if Ω is an open set, we have $\text{Int}(\Omega) = \Omega$.

where "ess" stands for "essentially," i.e., the norm is the smallest value M for which $|f(\mathbf{x})| \leq M$ a.e.
If Ω is the whole $\mathbb{R}^n$, the integral may be defined only locally, i.e., on any compact subset. $L^p_{\text{loc}}(\mathbb{R}^n)$ is the space of functions f such that $f\varphi \in L^p(\mathbb{R}^n)$ for any $\varphi \in C_0^\infty(\mathbb{R}^n)$.
One of the main results of the integration theory is the *Lebesgue dominated convergence theorem*: If $\{f_n\}$ is a sequence of functions in $L^1(\Omega)$ and if there exists a function $g \in L^1(\Omega)$ so that $|f_n(\mathbf{x})| \leq g(\mathbf{x})$ a.e. Then, if this sequence is convergent a.e. to a function f, this function is also integrable and $\int_\Omega f_n(\mathbf{x})d\mathbf{x} \to \int_\Omega f(\mathbf{x})d\mathbf{x}$. This result is sometimes called the *Lebesgue bounded convergence theorem* if one considers a constant $M > 0$ instead of the function g.
As an easy corollary, series and integral can be swapped with confidence provided the following holds: Let us consider that for every integer N, the previous sequence $f_n \in L^1(\Omega)$ is such that $|\sum_{n=1}^N f_n(\mathbf{x})| \leq g(\mathbf{x})$ a.e. Then, if $\sum_{n=1}^\infty f_n(\mathbf{x})$ is convergent a.e. to a sum function $f(\mathbf{x})$, this function is also integrable, the series $\sum_{n=1}^\infty \int_\Omega f_n(\mathbf{x})d\mathbf{x}$ is convergent, and $\int_\Omega f(\mathbf{x})d\mathbf{x} = \int_\Omega \sum_{n=1}^\infty f_n(\mathbf{x})d\mathbf{x} = \sum_{n=1}^\infty \int_\Omega f_n(\mathbf{x})d\mathbf{x}$.

- *Hilbert spaces*: Important examples of infinite dimensional vector spaces are the Hilbert spaces. A Hilbert space is a complete space with a scalar product (i.e., a pre-Hilbert Banach space). For instance, given a domain Ω of $\mathbb{R}^n$, the space of square integrable functions $L^2(\Omega)$ (introduced in the previous section) with the scalar product

$$(f, g) = \int_\Omega f(\mathbf{x})\overline{g(\mathbf{x})}d\mathbf{x} \text{ is a Hilbert space.}$$

One of the important properties of the Hilbert spaces is the *Riesz representation theorem*: Given any linear form f on a Hilbert space H, there exists one and only one vector $\mathbf{u}$ such that for all vectors $\mathbf{v} \in H$, one has $f(\mathbf{v}) = (\mathbf{u}, \mathbf{v})$. Hilbert spaces are, therefore, *reflexive spaces*, i.e., they are isomorphic to their dual spaces. A common abuse of notation is to identify a Hilbert space with its dual and to write $(\mathbf{u}, \mathbf{v}) = <\mathbf{u}, \mathbf{v}>$, where $\mathbf{v}$ denotes both the element of the Hilbert

space in the scalar product and its corresponding linear form by the Riez theorem in the duality product.

Via the scalar product, the notion of orthogonality is available in Hilbert spaces. Given a Hilbert space H and V a Hilbert subspace, i.e. a subspace which is also a Hilbert space, the set $V^\perp$ of elements of H orthogonal to all the elements of V is also a Hilbert subspace of H and one has $H = V \oplus V^\perp$.

- Operators and functionals: Various operations may be defined on vector spaces. Mappings from a vector space onto a vector space are usually called *operators*, while mappings from a vector space onto scalars are called *functionals*. Given two Banach spaces V and W, $\mathcal{L}(V, W)$ is the set of *linear operators* from V to W, i.e., operators that preserve linear combinations. If $L \in \mathcal{L}(V, W)$, for all $\mathbf{v}^1, \mathbf{v}^2 \in V$ and $a, b \in \mathbb{C}$, we have $L(a\mathbf{v}^1 + b\mathbf{v}^2) = aL(\mathbf{v}^1) + bL(\mathbf{v}^2)$. The set $\mathcal{L}(V, V)$ is written $\mathcal{L}(V)$. The space V is called the *domain* $\mathrm{dom}(L)$ of the operator L, and the vector space spanned by the elements of W, which can be obtained by the action of the operator on an element of V, is called the *range*, *image*, or *codomain* $\mathrm{cod}(L)$ of L. The *kernel* or *nullspace* $\ker(L)$ is the subspace $\ker(L) = \{\mathbf{v} \in V = \mathrm{dom}(L),\ L(\mathbf{v}) = \mathbf{0} \in W\}$.

  A linear operator L is a *bounded* if there is a constant C such that $\|L(\mathbf{v})\|_W \leq C\|\mathbf{v}\|_V$ for all $\mathbf{v} \in V$. A linear operator is continuous (i.e., if $\mathbf{v}^n \to \mathbf{v}$, then $L(\mathbf{v}^n) \to L(\mathbf{v})$) if and only if it is bounded.

  The subset $\mathcal{B}(V, W)$ of bounded linear operators is a normed linear space with the norm $\|L\|_{\mathcal{B}} = \sup_{\|\mathbf{v}\|_V = 1} \|L(\mathbf{v})\|_W$.

  A bounded operator is *compact* if the image of every bounded sequence contains a convergent sub-sequence. In particular, the identity map I such that $I(\mathbf{v}) = \mathbf{v}$ is not compact on an infinite dimensional space.

  If the image of a linear operator has a finite dimension, the operator is a *finite rank operator*. All the finite rank operators are compact.

  Linear operators in $\mathcal{L}(V, W)$ can be combined linearly by just taking the linear combination of their action on the resulting vectors, which gives a structure of vector space to $\mathcal{L}(V, W)$

itself. The *composite* of two operators L, M in $\mathcal{L}(V)$ is given by their successive applications: $L \circ M : \mathbf{v} \to L(M(\mathbf{v}))$. This can be viewed as a product that makes $\mathcal{L}(V)$ an algebra.
In functional spaces, important linear operators are obtained by combinations of multiplications by functions and partial derivatives. Parenthesis are often omitted when expressing the action of an operator and one writes, for instance, $L\mathbf{v}$ instead of $L(\mathbf{v})$.

- A *matrix* is a linear operator A between finite dimensional vector spaces, which can be represented by a rectangular array of scalars $[a_{ij}]$ (called the elements of the matrix) so that $(A\mathbf{v})_i = \sum_j a_{ij} v_j$. Note that in the context of matrix algebra, one is often not very careful about upper and lower indices. The multiplication by a scalar and the addition of matrices are obvious from the vector space structure of linear operators. The composite of two matrices is given by their successive applications: $C\mathbf{v} = AB\mathbf{v}$ and the corresponding array for C is given by $[c_{ij}] = [\sum_k a_{ik} b_{kj}]$ called the *matrix product*, which is not commutative. If the domain and image vector spaces have the same dimension, the matrix is a *square matrix*. A square matrix I such that $AI = IA = A$ for all the square matrices A of the same dimensions is a *unit matrix*. The *trace* $Tr(A)$ of a square matrix $A = [a_{ij}]$ is the sum of its diagonal elements: $Tr(A) = \sum_i a_{ii}$. If the columns of elements of a square matrix A are considered a set of n vectors of dimension n (they are, in fact, the $A\varepsilon_i$ vectors), a scalar can be built from the matrix via a n-linear form. One considers the totally skew-symmetric n-linear form,[b] which gives 1. when it is fed with the columns of the unit matrix, i.e., the ε_i. The scalar resulting from the action of this form on the columns of the matrix A is called the *determinant* $det(A)$. If the determinant of a matrix A is not equal to zero, the matrix is *regular* and there exists a matrix A^{-1} called the *inverse matrix* of A such that $A^{-1}A = AA^{-1} = I$ and else the matrix is *singular*.

[b] See the section below on n-covectors for further details.

A vector $\mathbf{v}$ can be considered a single column matrix and the action of a matrix on such a vector is merely a matrix product.

- *Fourier transformation*: The Fourier transformation $\mathcal{F}$ is an extremely important linear operator, which can be defined on $L^2(\mathbb{R}^n)$.
  Given a function $f \in L^2(\mathbb{R}^n) : \mathbf{x} \in \mathbb{R}^n \to \mathbb{R}$ and if the duality product on $\mathbb{R}^n$ is denoted by
  $$< \mathbf{k}, \mathbf{x} >= k_i x^i,$$
  the Fourier transform $\mathcal{F}[f] = \hat{f}$ is given by
  $$\hat{f}(\mathbf{k}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{-i<\mathbf{k},\mathbf{x}>} f(\mathbf{x}) d\mathbf{x}.$$
  It is not obvious that the previous integrals exist for all functions in $L^2(\mathbb{R}^n)$. Technically, one has to start with a space where the existence is obvious, $L^1(\mathbb{R}^n)$ or $C_0^\infty(\mathbb{R}^n)$, and then one extends the operator to $L^2(\mathbb{R}^n)$.
  The inverse transform is given by
  $$f(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}} \int_{\mathbb{R}^n} e^{i<\mathbf{k},\mathbf{x}>} \hat{f}(\mathbf{k}) d\mathbf{k}.$$
  One of the fundamental properties of the Fourier transformation is that it makes differentiation algebraic:
  $$\widehat{D^\alpha u}(\mathbf{k}) = \mathbf{k}^\alpha \hat{u}(\mathbf{k}) \text{ and } \widehat{\mathbf{x}^\alpha u}(\mathbf{k}) = D^\alpha \hat{u}(\mathbf{k})$$
  where $\mathbf{x}^\alpha = (x^1)^{\alpha_1} \cdots (x^n)^{\alpha_n}$ and $\mathbf{k}^\alpha = k_1^{\alpha_1} \cdots k_n^{\alpha_n}$.
  A fundamental property is the *Parseval–Plancherel theorem*: The Fourier transform is an isometry in the L^2 norm, i.e., it conserves the scalar product and the norm:
  $$(f, g)_{L^2} = (\mathcal{F}[f], \mathcal{F}[g])_{L^2}.$$
- *Convolution*: Another useful operation is the convolution. Given two functions $f, g \in L^2(\mathbb{R}^n)$, the convolution product is defined by
  $$f \star g(\mathbf{x}) = \int_{\mathbb{R}^n} f(\mathbf{y}) g(\mathbf{x} - \mathbf{y}) d\mathbf{y} =< f(\mathbf{y}), g(\mathbf{x} - \mathbf{y}) >_{\mathbf{y}}$$
  where the subscript $\mathbf{y}$ indicates that the integration involved in the duality product is performed along this variable.
  Surprisingly, this operation is commutative and associative: $f \star g = g \star f$ and $(f \star g) \star h = f \star (g \star h) = f \star g \star h$.

The fundamental result relating the Fourier transform and the convolution is the *Faltung theorem*: If f and g are two functions having Fourier transform and such that the convolution exists and is integrable: $\mathcal{F}[f \star g] = \mathcal{F}[f]\mathcal{F}[g]$.

- *Distribution*: The duality in infinite dimensional function space leads to a fundamental tool of mathematical physics: distributions or generalized functions. The space of test functions $\mathcal{D}$ is defined as the set of functions infinitely differentiable on $\mathbb{R}^n$ and with a bounded support, i.e., $C_0^\infty(\mathbb{R}^n)$. Those functions are gentle enough to be integrated on their whole domain and differentiated everywhere as many times you like. The *dual topological space* $\mathcal{D}'$ of continuous (this requirement makes the difference with the algebraic dual), linear forms on $\mathcal{D}$ is the space of distributions. A fundamental example is the Dirac delta distribution (often improperly called Dirac delta function in physics) δ, which associates to a test function $\phi \in \mathcal{D}$ its value at 0: $\delta(\phi) =< \delta, \phi >= \phi(0)$. The space of distributions is larger than the space of test functions and in a sense, contains it. On one hand, with any test function ϕ is associated a distribution D_ϕ such that the action of this distribution on another test function χ is given by $< D_\phi, \chi >= \int_{\mathbb{R}^n} \phi(\mathbf{x})\chi(\mathbf{x})\, d\mathbf{x}$, and on the other hand, there is no test function associated with the Dirac distribution.
  Another important distribution is $\mathrm{vp}\{1/\|\mathbf{x}\|\}$ (where vp stands for the *Cauchy principal value*) defined by

$$< \mathrm{vp}\{1/\|\mathbf{x}\|\}, \varphi >= \lim_{\epsilon \to 0} \int_{\mathbb{R}^n - \{\|\mathbf{x}\| < \epsilon\}} \varphi(\mathbf{x})/\|\mathbf{x}\|\, d\mathbf{x}.$$

  Nevertheless, there is a common and useful abuse of notation in physics, which writes the action of the Dirac delta distribution and other singular distributions (those which do not correspond to a test function) using the integral symbol $< \delta, \phi >= \int_{\mathbb{R}^n} \delta(\mathbf{x})\phi(\mathbf{x})d\mathbf{x} = \phi(0)$. In the same spirit, one writes $< \delta_\mathbf{y}, \phi >= \int_{\mathbb{R}^n} \delta(\mathbf{x} - \mathbf{y})\phi(\mathbf{x})d\mathbf{x} = \phi(\mathbf{y})$.
  We will often use the duality product notation to write the integral of the product of two functions f, g on their common domain of definition Ω: $< f, g >= \int_\Omega fg\, d\mathbf{x}$ in the real case and $< f, g >= \int_\Omega f\overline{g}\, d\mathbf{x}$ in the complex case.

Note that this corresponds also to the scalar product on L^2: $(f, g)_{L^2} =< f, g >$.[c]

- Operations such as derivation, Fourier transformation, and convolution can be applied to distributions. As far as the Fourier transformation is concerned, not all the distributions have a transformation and one has to reduce the space of distributions by increasing the space of test functions (see here how the duality plays). The set of rapidly decreasing functions, $\mathcal{S}$, is introduced, i.e., functions that decrease more rapidly to zero than any power of $\|\mathbf{x}\|$ when $\|\mathbf{x}\| \to \pm\infty$. The set of tempered distribution, $\mathcal{S}'$, is the topological dual space of $\mathcal{S}$, i.e., the set of continuous linear functionals on $\mathcal{S}$ and any tempered distribution admit a Fourier transform. Operators on distributions are described by giving the resulting distribution, but this one is itself described through its action on test functions φ. Therefore, one has the following definitions,[d] given only in the case of test functions and distributions on $\mathbb{R}$ for the sake of simplicity:
  - Derivative of a distribution: $< \frac{dD}{dx}, \varphi >= - < D, \frac{d\varphi}{dx} >$.
    One often introduces the notation $\{dD(x)/dx\}$ to mean that the derivation is taken in the sense of the functions and not of the distributions. For instance, if a function $f : \mathbb{R} \to \mathbb{R}$ is C^∞ except in a single point $x = a$ where there is a discontinuity
    $$\lim_{\epsilon \to 0^+} f(a+\epsilon) - \lim_{\epsilon \to 0^-} f(a+\epsilon) = disc_f(a)$$
    (also denoted $[f]_a$), one has $\frac{df}{dx} = \{\frac{df}{dx}\} + disc_f(a)\ \delta(x-a)$.
  - Multiplication of a distribution by a function:
    $$< fD, \varphi >=< D,\ f\varphi > .$$
    Yes, this seemingly trivial operation needs a definition! You can always multiply a distribution by a test function, but you cannot, for instance, multiply a function with a discontinuity at the origin with the Dirac distribution. An

[c] See the remark above on the abuse of notation due to the reflexive nature of the Hilbert spaces.

[d] Which are chosen to match the definitions for functions when the distribution can be associated with a function.

important negative property of distributions is that the product of two distributions does not always exist.

- Fourier transform of a tempered distribution:

$$< \mathcal{F}[D], \varphi >=< D, \mathcal{F}[\varphi] > .$$

For instance, we have $\mathcal{F}[\delta(x)] = \chi_{\mathbb{R}}(k)$ (where χ_Ω is the characteristic function of the set Ω: $\chi_\Omega(x) = 1$ if $x \in \Omega$ and $\chi_\Omega(x) = 0$ if $x \notin \Omega$) and $\mathcal{F}[\mathrm{sgn(x)}] = \frac{i}{\pi}\,\mathrm{pv}\left(\frac{1}{k}\right)$ where $\mathrm{sgn}(x) = \frac{x}{|x|}$.

- Convolution of two distributions:

$$< D(x) \star E(x), \varphi(x) >=< D \otimes E(x, y), \varphi(x + y) >,$$

where the right-hand duality product takes place in $\mathbb{R}^2$ with x and y variables.

Another interesting example of tempered distribution is the *Dirac comb* $Ш(x) = \sum_{n\in\mathbb{Z}} \delta(x - n)$, which is its own Fourier transform $\mathcal{F}[Ш(x)] = Ш(k) = \sum_{n\in\mathbb{Z}} \delta(k - n)$. Introducing this distribution in the Parseval–Plancherel theorem (extended to the duality product between $\mathcal{S}$ and $\mathcal{S}'$) together with a function $\varphi \in \mathcal{S}$ gives the *Poisson summation formula* $\sum_{n\in\mathbb{Z}} \varphi(n) = \sum_{n\in\mathbb{Z}} \widehat{\varphi}(n)$, a useful trick for convergence acceleration.

- *Sobolev spaces*: These are the Banach spaces defined for integers m, p by $W^{m,p}(\Omega) = \{u : u \in L^p(\Omega),$
$D^\alpha u \in L^p(\Omega)$ for $|\alpha| \le m\}$.
The spaces $W^{m,2}(\Omega) = H^m(\Omega)$ are of particular interest because they are Hilbert spaces with the scalar product defined by $(u, v)_{H^m} = \sum_{|\alpha|\le m}(D^\alpha u, D^\alpha v)_{L^2}$ and the corresponding norm $\|u\|^2_{H^m} = (u, u)_{H^m}$.
It can be shown that $u \in H^m(\mathbb{R}^n)$ if and only if

$$(1 + k^2)^{m/2}\widehat{u}(\mathbf{k}) \in L^2(\mathbb{R}^n)$$

and that the norm

$$\|u\|'_{H^m} = \left(\int_{\mathbb{R}^n} (1 + k^2)^{m/2}\widehat{u}(\mathbf{k})\, d\mathbf{k}\right)^{1/2}$$

with $k^2 = \sum_i (k_i)^2$ is equivalent to the first defined norm.
The advantage of this definition is that it can be extended to any real value s and, therefore, allows the definition of

negative and/or fractional index spaces H^s. Note that in the case of negative indices, the elements of H^s are not all in $L^2 = H^0$ and these spaces are rather distribution spaces than genuine function spaces.
The *Rellich–Kondrachov theorem* states that we have the embeddings $H^s(\mathbb{R}^n) \subset H^t(\mathbb{R}^n)$ if $s > t$ and that the corresponding inclusion maps are compact .
The *Sobolev lemma* states that

$$H^s(\mathbb{R}^n) \subset C^k(\mathbb{R}^n) \text{ if } s > k + n/2.$$

$H^{-s}(\mathbb{R}^n)$ is the topological dual space of $H^s(\mathbb{R}^n)$ according to the classical duality pairing corresponding to the L^2 scalar product.
Given Ω, the space $H_0^s(\Omega)$ is the closure of $C_0^\infty(\Omega) = \mathcal{D}(\Omega)$ in $H^s(\Omega)$ (i.e., every element in $H_0^s(\Omega)$ is the limit according to the norm of $H^s(\Omega)$ of a sequence of functions in $C_0^\infty(\Omega)$). If $\partial\Omega$ is "regular enough," $H_0^s(\Omega)$ is the subspace of elements of $H^s(\Omega)$ equal to zero on $\partial\Omega$. In the case $\Omega = \mathbb{R}^n$, $H_0^s(\Omega)$ is the same as $H^s(\Omega)$, but it is a strict subset in the other cases. $H^{-s}(\Omega)$ is the topological dual of $H_0^s(\Omega)$.
One has the following situation: $H_0^s(\Omega) \subset L^2(\Omega) \subset H^{-s}(\Omega)$ called a *Gelfand triplet* or a *rigged Hilbert space*. The space L^2 is called the *pivot space* and its scalar product provides the duality pairing. $H^{-s}(\Omega)$ can be, without contradiction, "larger" than $H_0^s(\Omega)$ (it is a distribution space) and "equal" since it is isomorphic according to the Riesz theorem.
Another fundamental example of rigged Hilbert space is $\mathcal{D}(\Omega) \subset L^2(\Omega) \subset \mathcal{D}'(\Omega)$ where the "rigging" spaces are not Hilbert ones.

- *Trace theorems* answer the question of the regularity and restriction of functions (and distributions) on the boundary of a domain or another lower dimensional sub-domain. Consider Ω an open domain of $\mathbb{R}^n$ and a piecewise smooth hyper-surface Γ of co-dimension 1 (i.e., of dimension $n-1$) contained in $\overline{\Omega}$ (and in particular which may coincide with the boundary $\partial\Omega$). The *restriction operator* γ is introduced by $\gamma : C^\infty(\overline{\Omega}) \to C^\infty(\Gamma)$, $\gamma(\mathbf{u}) = \mathbf{u}|_\Gamma$. This operator can be extended to a continuous operator on some Sobolev spaces:

If the hyper-surface Γ is either compact or a portion of a hyper-plane, for $s > 1/2$ the operator γ can be extended to a continuous operator $\gamma : H^s(\Omega) \to H^{s-1/2}(\Gamma)$.
Considering the case $s = 1$ as an example, an alternative definition of the space $H^{1/2}(\Gamma)$ can be defined as the quotient space $H^1(\Omega)/H_0^1(\Omega)$ with the quotient norm $\|\mathbf{u}\|_{H^{1/2}(\Gamma)} = \inf_{\{\mathbf{v}\in H^1(\Omega),\gamma(\mathbf{v})=\mathbf{u}\}} \|\mathbf{v}\|_{H^1(\Omega)}$. Accordingly, one has $H^1(\Omega) = H_0^1(\Omega) \oplus H^{1/2}(\Gamma)$ and the trace inequality $\|\mathbf{u}|_\Gamma\|_{H^{1/2}(\Gamma)} \leq C\|\mathbf{u}\|_{H^1(\Omega)}$.

- Given a linear operator $L \in \mathcal{L}(V, V)$, the *adjoint operator* is $L^* \in \mathcal{L}(V^*, V^*)$ such that

$$< L\mathbf{u}, \mathbf{v} >=< \mathbf{u}, L^*\mathbf{v} >, \forall \mathbf{u} \in V, \forall \mathbf{v} \in V^*.$$

In the functional case, given a differential operator L: $C^\infty(\Omega) \to C^\infty(\Omega)$ of order m defined by $L = \sum_{|\alpha|\leq m} a_\alpha(x) D^\alpha$, the *formal adjoint* L^* is such that

$$< L\varphi, \chi >=< \varphi, L^*\chi >, \forall \varphi, \forall \chi \in C_0^\infty(\Omega).$$

Integration by parts shows that $L^*\varphi = \sum_{|\alpha|\leq m} (-1)^{|\alpha|} D^\alpha(a_\alpha \varphi)$.
For real matrices, the adjoint of a matrix $A = [a_{ij}]$ is called the *transposed matrix* $A^T = [a_{ji}]$ where lines are written as columns. For complex matrices, the adjoint is the complex conjugate transposed or Hermitian transposed $A^H = [\overline{a_{ji}}]$. A matrix A is *symmetric* if $A = A^T$ and it is (symmetric) *definite positive* if $(\mathbf{v}, A\mathbf{v}) > 0$, $\forall \mathbf{v} \neq \mathbf{0}$. A matrix A is *Hermitian* if $A = A^H$. The transposed $\mathbf{v}^T$ of a column vector $\mathbf{v}$ is a row vector, i.e., a single row matrix. The matrix product $\mathbf{v}^T\mathbf{w}$ is the canonical scalar product of real finite dimensional vectors. The matrix product $\mathbf{v}\mathbf{w}^T = [v_i w_j]$ resulting in a general matrix is called the *dyadic product* of the two vectors.
An operator L is *self-adjoint* if $L = L^*$. In the case of unbounded operators, physicists often confuse self-adjoint operators with merely Hermitian (in the complex case) or symmetric (in the real case) operators or even with formally self-adjoint operators and so we do in this book. An *extension* B of an operator A of domain $\mathrm{dom}(A)$ is an operator with a domain $\mathrm{dom}(B) \supset \mathrm{dom}(A)$ such that $B|_{\mathrm{dom}(A)} = A$

and this situation is denoted by $B \supset A$. An operator is self-adjoint if $A^* = A$, but the most common situation is $A^* \supset A$, which corresponds to symmetric and Hermitian operators. We will disregard this distinction except in the following simple example. Consider A_0 an operator acting on $H_0^1([a,b])$ (i.e., the set of square integrable functions φ defined on the interval $[a,b]$, with a square integrable derivative, and verifying the boundary conditions $\varphi(a) = \varphi(b) = 0$), defined by $A_0\varphi = i\frac{d\varphi}{dx}$. This operator is Hermitian:

$$< A_0\varphi,\, \chi >= \int_a^b i\frac{d\varphi}{dx}\overline{\chi}dx = \int_a^b i\overline{\frac{d\chi}{dx}}dx + i[\varphi\overline{\chi}]_a^b =< \varphi,\, A_0^*\chi >$$

with $A_0^*\chi = \frac{d\chi}{dx}$ for any function $\chi \in H^1([a,b])$, which is obtained because of the boundary conditions on φ. There are no boundary conditions needed on $\chi \in H^1([a,b]) \supset H_0^1([a,b])$ and obviously $A_0^* \supset A_0$. Given $\theta \in [0, 2\pi[$, the operator $A_\theta\varphi = i\frac{d\varphi}{dx}$ with the domain $\mathrm{dom}(A_\theta) = H_\theta^1([a,b]) = \{\varphi \in H^1([a,b]), \varphi(a) = e^{i\theta}\varphi(b)\}$ is now considered, the boundary term becomes

$$i(\varphi(b)\overline{\chi(b)} - \varphi(a)\overline{\chi(a)}) = i\varphi(b)(\overline{\chi(b)} - \overline{e^{-i\theta}\chi(a)}).$$

This term vanishes if and only if $\chi(a) = e^{i\theta}\chi(b)$, i.e., $\chi \in H_\theta^1([a,b])$ and $A_\theta^* = A_\theta$ is a self-adjoint operator. This is, in fact, a one-parameter family of self-adjoint operators with $A_0 \supset A_\theta \supset A_0^*$ but with $A_\theta \neq A_{\theta'}$ for $\theta \neq \theta'$.
Choosing $\chi \in C_0^\infty$ in the duality product avoids any question about the boundary condition and leads to the careless concept of formally self-adjoint operator A such that

$$< A\varphi,\, \chi >=< \varphi,\, A\chi > \text{ for any } \chi \in C_0^\infty.$$

One says that $Lf = g$ *weakly* for $f, g \in L^1_{\mathrm{loc}}$ if

$$< g, \varphi >=< f, L^*\varphi > , \forall\varphi \in C_0^\infty(\Omega).$$

- The *Green's function* of an operator L with constant coefficients is a distribution $G(\mathbf{x}-\mathbf{y}) \in \mathcal{D}' \otimes \mathcal{D}'$ such that $L^*_{\mathbf{x}}G(\mathbf{x}-\mathbf{y}) = \delta(\mathbf{x}-\mathbf{y})$ where the fact that the operator acts on the $\mathbf{x}$ variables is emphasized. One also has to request that $G|_{\partial\Omega} = 0$ or to consider the free space Green's

function, but we stay on a rather formal level here and let such considerations on side. The Green's function is usually used to (formally) invert the differential operator. Consider once again $Lf = g$. By means of the fact that the Dirac distribution is the neutral element of the convolution, one has $f(\mathbf{x}) = f \star \delta =< f(\mathbf{y}), \delta(\mathbf{x}-\mathbf{y}) >_{\mathbf{y}} =< f(\mathbf{y}), L^* G(\mathbf{x}-\mathbf{y}) >_{\mathbf{y}} =< Lf(\mathbf{y}), G(\mathbf{x}-\mathbf{y}) >_{\mathbf{y}} =< g(\mathbf{y}), G(\mathbf{x}-\mathbf{y}) >_{\mathbf{y}} = g \star G$. The formula $f = g \star G$ is only valid in free space, i.e., if the domain is $\mathbb{R}^n$. In the case of a bounded domain, boundary conditions have to be taken into account as it will be explained later.

- Up to now, the geometrical domains were open sets of $\mathbb{R}^n$. A serious treatment of the geometrical framework of physics requires the concept of *manifold*. A *manifold* M is a set of points that is locally homeomorphic to $\mathbb{R}^n$ in the sense that any neighborhood of a point can be continuously mapped on an open set of $\mathbb{R}^n$ (n is the same for all the points of M and is called the *dimension* of the manifold). With such a mapping, the points in the neighborhood can be distinguished by an ordered set of n real numbers called the *(local) coordinates*. Nevertheless, we cannot hope to be always able to find a set of coordinates that covers the whole manifold at once. Therefore, it is allowed to cover the manifold with several overlapping open sets each endowed with a particular coordinate system. The regularity of the manifold is given by the regularity of the so-called *transition functions*: Considering two coordinate systems defined on overlapping opens sets U and $\tilde{U}$ by the mappings $\varphi : U \to \mathbb{R}^n$ and $\tilde{\varphi} : \tilde{U} \to \mathbb{R}^n$, the invertibility of the continuous coordinate mappings induces a transition mapping $\tilde{\varphi}\varphi^{-1} : \varphi(U \cap \tilde{U}) \subset \mathbb{R}^n \to \tilde{\varphi}(U \cap \tilde{U}) \subset \mathbb{R}^n$. The regularity of the manifold is the one of the transition functions: A differentiable C^m manifold is such that all the transition functions are C^m.

  The $\mathbb{R}^n$ are trivial examples of manifolds, and in fact the only manifold used here is $\mathbb{R}^3$. In this case, why bother about manifolds? Because $\mathbb{R}^n$ can also be considered a vector space, for instance. On the one hand, in "vector space $\mathbb{R}^n$," two elements can be added but not in "manifold $\mathbb{R}^n$" and

on the other hand, changing all the n-tuples $\{x^1, \cdots, x^n\}$ in "manifold $\mathbb{R}^n$" to n-tuples $\{y^1 = (x^1)^3, \cdots, y^n = (x^n)^3\}$ is a valid global change of coordinates leaving the manifold unchanged but makes no sense in "vector space $\mathbb{R}^n$"!

- The position of a point in a manifold M of dimension n is given by an ordered set of n numbers $(x^1, \cdots, x^n)$ called the coordinates. Each coordinate can also be viewed as a function on the manifold.

  A *curve* γ is an application from an interval of $\mathbb{R}$ on the manifold M: $\mathbf{r}(t) = (x^1(t), \cdots, x^n(t))$ where t is the parameter. If $f(x^1, \cdots, x^n)$ is a scalar function on the space, the composition of this function with the curve gives a function from $\mathbb{R}$ to $\mathbb{R}$: $f(x^1(t), \cdots, x^n(t))$. The derivation with respect to t of this function gives, applying the chain rule: $\frac{d}{dt} f(x^1(t), \cdots, x^n(t)) = \frac{\partial f}{\partial x^1}\frac{dx^1}{dt} + \cdots + \frac{\partial f}{\partial x^n}\frac{dx^n}{dt}$. This expression can be viewed as the duality product $< df, \mathbf{v}_\gamma >$ of the covector $df = \frac{\partial f}{\partial x^1}dx^1 + \cdots + \frac{\partial f}{\partial x^n}dx^n = \frac{\partial f}{\partial x^i}dx^i$ with the vector $\mathbf{v}_\gamma = \frac{dx^1}{dt}\frac{\partial .}{\partial x^1} + \cdots + \frac{dx^n}{dt}\frac{\partial .}{\partial x^n} = \frac{dx^i}{dt}\frac{\partial .}{\partial x^i}$. The vector $\mathbf{v}_\gamma$ is the tangent to the curve given in the form of a first-order linear differential operator and the covector df is the *differential* of the function. One takes the general definitions:

  - A *vector* (in the geometric sense) at a point of the manifold is a first-order linear differential operator on the functions on the manifold. In a coordinate system $(x^1, \cdots, x^n)$, a vector $\mathbf{v}$ at a point of coordinates $(p^1, \cdots, p^n)$ is represented by a set of n numbers $(v^1, \cdots, v^n)$, the *components* of the vector, so that the result of the action of this vector on a function $f(x^1, \cdots, x^n)$ is the scalar: $\mathbf{v}(f) = v^1\frac{\partial f}{\partial x^1} + \cdots + v^n\frac{\partial f}{\partial x^n}|_{(x^1=p^1,\cdots,x^n=p^n)}$.
  - In a coordinate system, a basis for the vectors are the partial derivatives with respect to the coordinates, so that a vector can be written $\mathbf{v} = v^1\frac{\partial}{\partial x^1} + \cdots + v^n\frac{\partial}{\partial x^n}$.
  - A *covector* at a point of the space is a linear form on the vectors at this point.
  - In a coordinate system, a basis for the covectors is the differential of the coordinates so that a covector can be written $\alpha = \alpha_1 dx^1 + \cdots + \alpha_n dx^n$. They form a basis dual

to the partial derivatives: $< dx^i, \frac{\partial .}{\partial x^j} >= \delta^i_j$ and, therefore, $< \mathbf{v}, \alpha >= \alpha_i v^i$.

- A *(co)vector field* is a set of (co)vectors so that with each point of the manifold is associated a (co)vector. In a coordinate system, it is represented by a set of n functions on the coordinates. The result of the action of a covector field on a vector field is a scalar function on the coordinates.
- A covector field is also called a *1-form*.
- The differential of a scalar function is a 1-form.
- In older terminology, vectors were called *contravariant vectors* and covectors were called *covariant vectors*.
- Vector fields and 1-forms have both n components, i.e., they can be represented (at least locally) by sets of n functions of the coordinates. There is a strong temptation to say that the sets of vector fields and 1-forms are n-dimensional vector spaces, but this forgets the fact that the components are functions that are themselves elements of infinite dimensional functional spaces.

- Geometrical tensor spaces are generated by tensor products of vector and covectors. For example, $A = A^i_{jk} dx^i \otimes \frac{\partial .}{\partial x^j} \otimes \frac{\partial .}{\partial x^k}$ is a rank 3 tensor, once contravariant and twice covariant and the A^i_{jk} are its n^3 components in the x^i coordinate system. A tensor field is a set of tensors so that with each point of the manifold is associated a tensor. A rank k ($k \in \mathbb{N}$) tensor field has n^k components.
- Skew-symmetric tensors play a fundamental role in differential geometry. A *k-covector* ω is a totally skew-symmetric tensor of rank k ($k \in \mathbb{N}$), i.e., it is a k-linear form on vectors such that the swapping of two vectors changes the sign of the resulting scalar: $\omega(\cdots, \mathbf{v}^i, \cdots, \mathbf{v}^j, \cdots) = -\omega(\cdots, \mathbf{v}^j, \cdots, \mathbf{v}^i, \cdots)$. Note that k-covectors are identically zero if $k > n$.

  Given $\{1, \cdots, n\}$, the set of integers from 1 to n, a *permutation* is a bijection $\sigma \in \mathcal{P}_n : \{1, \cdots, n\} \to \{1, \cdots, n\}$. A *transposition* is a permutation that swaps two elements and leaves the others at the same place. Any permutation can

be decomposed in a finite sequence of transpositions. The *signature* $\varepsilon(\sigma)$ of a permutation is the number equal to 1 if the number of transposition is even and to -1 if it is odd. It does not depend, of course, on the particular set of transpositions used to describe the permutation. The *Levi–Civita symbol* $\varepsilon_{i_1 i_2 \cdots i_n}$ is equal to $\varepsilon(\sigma)$ if the indices are a permutation σ of $\{1, \cdots, n\}$: $\{i_1 = \sigma(1), \cdots, i_n = \sigma(n)\}$ and equal to 0 if some indices are repeated. The general properties of k-covectors is $\omega(\mathbf{v}^1, \cdots, \mathbf{v}^k) = \varepsilon(\sigma)\omega(\mathbf{v}^{\sigma(1)}, \cdots, \mathbf{v}^{\sigma(k)})$.
The dimension of the vector space formed by the k-covectors on a vector space of dimension n is $\binom{n}{k} = \frac{n!}{k!(n-k)!}$.
A *k-form* is a field of k-covectors so that it is a map from the sets of k vector fields on M to the scalar functions on M. Note that k-forms are usually called *differential forms* or *exterior forms*. The vector space of k-forms on a manifold M is denoted by $\bigwedge^k(M)$.
The *exterior product* of $\alpha \in \bigwedge^k(M)$ and $\beta \in \bigwedge^j(M)$ is $\alpha \wedge \beta \in \bigwedge^{k+j}(M)$ such that

$$\alpha \wedge \beta(\mathbf{v}^1, \cdots, \mathbf{v}^{k+j}) = \frac{1}{k!j!}\sum_{\sigma \in \mathcal{P}_{k+j}} \varepsilon(\sigma)\alpha(\mathbf{v}^{\sigma(1)}, \cdots, \mathbf{v}^{\sigma(k)})\beta(\mathbf{v}^{\sigma(k+1)}, \cdots, \mathbf{v}^{\sigma(k+j)}).$$

The main properties of the exterior product of forms are:

$\alpha \wedge \beta = (-1)^{jk}\beta \wedge \alpha, \quad \forall \alpha \in \bigwedge^k(M), \beta \in \bigwedge^j(M)$.
$\alpha \wedge (\beta + \gamma) = \alpha \wedge \beta + \alpha \wedge \gamma$, for all forms α, β, γ.
$(\alpha \wedge \beta) \wedge \gamma = \alpha \wedge (\beta \wedge \gamma) = \alpha \wedge \beta \wedge \gamma$, for all forms α, β, γ.

The exterior products of coordinate differentials form a basis for the forms so that any k-form α can be written as:

$$\alpha = \alpha(x^1, \cdots, x^n)_{i_1 \cdots i_k} dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

- n-forms have a single component and can be written in the form $f(x^1, \cdots, x^n)dx^1 \wedge \cdots \wedge dx^n$. If f is everywhere different from zero on the manifold, the n-form is called a *volume form*. Given an n-form ω, n vector fields $\mathbf{v}^i$ and a matrix A with constant elements acting on the vector fields as a linear operator, we have $\omega(A\mathbf{v}^1, \cdots, A\mathbf{v}^n) = det(A)\omega(\mathbf{v}^1, \cdots, \mathbf{v}^n)$, which gives the geometrical meaning of the determinant.

- Another fundamental operation is the *exterior derivative* of a form defined by:

$$d\alpha = d\alpha_{i_1\cdots i_k} \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k} = \frac{\partial \alpha_{i_1\cdots i_k}}{\partial x^i}\, dx^i \wedge dx^{i_1} \wedge \cdots \wedge dx^{i_k}.$$

In the right hand, member $d\alpha_{i_1\cdots i_k}$ denotes the differential of the component $\alpha_{i_1\cdots i_k}$ (these coefficients are functions of the coordinates) and the implicit summation on repeated indices is used. From this definition, it is clear that

$$d : \bigwedge^k(M) \to \bigwedge^{k+1}(M).$$

The exterior derivative of a function is its differential; the exterior derivative of an n-form is zero. The main properties of the exterior derivative of forms are:

d is linear.
$d(\alpha \wedge \beta) = (d\alpha) \wedge \beta + (-1)^k \alpha \wedge d\beta, \quad \forall \alpha \in \bigwedge^k(M)$
(*Leibnitz rule*).
$dd\alpha = 0$, for all forms α.

- Given an open set Σ of dimension $p \leq n$ in a manifold M of dimension n, sufficiently regular, and, for the simplicity of the presentation, which can be covered by a single local coordinate system $\{x^1, \cdots, x^n\}$ so that Σ can be parameterized by $\Sigma : \{x^1(\xi^1, \cdots, \xi^p), \cdots, x^n(\xi^1, \cdots, \xi^p), \forall \{\xi^1, \cdots, \xi^p\} \in \Omega \subset \mathbb{R}^p\}$. Depending on the context, Σ can be, in fact, a *submanifold* (i.e., a manifold of dimension p together with a regular map of this manifold into M), a hyper-surface or a chain that corresponds to different objects from a formal point of view.[e] The *integral* $\int_\Sigma \alpha$ of a p-form α on Σ is defined by:

$$\int_\Sigma \alpha = \int \cdots \int_\Omega \alpha_{i_1\cdots i_p} \det\left(\frac{\partial(x^{i_1}, \cdots, x^{i_p})}{\partial(\xi^1, \cdots, \xi^p)}\right) d\xi^1 \cdots d\xi^p,$$

where the indices $i_1, \cdots, i_p$ run between 1 and n (without repetition), $\det\left(\frac{\partial(x^{i_1}, \cdots, x^{i_p})}{\partial(\xi^1, \cdots, \xi^p)}\right)$ are the Jacobians (i.e., the determinants of the matrices whose elements are the partial derivatives of the x^i with respect to the ξ^j) and

[e] The case $p = 1$ corresponds to a curve.

$d\xi^1 \cdots d\xi^p$ the Lebesgue measure on $\mathbb{R}^p$. This definition is, in fact, independent of the system of coordinates and of the parametrization used in the definition. The extension to the case where several local coordinate systems are necessary is purely technical and is based on the use of a partition of the unity.

The integration of a form is a linear operation. In order to emphasize the duality between the p-dimensional Σ and the p-forms, the abstract notation $< \Sigma, \alpha >= \int_\Sigma \alpha$ can be used.

The main property is the *Stokes theorem*. Given a $(p-1)$-form α and an open set Σ of dimension p such that its boundary $\partial\Sigma$ (of dimension $p-1$) is regular enough, we have:

$$\int_\Sigma d\alpha = \int_{\partial\Sigma} \alpha.$$

- The marriage of differential forms and distribution theory leads to the concept of *de Rham current*. First, a test p-form on a smooth manifold M (of dimension n) is a p-form the coefficients of which are $C_0^\infty(M)$ functions, i.e., infinitely differentiable and with a compact support. A p-current C is a continuous linear form on the test $(n-p)$-forms φ. The associated duality product is denoted by $< C, \varphi >$. The set of currents is, therefore, the topological dual of the space of test forms.

  With a p-form α is associated a p-current C_α, denoted as α by abuse of notation, such that $< C_\alpha, \varphi >= \int_M \alpha \wedge \varphi$, for all test $(n-p)$-forms φ.

  The lower dimensional submanifolds or hyper-surfaces give currents similar to the singular distributions[f] with an $(n-p)$-dimensional submanifold Σ is associated a p-current C_Σ, denoted as Σ by abuse of notation, such that

  $$< C_\Sigma, \varphi >= \int_\Sigma \varphi, \text{ for all test } (n-p) \text{ -forms } \varphi.$$

  The product of a p-current C by a C^∞ q-form α (not necessarily with a bounded support) is a $(p+q)$-current

[f] A first technical step is to consider formal linear combinations of those geometrical objects in order to give them a vector space structure.

defined by

$$< C \wedge \alpha, \varphi >=< C, \alpha \wedge \varphi >,$$

for all test $(n - p - q)$-form φ, and $C \wedge \alpha = (-1)^{pq}\alpha \wedge C$. For instance, with a pair (Σ, α), where Σ is an $(n - p)$-dimensional submanifold and α is a $C^\infty(\Sigma)$ q-form (which needs only be defined on the support of Σ), is associated a $(p+q)$-current $C_{\Sigma\wedge\alpha}$, denoted as $\Sigma \wedge \alpha$ by abuse of notation, such that $< C_{\Sigma\wedge\alpha}, \varphi >= \int_\Sigma \alpha \wedge \varphi$, for all test $(n-p-q)$-forms φ.

The exterior derivative of a p-current C is defined by $< dC, \varphi > = (-1)^{p-1} < C, d\varphi >$, for all test $(n - p - 1)$-forms φ. If the current is associated with a form, the definition coincides with the former definition of the exterior derivative. In the case of a p-current associated with an $(n - p)$-dimensional manifold Σ, the previous definition via the Stokes theorem gives:

$$< d\Sigma, \varphi >= (-1)^{p-1} < \Sigma, d\varphi >= (-1)^{p-1} < \partial\Sigma, \varphi >$$

hence $d\Sigma = (-1)^{p-1}\partial\Sigma$. For the p-current C and the C^∞ q-form α:

$$d(C \wedge \alpha) = dC \wedge \alpha + (-1)^p C \wedge d\alpha.$$

If Σ is an $(n-1)$-dimensional submanifold (and its associated 1-current) and if ω is a p-form discontinuous on Σ, i.e., the components in any coordinate system are differentiable in the complement of Σ in M except across Σ where they suffer a jump $[\omega]_\Sigma$, then[g] $d\omega = \{d\omega\} + \Sigma \wedge [\omega]_\Sigma$.

- All those geometrical notions, the exterior product, the exterior derivative, and the integration of a form, do not rely on the definition of a scalar product or a norm and are, therefore, purely topological and differential but not metric.
- **Riemmanian spaces**: A scalar product on the tangent vectors of a manifold can be defined as a rank 2 totally covariant symmetric tensor (field) **g** called the *metric*. In a coordinate system, this tensor can be written as $\mathbf{g} =$

[g] See the derivation of a discontinuous function in the section on the distributions above for the notation $\{d\omega\}$.

$g_{ij}dx^i \otimes dx^j$ where $g_{ij} = g_{ji}$ (the n^2 coefficients form a positive definite matrix with $n(n-1)/2$ independent values) and the scalar product of two vectors can be written as $(\mathbf{v}, \mathbf{w}) = g_{ij}v^i w^j$.

If the coefficients g_{ij} are considered to form a matrix, the coefficients of the inverse matrix are denoted by g^{ij} (with $g^{ik}g_{kj} = \delta^i_j$) and define a rank 2 totally contravariant symmetric tensor $g^{ij}\frac{\partial}{\partial x^i} \otimes \frac{\partial}{\partial x^j}$.

In the context of differential forms, the metric is mostly involved in the *Hodge star operator* $* : \bigwedge^p(M) \to \bigwedge^{(n-p)}(M)$, which maps p-forms on $(n-p)$-forms. The p- and $(n-p)$-covectors have the same number of components and the map is linear, one-one, and $** = (-1)^{p(n-p)}$. For any p-form α expressed in an arbitrary (covector) basis $\{\varepsilon^{i_1}, \cdots, , \varepsilon^{i_n}\}$ by $\alpha_{j_1\cdots j_p}\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_p}$, the action of the Hodge operator $*$ is given by

$$*(\alpha_{j_1\cdots j_p}\varepsilon^{i_1} \wedge \cdots \wedge \varepsilon^{i_p}) = \frac{1}{(n-p)!}\varepsilon_{i_1\cdots i_n}|\det[g_{ij}]|^{1/2}\alpha_{j_1\cdots j_p}g^{i_1 j_1}\cdots g^{i_p j_p}\varepsilon^{i_{p+1}} \wedge \cdots \wedge \varepsilon^{i_n}.$$

The Hodge operator allows the definition of a scalar product on the vector spaces $\bigwedge^p(M)$ of p-forms, which makes them Hilbert spaces[h] by setting

$$(\alpha, \beta) = \int_M \alpha \wedge *\beta.$$

The *coderivative* $\delta = (-1)^{n(p+1)+1} * d*$ is the formal adjoint of the exterior derivative[i] since one has $(d\alpha, \beta) = (\alpha, \delta\beta)$ for forms with $C_0^\infty(M)$ components. One has also $\delta\delta = 0$.

The *Laplace–Beltrami operator* Δ (*Laplacian* for short) is defined by $\Delta = (d+\delta)(d+\delta) = \delta d + d\delta$ and is self-adjoint.

- Of course, many of the previous operations are simpler and indeed even trivial in the *three-dimensional Euclidean space* $\mathbb{E}^3$ that is $\mathbb{R}^3$. This is considered a manifold and is equipped with a special metric such that there exist global coordinates

[h] The situation is, in fact, more subtle since it depends on the regularity of the components of the forms as functions of the coordinates, but we consider here that they are all in $L^2(M)$.

[i] Which could have then been denoted by d^*.

called *Cartesian coordinates* $\{x^1 = x, x^2 = y, x^3 = z\}$ where the metric has the form $\mathbf{g} = dx \otimes dx + dy \otimes dy + dz \otimes dz$. In these coordinates, the Hodge operator has the following action:
$*dx = dy \wedge dz, *dy = dz \wedge dx, *dz = dx \wedge dy,$
$*(dx \wedge dy) = dz, *(dz \wedge dx) = dy, *(dy \wedge dz) = dx,$
$*1 = dx \wedge dy \wedge dz, *(dx \wedge dy \wedge dz) = 1.$

- There are, in fact, several mathematical structures that can be considered "natural" descriptions of our three-dimensional perception of "space." $\mathbb{R}^3$ as a bare manifold obviously lacks structures, but defining Cartesian coordinates is too arbitrary since it involves, for instance, the choice of a distinguished point, the origin O of coordinates $\{0, 0, 0\}$. Another candidate is the "vector space $\mathbb{R}^3$," which we note $\mathbb{V}^3$ to make the distinction with the manifold. The addition of two points is now a valid but meaningless operation. A sounder choice is to consider $\mathbb{A}^3$, the affine space associated with $\mathbb{V}^3$ (loosely speaking obtained by forgetting the origin). The physical points are elements of $\mathbb{A}^3$; they cannot be added, but their differences are vectors of $\mathbb{V}^3$ and the elements of $\mathbb{V}^3$ operate on points of $\mathbb{A}^3$ as displacements. The introduction of the Euclidean distance gives the affine Euclidean space, which we still call $\mathbb{E}^3$ and where displacements (and not points) have a length. The distance between two points is then defined. Of course, in practice, an origin point and three mutually orthogonal unit vectors are chosen such that they define a particular Cartesian coordinate system.
- The peculiarities of $\mathbb{E}^3$ allow a simpler setting called *vector analysis*, which takes advantage of the Cartesian coordinates but forget almost all the geometric relevance! We do not want to advocate for giving up vector analysis, but we just want it to be considered as a computational trick rather than a genuine geometrical framework. In $\mathbb{E}^3$, 0-forms are scalar functions $v(x, y, z)$ and 3-forms are pseudo-scalars or *densities* $\rho(x, y, z)dx \wedge dy \wedge dz = \rho *1$. Both fields have only one single component so that they are merged into the single concept of *scalar field* using the Hodge operator. Similarly, the

1-forms $\alpha_x dx + \alpha_y dy + \alpha_z dz$ are the *field intensities* and the 2-forms $\beta_x dy \wedge dz + \beta_y dz \wedge dx + \beta_z dx \wedge dy$ are the *flux densities* and both have three components so that they are merged in the concept of *vector field*. In this case, those vector fields are considered *proxies* for 1-forms and 2-forms.

A vector field is written as $\mathbf{v} = v_x \varepsilon^x + v_y \varepsilon^y + v_y \varepsilon^y$ where the unit vectors ε are as well unit 1-forms as unit 2-forms associated with Cartesian coordinates. This is safe as long as only Cartesian coordinates are used. The scalar product of two vectors is called the *dot product* and is defined by $\mathbf{v} \cdot \mathbf{w} = v_x w_x + v_y w_y + v_z w_z$ and the associated norm is, of course, $|\mathbf{v}|^2 = \mathbf{v} \cdot \mathbf{v}$. Dot product can be traced back in differential geometry as the scalar product of two 1-forms or of two 2-forms but also to the metric free exterior product of a 1-form and a 2-form. This is usually enough to cloud the geometrical meaning of vector analysis computations! The *cross product* of two vectors is defined by

$$\mathbf{v} \times \mathbf{w} = (v_y w_z - v_z w_y)\varepsilon^x + (v_z w_x - v_x w_z)\varepsilon^y + (v_x w_y - v_y w_x)\varepsilon^z.$$

It comes mostly from the exterior product of two 1-forms.[j]

The exterior derivative gives rise to several operators.

The exterior derivative of a 0-form $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy + \frac{\partial f}{\partial z} dz$ corresponds to the *gradient* of a scalar field:

$$\mathbf{grad}\, f = \frac{\partial f}{\partial x}\varepsilon^x + \frac{\partial f}{\partial y}\varepsilon^y + \frac{\partial f}{\partial z}\varepsilon^z.$$

For a 1-form $\mathbf{v} = v_x dx + v_y dy + v_z dz$, the exterior derivative $d\mathbf{v} = (\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}) dy \wedge dz + (\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}) dz \wedge dx + (\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}) dx \wedge dy$ corresponds to the *curl* of a vector field:

$$\mathbf{curl}\, \mathbf{v} = \left(\frac{\partial v_z}{\partial y} - \frac{\partial v_y}{\partial z}\right)\varepsilon^x + \left(\frac{\partial v_x}{\partial z} - \frac{\partial v_z}{\partial x}\right)\varepsilon^y + \left(\frac{\partial v_y}{\partial x} - \frac{\partial v_x}{\partial y}\right)\varepsilon^z.$$

For a 2-form $\mathbf{w} = w_x dy \wedge dz + w_y dz \wedge dx + w_z dx \wedge dy$, the exterior derivative $d\mathbf{w} = (\frac{\partial w_x}{\partial w} + \frac{\partial w_y}{\partial y} + \frac{\partial w_z}{\partial z}) dx \wedge dy \wedge dz$ corresponds to the *divergence* of a vector field:

$$\text{div}\, \mathbf{w} = \frac{\partial w_x}{\partial w} + \frac{\partial w_y}{\partial y} + \frac{\partial w_z}{\partial z}.$$

[j] There are many other features in differential geometry not introduced here, such as the Lie derivative and the inner product. In electromagnetism, the cross product of the velocity together with the magnetic flux density in the Lorentz force is, in fact, the inner product of a vector with a 2-form.

Alternative notations for these operations use the nabla operator ∇: $\mathbf{grad}\ f = \nabla f$, $\mathbf{curl}\,\mathbf{v} = \nabla \times \mathbf{v}$, and $\mathrm{div}\ \mathbf{v} = \nabla \cdot \mathbf{v}$. The Leibnitz rule is expressed in the classical formulae:

$$\begin{aligned}
\mathbf{grad}(fg) &= f\,\mathbf{grad}\,g + g\ \mathbf{grad}\ f.\\
\mathbf{curl}(f\mathbf{v}) &= f\,\mathbf{curl}\,\mathbf{v} - \mathbf{v} \times \mathbf{grad}\ f.\\
\mathrm{div}(f\mathbf{v}) &= f\,\mathrm{div}\,\mathbf{v} + \mathbf{v} \cdot \mathbf{grad}\ f.\\
\mathrm{div}(\mathbf{v} \times \mathbf{w}) &= \mathbf{w} \cdot \mathbf{curl}\,\mathbf{v} - \mathbf{v} \cdot \mathbf{curl}\,\mathbf{w}.
\end{aligned}$$

and of course $dd = 0$ is nothing else than $\mathbf{curl}\ \mathbf{grad} = \mathbf{0}$ and $\mathrm{div}\ \mathbf{curl} = 0$.

The Laplacian of a scalar field is

$$\Delta f = \mathrm{div}\ \mathbf{grad}\ f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

The Laplacian of a vector field is

$$\Delta \mathbf{v} = \mathbf{grad}\ \ \mathrm{div}\,\mathbf{v} - \mathbf{curl}\,\mathbf{curl}\,\mathbf{v} = (\Delta v_x)\varepsilon^x + (\Delta v_y)\varepsilon^y + (\Delta v_z)\varepsilon^z.$$

Note that Laplacians are self-adjoint operators.

In vector analysis, a (rank 2) tensor (field) is a linear operator transforming a vector (field) into a vector (field). In a particular system of coordinates, it is usually given in the form of an array of 9 coefficients, hence the common confusion with a square matrix. A "vector analysis" tensor is denoted by $\underline{\underline{\alpha}}$ and its action on a vector $\mathbf{v}$ is simply denoted by $\underline{\underline{\alpha}}\mathbf{v}$.

The Stokes theorem corresponds to various integral equalities:

If Γ is a curve with initial point $\mathbf{a}$ and end point $\mathbf{b}$, for any scalar field whose gradient exists:

$$\int_\Gamma \mathbf{grad}\ f \cdot d\mathbf{l} = f(\mathbf{b}) - f(\mathbf{a})$$

where $d\mathbf{l}$ is the line element.

If Σ is a surface with boundary $\partial\Sigma$ (with an orientation inherited from Σ), for any vector field $\mathbf{v}$ whose curl exists:

$$\iint_\Sigma \mathbf{curl}\,\mathbf{v} \cdot \mathbf{n}\, ds = \int_{\partial\Sigma} \mathbf{v} \cdot\ d\mathbf{l}$$

where ds is the surface element and $\mathbf{n}$ is the normal vector on Σ.

If V is a volume with boundary ∂V, for any vector field $\mathbf{v}$ whose divergence exists:

$$\iiint_V \operatorname{div} \mathbf{v}\, d\mathbf{v} = \iint_{\partial V} \mathbf{v} \cdot \mathbf{n}\, ds$$

where $d\mathbf{v}$ is the volume element and $\mathbf{n}$ is the outer normal vector on ∂V.

Often, multiple integrals are simply denoted by a single $\int$ just as in differential geometry.

The simplest way to give natural definitions of all the integrals involved here above is to go back to the definition of the integration of differential forms. *Moreover, the traditional notations of vector analysis rely on metric concepts such as the normal vector (involving both orthogonality and unit length) and the scalar product while the definition of these integrals do not require any metric.*

Given a unit vector ε, the *directional derivative* $\frac{\partial \cdot}{\partial \varepsilon}$ is defined by $\frac{\partial f}{\partial \varepsilon} = \mathbf{grad}\ f \cdot \varepsilon$. Another fundamental integral identity is the *Green's formula*:

$$\iiint_V (f \Delta g - g \Delta f) d\mathbf{v} = \iint_{\partial V} \left(f \frac{\partial g}{\partial \mathbf{n}} - g \frac{\partial f}{\partial \mathbf{n}} \right) ds\,.$$

- Using the Leibnitz rule to integrate by parts, it is easy to find the formal adjoint to the vector analysis operators:

$$\mathbf{grad}^* = -\operatorname{div},\ \mathbf{curl}^* = \mathbf{curl},\ \text{and}\ \operatorname{div}^* = -\mathbf{grad}\,.$$

- The identity $d\omega = \{d\omega\} + \Sigma \wedge [\omega]_\Sigma$ for de Rham currents leads to the following identities in vector analysis in the case of a function f and a vector field $\mathbf{v}$ differentiable in the complement of a surface Σ but undergoing the jumps $[f]_\Sigma$ and $[\mathbf{v}]_\Sigma$ across Σ, respectively (the direction chosen to cross the surface is given by $\mathbf{n}$, the normal vector to the surface Σ):

$$\begin{aligned} \mathbf{grad}\, f &= \{\mathbf{grad}\, f\} + \mathbf{n}[f]_\Sigma \delta_\Sigma. \\ \mathbf{curl}\, \mathbf{v} &= \{\mathbf{curl}\, \mathbf{v}\} + \mathbf{n} \times [\mathbf{v}]_\Sigma \delta_\Sigma. \\ \operatorname{div} \mathbf{v} &= \{\operatorname{div} \mathbf{v}\} + \mathbf{n} \cdot [\mathbf{v}]_\Sigma \delta_\Sigma. \end{aligned}$$

where it is necessary to denote explicitly the singular distribution δ_Σ associated with the surface Σ (by definition $\int_{\mathbb{R}} \delta_\Sigma \varphi(\mathbf{x}) d\mathbf{x} = \varphi|_\Sigma$). Note that metric concepts creep again in a place where they are not necessary.

- $L^2(\mathbb{R}^3, \mathbb{C}^3) = [L^2(\mathbb{R}^3)]^3 = \mathbb{L}^2(\mathbb{R}^3)$ denotes the space of square integrable functions on $\mathbb{R}^3$ with values in $\mathbb{C}^3$ where the scalar product is defined by $(\mathbf{v}, \mathbf{w}) = \int_{\mathbb{R}^3} \mathbf{v}(\mathbf{x}) \cdot \overline{\mathbf{w}(\mathbf{x})} d\mathbf{x}$.
- The Green's function G of three-dimensional scalar Laplacian (in free space) is such that $\Delta_{\mathbf{p}} G(\mathbf{p} - \mathbf{q}) = \delta(\mathbf{p} - \mathbf{q})$ where $\mathbf{p}, \mathbf{q} \in \mathbb{E}^3$ and $\Delta_{\mathbf{p}}$ indicates that the derivatives in the Laplacian are taken with respect to the coordinates of $\mathbf{p}$ (but this is seldom explicitly indicated when there is no ambiguity). The expression of this Green's function is $G(\mathbf{p} - \mathbf{q}) = \frac{-1}{4\pi|\mathbf{p}-\mathbf{q}|}$. The displacement vector $\mathbf{p} - \mathbf{q}$ is traditionally called $\mathbf{r}$ and its norm r and one can be written $G = \frac{-1}{4\pi r}$.
  In the two-dimensional scalar Laplacian case, $G = \frac{1}{2\pi} \ln(r)$.
  For the scalar Helmholtz equation, one has $(\Delta + k^2)H = \delta$ and $H = -\frac{e^{ikr}}{4\pi r}$.
  In the two-dimensional scalar Helmholtz equation case, $H = \frac{1}{4i} H_0^{(1)}(kr)$ where $H_0^{(1)}$ denotes a Hankel function.
  The Green's function is physically interpreted as the field (or potential) generated by a (monopolar) point source. The introduction of Green's functions in Green's formula gives identities, which are the basis of the *boundary element method*.
- The differential geometry on $\mathbb{R}^2$ provides a nice framework to introduce *complex analysis* on $\mathbb{C}$. Given two Cartesian coordinates x and y and the Euclidean metric $\mathbf{g} = dx \otimes dx + dy \otimes dy$, the associated Hodge star operator has the following action on 1-forms: $*dx = dy$ and $*dy = -dx$. This action is a $\pi/2$ rotation counterclockwise equivalent to the action of a multiplication by i in the complex plane. The knowledge of this Hodge star on the plane does not determine the Euclidean geometry but only the conformal geometry, i.e., the metric up to a scalar factor so that only angles but not lengths are relevant. In the complex plane, the complex variable $z = x + iy$ and its complex conjugate $\overline{z} = x - iy$ play the role of independent variables. Their differentials $dz = dx + i\,dy$ and $d\overline{z} = dx - i\,dy$ are a basis for the 1-forms. The dual basis for the vectors are the differential

operators $\frac{\partial}{\partial z} = \frac{1}{2}(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y})$ and $\frac{\partial}{\partial \overline{z}} = \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})$. Note that $\frac{\partial}{\partial \overline{z}}\frac{\partial}{\partial z} = \frac{\partial}{\partial z}\frac{\partial}{\partial \overline{z}} = \frac{1}{4}\Delta$ and $dz \wedge d\overline{z} = -2i(dx \wedge dy)$.

A complex function of the complex variable $f : \mathbb{C} \to \mathbb{C}$ is associated with a pair (p, q) of real functions on $\mathbb{R}^2$ so that $f(z, \overline{z}) = p(x, y) + iq(x, y)$. Those complex functions are usually denoted by $f(z)$, which is fully justified in the case of holomorphic functions as explained below.

Given two complex functions f and g, the exterior derivative of the 0-form f is

$$df = \frac{\partial f}{\partial z}dz + \frac{\partial f}{\partial \overline{z}}d\overline{z} = \left(\frac{\partial p}{\partial x} + i\frac{\partial q}{\partial x}\right)dx + \left(\frac{\partial p}{\partial y} + i\frac{\partial q}{\partial y}\right)dy$$

and the exterior derivative of the 1-form $f\,dz + g\,d\overline{z}$ is

$$d(f\,dz + g\,d\overline{z}) = \left(\frac{\partial g}{\partial z} - \frac{\partial f}{\partial \overline{z}}\right)dz \wedge d\overline{z}.$$

Let γ be a closed curve in $\mathbb{C}$ (and, therefore, also in $\mathbb{R}^2$), i.e., a map $\gamma : t \in [a, b] \subset \mathbb{R} \to z(t) \in \mathbb{C}$ such that $\gamma(a) = \gamma(b)$, and such that there are no other multiple points. The line integral of the complex 1-form $f dz$ associated with f (note that it is not a general 1-form since there is no term in $d\overline{z}$) is $\int_\gamma f dz = \int_\gamma (p\,dx - q\,dy) + i\int_\gamma (p\,dy + q\,dx)$. The exterior derivative is $d(f dz) = -\frac{\partial f}{\partial \overline{z}}\,dz \wedge d\overline{z}$.

If D is a bounded domain of $\mathbb{C}$ (and, therefore, also of $\mathbb{R}^2$), the domain integral of a complex 2-form $f\,dz \wedge d\overline{z}$ is

$$\int_D f\,dz \wedge d\overline{z} = -2\int_D p\,dx \wedge dy + 2i\int_D q\,dx \wedge dy.$$

In this case, the Stokes theorem becomes $\int_D d(f dz) = \int_{\partial D} f dz$. As for a 1-form $f\,d\overline{z}$, one has

$$\int_D d(f d\overline{z}) = \int_{\partial D} f d\overline{z}$$

where $\int_\gamma f d\overline{z} = \int_\gamma (p\,dx + q\,dy) + i\int_\gamma (q\,dx - p\,dy)$.

A function $f(z)$ is *holomorphic* or *analytic* on a topologically simple domain (contractible to a point) $D \subset \mathbb{C}$ if one of the following five equivalent conditions are satisfied:

(1) If at each point of D, f is $\mathbb{C}$-differentiable, i.e., if the limit $f'(z) = \lim_{|h|\to 0}\frac{f(z+h)-f(z)}{h}$ exists and is independent of the direction of $h \in \mathbb{C}$ (Fréchet derivative). In this case $f'(z) = \frac{\partial f(z)}{\partial z} = \frac{1}{2}(\frac{\partial p}{\partial x} + \frac{\partial q}{\partial y}) - \frac{i}{2}(\frac{\partial p}{\partial y} - \frac{\partial q}{\partial x})$.

(2) If the 1-form $f\,dz$ is closed, i.e., its exterior derivative vanishes, on D: $d(f\ dz) = -\frac{\partial f}{\partial \overline{z}}\ dz = 0$ at each point of D, i.e., if the *Cauchy–Riemann conditions* $\frac{\partial p}{\partial x} - \frac{\partial q}{\partial y} = 0$ and $\frac{\partial p}{\partial y} + \frac{\partial q}{\partial x} = 0$ are verified. This condition means that $f(z)$ behaves as if it were independent of $\overline{z}$. The operator $\frac{\partial}{\partial \overline{z}}$ is called the Cauchy–Riemann operator, and since it is a factor of the Laplacian, a holomorphic function f is also a harmonic function, i.e., $\Delta f = 0$.

(3) If $df = g\ dz$, i.e., if the exterior derivative of f has no $d\overline{z}$ component.

(4) If the line integrals $\int_\gamma f\ dz = 0$ for any closed curve γ contained together with its interior in D. As any closed curve can be defined as a boundary $\partial\Omega$ with $\Omega \subset D$, by the Stokes theorem $\int_{\partial\Omega} f dz = \int_\Omega d(f dz) = 0$.

(5) If f can be represented by an infinite power series $f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$ in a neighborhood of any point of z_0 of D.

In the case of the function $1/z$ on a domain Ω, including the origin $z = 0$, this function is holomorphic everywhere on Ω except at the origin where there is a singularity called a *pole* of order 1. Such a function, which is holomorphic except for a discrete set of points, is called *meromorphic*. The fact that the domain in which the function is holomorphic is no more topologically simple has important consequences: Consider for instance a disc $D(r, 0)$ of radius r and center $z = 0$. Using the polar representation $z = re^{i\theta}$, $dz = ire^{i\theta} d\theta$ on $\partial D(r, 0)$ and one has, for any $n \in \mathbb{Z}$,

$$\int_{\partial D(r,0)} z^n dz = \int_0^{2\pi} e^{in\theta} r^n ire^{i\theta} d\theta = ir^{n+1} \int_0^{2\pi} e^{i(n+1)\theta} d\theta = 2i\pi$$

if $n = -1$ and 0 if $n \neq -1$. As $1/z$ is holomorphic on any domain not including the origin, its domain integral is null and this fact can be used to easily prove that $\int_{\partial\Omega} \frac{1}{z} dz = 2i\pi$ for any Ω, including the pole $z = 0$. A holomorphic function f on Ω, including the origin $z = 0$, can be written as an infinite power series $f(z) = \sum_{n=0}^{\infty} a_n z^n$ where the z^n are all holomorphic functions and where $a_0 = f(0)$. Therefore, $\int_{\partial\Omega} \frac{f(z)}{z} dz = \int_{\partial\Omega} \sum_{n=0}^{\infty} a_n z^{n-1} dz = 2i\pi f(0)$. It is easy to shift the pole at any ζ inside a domain Ω where f is holomorphic

to obtain the *Cauchy's integral formula* or *Cauchy theorem*:

$$f(\zeta) = \frac{1}{2i\pi} \int_{\partial\Omega} \frac{f(z)dz}{z-\zeta}.$$

If the function f is meromorphic, it can be represented by a *Laurent series* $f(z) = \sum_{n=-p}^{\infty} a_n (z-z_0)^n$ (involving negative powers of z) in a neighborhood D of a pole z_0 of order p. The coefficient a_{-1} denoted by $R(z_0; f)$ is called the *residue* and we have: $R(z_0; f) = \lim_{z\to z_0}(z-z_0)f(z) = \frac{1}{2i\pi}\int_\gamma f(z)dz$ for a curve γ such that it is contained in D, and that z_0 is the only pole contained in its interior.

Given two complex functions f and g on D, a duality product $< f, g >= \int_D f\overline{g}\, dz \wedge d\overline{z}$ can be introduced so that $-\frac{\partial}{\partial \overline{z}}$ is the formal adjoint of $\frac{\partial}{\partial z}$ since

$$< \frac{\partial f}{\partial z}, g >= - < f, \frac{\partial g}{\partial \overline{z}} > + \int_{\partial D} f\overline{g}\, d\overline{z}.$$

The Green's function of $\frac{\partial}{\partial z}$ is given by $\frac{\partial}{\partial \overline{z}}\frac{1}{\pi z} = \delta$, where δ is the Dirac distribution on $\mathbb{R}^2$, and it can be justified by $\int_{\partial D} f \frac{1}{2i\pi z} dz = f(0) = \int_D f\delta \frac{dz\wedge d\overline{z}}{-2i} = \int_D f \frac{\partial g}{\partial \overline{z}} \frac{dz\wedge d\overline{z}}{-2i} = \frac{1}{2i}\int_D d(fg\, dz) = \frac{1}{2i}\int_{\partial D} fg\, dz$. The theory of hyperfunctions is a distribution theory on $\mathbb{C}$ where test functions are holomorphic.

- *Landau notation*: Given x_0 and two functions f and g defined in a neighborhood of $x_0 \in \overline{\mathbb{R}}$ (i.e., real numbers, including the cases $x_0 \to \pm\infty$), $f = O(g)$ if and only if $\frac{f(x)}{g(x)} = O(1)$, which means that $|\frac{f(x)}{g(x)}|$ stays bounded in a neighborhood of x_0. Moreover, $f = o(g)$ if and only if $\frac{f(x)}{g(x)} = o(1)$, which means that $\lim_{x\to x_0} \frac{f(x)}{g(x)} = 0$. It is legitimate to write, say, $2x = O(x) = o(x^2)$ for $x \to \infty$ with the understanding that we are using the equality sign in an unsymmetrical (and informal) way, in that we do not have, for example, $o(x^2) = O(x)$.

  Landau notation is a practical tool to give asymptotic behaviors, e.g., for $\nu \in \mathbb{R}$,

$$J_\nu(x) = \sqrt{\frac{2}{\pi x}} \cos\left(x - (2\nu+1)\frac{\pi}{4}\right) + O\left(\frac{1}{x\sqrt{x}}\right) \text{ as } x \to +\infty.$$

Landau notation is also handy in computer science, e.g., in describing the efficiency of an algorithm. It is common to say that an algorithm requires $O(n^3)$ steps, e.g., without needing to specify exactly what is a step; for if $f = O(n^3)$, then $f = O(An^3)$ for any positive constant A.

- Perfect notations probably do not exist since absolutely rigorous ones should be intractable! It is difficult to avoid ambiguities and collisions. For instance, using $\star$ for the convolution, $*$ for the Hodge star operator, and the upper index * for algebraic duals may not be appreciated by presbyopic readers. Adopting nonambiguous but nonstandard notations may be a solution to avoid collisions. The danger is to lose the reader in a cumbersome formal deciphering game. Therefore, we prefer to leave some ambiguities, which may be removed by understanding.

  Moreover, we leave on side with regret some important issues that are not explicitly used here, such as the orientation of manifolds and twisted forms, the Hodge orthogonal decomposition theorem for forms that generalizes the Helmholtz decomposition theorem for vector fields.

  There is, of course, a huge number of books on the mathematical tools for physics, but almost everything in this appendix can be found in a more detailed and rigorous version in the formidable Ref. (Choquet-Bruhat et al., 1982). The classical reference for functional analysis is Ref. (Yosida, 1980), but a more readable book for the physicist interested in the functional analysis for partial differential equations is Ref. (Folland, 1995) and a concise introduction to basic analysis and functional analysis is Ref. (Friedman, 1982). One of the best reference for distribution theory, including de Rham currents, is still Ref. (Schwartz, 1966). There are many good books on geometrical methods in physics, such as the very pedagogical Ref. (Bamberg and Sternberg, 1991) or the quite comprehensive Ref. (Nakahara, 1990) but a very concise presentation aimed at electromagnetism is Ref. (Bossavit, 1991).

References

Bamberg, P. and Sternberg, S. (1991). *A Course in Mathematics for Students of Physics: vol. 1 and vol. 2* (Cambridge University Press, Cambridge).

Bossavit, A. (1991). *Notions de géométrie différentielle pour l'étude des courants de Foucault et des méthodes numériques en électromagnétisme*, Vol. Méthodes numériques en électromagnétisme (A. Bossavit, C. Emson, I. Mayergoyz) (Eyrolles, Paris), see also the English translation httpee also the English translation http://www.lgep.supelec.fr/mse/perso/ab/DGSNME.pdf.

Choquet-Bruhat, Y., Witt-Morette, C. D., and Dillard-Bleick, M. (1982). *Analysis, Manifolds and Physics*, revised edn. (Elsevier, Amsterdam).

Folland, G. B. (1995). *Introduction to Partial Differential Equations* (Princeton University Press, Princeton).

Friedman, A. (1982). *Foundation of Modern Analysis* (Dover, New York).

Nakahara, M. (1990). *Geometry, Topology and Physics* (IOP, Bristol).

Schwartz, L. (1966). *Théorie des Distributions*, 3rd edn. (Hermann, Paris).

Yosida, K. (1980). *Functional Analysis*, 6th edn. (Springer, Berlin).

Index

"*Beginning with a distillation of the Maxwell postulates, the authors move on first to the dielectric and magnetic response characteristics of actual materials and the band diagrams of diverse crystals, and then to a comprehensive overview of the electromagnetic characteristics of periodic distributions of matter in space. Homogenizable as well as nonhomogenizable periodic composites are described, along with a useful introduction to transformation optics.*"

Prof. Akhlesh Lakhtakia
Pennsylvania State University, USA

"*This entry-level book is suitable for everyone interested in understanding and mastering metamaterial design. Starting from the necessary details of basic principles of electromagnetism, it develops all the way to illustrate implementations of sophisticated numerical models and metamaterial homogenization approaches. Care is taken to underline physical insights behind mathematical notions.*"

Prof. Anatoly Zayats
King's College London, UK

"*I enjoyed reading the book edited by Felbacq and Bouchitté. It is a solid reference for the rigorous modeling and design of metamaterials, with special emphasis on the mathematical aspects. It will serve as a great reference for the field and as an inspiration for scientists and students entering this area of research.*"

Prof. Andrea Alù
University of Texas, USA

This book covers the fundamental physics, mathematics, and numerics necessary for entering the field of metamaterials. It presents advanced mathematical methods in a self-consistent way, along with numerous examples. It focuses on electromagnetic waves but is also useful in studying other types of metamaterials. It presents the structure of Maxwell equations, discusses the homogenization theory in detail, and includes important problems on resonance. It has an entire section devoted to numerical methods (finite elements, scattering theory), which motivates a reader to implement them. It offers numerous interesting examples at the forefront of research. The book is not written as a collection of independent chapters but as a textbook with a strong pedagogical flavor.

Didier Felbacq is full professor at the University of Montpellier, France. He graduated in mathematics and physics from the École Centrale de Marseille, France, and Aix-Marseille University, France, respectively. His current research activities cover electron transport in transistors for terahertz emission and detection, second harmonic emission in photonic crystals, excitons in 2D materials, quantum and thermal metamaterials, and modeling in biology.

Guy Bouchitté is full professor at the University of Toulon, France. He graduated from the École Centrale Paris, France, and defended a doctoral thesis in mathematics at the University of Montpellier. He is a specialist in the mathematical theory of homogenization, of optimization, and of variational calculus, and his current research activity covers applications of mathematics to photonic crystals and metamaterials.

V218
ISBN 978-981-4316-12-5
9 789814 316125

PAN STANFORD PUBLISHING
www.panstanford.com