MSE 487 - Exam sheet

Weeks 1 and 2: algebraic structures

Proof by induction

- For $n_0 \in \mathbb{N}$, for a proposition P(n) ($n \in \mathbb{N}$) to be true for all $n \geq n_0$, it is necessary and sufficient that:
  - P(n_0) is true
  - For all $n \geq n_0$, if P(n) is true, then P(n+1) is also true.

- Strong induction: It is equivalent to show that: if it is true for n_0, and for all integers < n, then P(n) is also true.

Set, permutation, Combinatorial

- If a finite set E contains *n* elements ($n \in \mathbb{N}$), *n* is also called the order or the cardinal, then the number of part, or sub-ensembles of E, including the "empty" part and E itself, is 2^n.
- If a set E contains *n* elements ($n \in \mathbb{N}$), the number of ways to arrange them is n!
- $n! = \prod_{k=1}^{n} k = 1 \times 2 \times 3 \times \ldots \times (n-2) \times (n-1) \times n$ with $0! = 1$
- A permutation, or arrangement, of *p* elements ($p \leq n$) of E is a sub-set of E with elements arranged in a certain way. The ordering is here important: a set of similar elements but arranged differently forms a different arrangement.
- Number of arrangements of p elements among n: $\prod_{k=0}^{p-1}(n-k) = \frac{n!}{(n-p)!} = A_n^p$

- A combination is the number of ways of selecting p elements among n, without considering their permutation. The number of ways to select *p* elements among *n*, is: $\binom{n}{p} = \frac{A_n^p}{p!} = \frac{n!}{p!(n-p)!}$

  - Pascal relation: $\binom{n+1}{p+1} = \binom{n}{p+1} + \binom{n}{p}$
  - Newton binomial: $\forall (a,b) \in \mathbb{C}^2,\ \forall\, n \in \mathbb{N}:\quad (a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$

Euclidean division

- Given two integers $(a,b) \in \mathbb{Z}^2$, with $b \neq 0$, there exist unique integers *q* and *r* such that: $a = bq + r$ and $0 \leq r < |b|$,
- Given two integers $(a,b) \in \mathbb{Z}^2$, *a* divides *b* if there exists an integer *q* such that $a = bq$.
- An equivalent definition is *a* divides *b* if and only if the remainder *r* of the Euclidean division is zero.

- We consider $\{x_k,\ k \in \mathbb{N}\ and\ 1 \leq k \leq n,\ and\ x_k \in \mathbb{Z}^*\}$.
  - The ensemble of the dividers of the x_k admits a maximum, called the greatest common divider and defined as gcd(x_k).
  - The ensemble of the multiples of the x_k admits a minimum, called the lowest common multiple and is defined as lcm(x_k)

Prime numbers, co-primes

- A prime number is a number greater than one that is only divided by 1 and itself.
- Every integer greater than 1 can be represented uniquely as a product of prime numbers, up to the order of the factors. For all integers *n,* there exists prime numbers p_i and integers n_i ($1 \leq i \leq k$), such that

$$n = \prod_{i=1}^{k} p_i^{n_i}$$

- Two integers *a* and *b* are mutually prime (or co-prime, relatively prime), if $gcd(a,b) = 1$.
- This definition can be extended to *n* integers x_i, which are called mutually prime if $gcd(x_1,\ldots,x_n) = 1$.
- Theorem of Bézout:
  For *n* non zero integers x_i, $gcd(x_1, \ldots, x_n) = d$. Then, $\exists\ (d_1, \ldots, d_n) \in \mathbb{Z}^n$ such that

$$\sum_{i=1}^{n} d_i x_i = d$$

- Important corollary to Bézout's theorem:
  If *n* non zero integers x_i are mutually prime, or co-prime, i.e. if $gcd(x_1, \ldots, x_n) = 1$, then $\exists\ (d_1, \ldots, d_n) \in \mathbb{Z}^n$ such that:

$$\sum_{i=1}^{n} d_i x_i = 1$$

- Gauss Theorem: $\forall\ (a,b,c) \in (\mathbb{Z}^*)^3,\ \{a|bc\ \&\ \gcd(a,b) = 1\} \Longrightarrow a|c$

- Euclid's lemma: If a prime *p* divides the product *ab* of two integers *a* and *b*, then *p* must divide at least one of those integers *a* or *b*. (as can be seen by a direct application of Gauss theorem).

- $\mathbb{Q}$
  - The ensemble of rational numbers is defined as the ensemble $\mathbb{Q} = \left\{\frac{p}{q},\ (p,q) \in \mathbb{Z} \times \mathbb{Z}^*\right\}$
  - $\mathbb{Q}$ is dense in $\mathbb{R}$: $\forall (x,y) \in \mathbb{R}^2,\ \exists\ z\ \in \mathbb{Q}$ such that $x < z < y$.
  - $\forall x \in \mathbb{Q},\ \exists!\ (p,q) \in \mathbb{Z} \times \mathbb{N}^*$ such that $x = \frac{p}{q}$ & gcd $(p,q) = 1$ (ie *p* and *q* are co-prime).
- $\mathbb{R}$
  - Absolute value: $\forall x \in \mathbb{R},\ |x| = (x\ if\ x \geq 0,\ -x\ if\ x \leq 0)$.
    $\forall (x,y) \in \mathbb{R}^2,\ |x+y| \leq |x| + |y|$, and $\big||x| - |y|\big| \leq |x - y|$
  - Inequality of Cauchy-Schwartz:
    $$\left(\sum_{i=1}^{n} x_i y_i\right)^2 \leq \left(\sum_{i=1}^{n} {x_i}^2\right)\left(\sum_{i=1}^{n} {y_i}^2\right)$$
  - Inequality of Minkowsky: $\sqrt{\sum_{i=1}^{n}(x_i+y_i)^2} \leq \sqrt{\sum_{i=1}^{n} {x_i}^2} + \sqrt{\sum_{i=1}^{n} {y_i}^2}$
  - <u>Nth root</u>: $\forall (y,n) \in \mathbb{R}_+ \times \mathbb{N}^*,\ \exists!\ x \in \mathbb{R}$ such that $x^n = y$.
  - Root of a second-degree polynomial
    For $(a,b,c) \in \mathbb{R}^3,\ a \neq 0$ we consider the trinomial for $x \in \mathbb{R}, T(x) = ax^2 + bx + c$ and its discriminant $\Delta = b^2 - 4ac$:
    - If $\Delta < 0$, $\forall x \in \mathbb{R}$, $aT(x) > 0$
    - If $\Delta = 0$, T has one root $-\frac{b}{2a}$, and $aT(x) \geq 0$
    - If $\Delta > 0$, T has two roots x' and x'' with $x' = \frac{-b-\sqrt{\Delta}}{2a}$ and $x'' = \frac{-b+\sqrt{\Delta}}{2a}$
- Polynomials:
  - Lagrange Polynomial: for a function in $\mathbb{R}$ or $\mathbb{C}$, for a given set of n numbers $(a_k)_{1\leq k\leq n}$, there is a unique polynomial P such that $\forall k,\ f(a_k) = \mathrm{P}(a_k)$.
    $$P(x) = \sum_{k-1}^{n} f(a_k) \frac{\prod_{j\neq k}(x - a_j)}{\prod_{j\neq k}(a_k - a_j)}$$
  - A polynomial of degree *n* in $\mathbb{R}$ can have a maximum of *n* roots, and the polynomials $(X - \alpha)^\beta$ are irreducible factors, very much like prime numbers for numbers.
  - A polynomial in $\mathbb{R}$ is said split, if $\exists \alpha_i \in \mathbb{R},\ \beta_i \in \mathbb{N}$ such that $P(X) = \prod_i (X - \alpha_i)^{\beta_i}$
    If $\beta_i > 1$, the root is said degenerate. If *deg(P) = n*, then $n = \sum_i \beta_i$
  - Every polynomials in $\mathbb{C}$ has at least one root. Corollary: every polynomial in $\mathbb{C}$ is split.

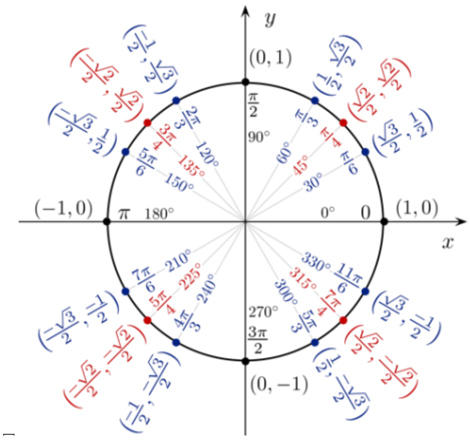
Euclidean geometry

- The following notation will be used: $\boldsymbol{a} = \begin{pmatrix} a_x \\ a_y \\ a_z \end{pmatrix}$
- The magnitude (or norm) of a vector: $\|\boldsymbol{a}\| = \sqrt{{a_x}^2 + {a_y}^2 + {a_z}^2}$
- Scalar (or dot) product: for two vectors in the **orthonormal basis** $\boldsymbol{i}, \boldsymbol{j}, \boldsymbol{k}$, we have: $\boldsymbol{a}.\boldsymbol{b} = a_x b_x + a_y b_y + a_z b_z$
  $\boldsymbol{a}.\boldsymbol{b} = \|\boldsymbol{a}\| \times \|\boldsymbol{b}\| \cos(\alpha)$ where α is the angle between the two vectors.
- The cross product of two vectors forming an angle α is a vector perpendicular to these vectors, with the magnitude:
  $\|\boldsymbol{a} \times \boldsymbol{b}\| = \|\boldsymbol{a}\|\ \|\boldsymbol{b}\| \sin(\alpha)$
- In an orthonormal basis (***i,j,k***), the Cross product of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$ is:
  $$\boldsymbol{a} \times \boldsymbol{b} = (a_y b_z - a_z b_y)\boldsymbol{i} + (a_z b_x - a_x b_z)\boldsymbol{j} + (a_x b_y - a_y b_x)\boldsymbol{k}$$
- **Line**: Parametric equation of a line passing by two points A and B: $L = \left\{M = \begin{pmatrix} x \\ y \\ z \end{pmatrix} such\ that\ \exists \lambda \in \mathbb{R}\ \boldsymbol{AM} = \lambda \boldsymbol{AB}\right\}$
- **Plane**:
  - A plane is defined by 3 points $A = \begin{pmatrix} x_A \\ y_A \\ z_A \end{pmatrix}$, $\mathrm{B} = \begin{pmatrix} x_B \\ y_B \\ z_B \end{pmatrix}$ and $\mathrm{C} = \begin{pmatrix} x_C \\ y_C \\ z_C \end{pmatrix}$ or a point A and a normal $\boldsymbol{n} = \begin{pmatrix} n_x \\ n_y \\ n_z \end{pmatrix}$
  - This can be expressed in a simple way as: $P = \left\{M = \begin{pmatrix} x \\ y \\ z \end{pmatrix},\ \boldsymbol{AM}.\boldsymbol{n} = 0\right\}$
  - One can extract the linear equation: for $(a,b,c,d) \in \mathbb{R}^4$, $P = \left\{M = \begin{pmatrix} x \\ y \\ z \end{pmatrix},\ ax + by + cz - d = 0\right\}$
- **Angles**
  - The angle between two vectors can be calculated from the dot or the cross products.
  - Angle between a line and a plane: Complementary of the angle between the line direction and the normal of the plan
  - Angle between two planes: Angle between their normals:
  - Volume formed by three vectors: $V = \vec{a}.(\vec{b} \times \vec{c}) = \vec{b}.(\vec{c} \times \vec{a}) = \vec{c}.(\vec{a} \times \vec{b})$

Week 3 : Complex Numbers

- The form $z = x + iy$ constitutes the **algebraic form** of a complex number z.
- x is called the real part and written $x = Re(z)$, and y is the Imaginary part with $y = Im(z)$.
- For two complex numbers z and z', *Re(z+z') = Re(z) + Re(z')* and *Im(z+z') = Im(z) + Im(z')*
  z = z' if and only if *Re(z) = Re(z')* and *Im(z) = Im(z')*
- **Conjugate**: $z^* = x - iy$. Also denoted by $\bar{z}$.
- The **modulus** of a complex number $z = x + iy$ is given by: $|z| = \sqrt{zz^*} = \sqrt{x^2 + y^2}$
- For $(z, z') \in \mathbb{C}^2$, the multiplication proceeds as follow: $z \times z' = (x + iy) \times (x' + iy') = (xx' - yy') + i(x'y + xy')$
- The division: $\frac{z}{z'} = \frac{x+iy}{x'+iy'} = \frac{(x+iy)(x'-iy')}{|z'|^2} = \frac{xx'+yy'}{x'^2+y'^2} + \mathrm{i}\frac{x'y-xy'}{x'^2+y'^2}$
- **Polar form**: $z = x + iy$ if we call *r* the magnitude of the depicted vector, then : $x = rcos\theta$, $y = rsin\theta$
  One can write : $z = rcos\theta + irsin\theta = r(cos\theta + isin\theta)$, r is the modulus and θ is the argument. $r = \sqrt{x^2 + y^2}$ and $\tan\theta = \frac{y}{x}$
- **Exponential form**: $z = rcos\theta + irsin\theta = re^{i\theta}$
- For $z \in \mathbb{C}$, $z = re^{i\theta}$, $z^* = re^{-i\theta}$
- $|e^{i\theta}| = 1 = \sqrt{x^2 + y^2}$, with $x = cos\theta$ and $y = sin\theta$
- $\cos(x) = \frac{e^{ix}+e^{-ix}}{2}$ *and* $\sin(x) = \frac{e^{ix}-e^{-ix}}{2i}$
- Trigonometric formulae :

$$\sin(\alpha+\beta) = \sin\alpha\,\cos\beta + \cos\alpha\,\sin\beta$$
$$\cos(\alpha+\beta) = \cos\alpha\,\cos\beta - \sin\alpha\,\sin\beta$$
$$\sin(\alpha-\beta) = \sin\alpha\,\cos\beta - \cos\alpha\,\sin\beta$$
$$\cos(\alpha-\beta) = \cos\alpha\,\cos\beta + \sin\alpha\,\sin\beta$$



- Roots:

$$(\cos\alpha + \mathrm{j}\sin\alpha)^n = [\cos(\alpha+2\pi k) + \mathrm{j}\sin(\alpha+2\pi k)]^n$$
$$= \cos(n\alpha + 2\pi nk) + \mathrm{j}\sin(n\alpha + 2\pi nk)$$
$$\text{where} \quad k = 0, \pm1, \pm2, \pm3, \cdots$$

$$\sqrt[n]{x + \mathrm{j}y} = \sqrt[n]{r}\left[\cos\left(\frac{\alpha}{n} + \frac{2\pi}{n}k\right) + \mathrm{j}\sin\left(\frac{k}{n} + \frac{2\pi}{n}k\right)\right]$$
$$\text{where} \quad k = 0, \pm1, \pm2, \cdots$$

- **Polynomials**
  - Polynomial in $\mathbb{C}$ of any degree are split, i.e. α_i, $\beta_i \in \mathbb{N}$ such that $P(X) = \prod_i (X - \alpha_i)^{\beta_i}$

- **Logarithmic**
  - For $(x, y) \in \mathbb{R}^2$, $y = e^x > 0$. So $x = lny$ defined with $y \in \mathbb{R}_+^*$
  - One can define ln on negative numbers using complex numbers: $\ln(-5) = 2\ln(i) + \ln(5) = i\pi + \ln(5) \in \mathbb{C}$

Week 4 -6: Linear Algebra

- Matrices :
  - For two matrices A (kxp) and B (pxn): $(A + B)_{ij} = (A)_{ij} + (B)_{ij}$
    $(\lambda A)_{ij} = \lambda(A)_{ij}, \quad \lambda \in \mathbb{C}$
  - Multiplication : $(AB)_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{il}b_{lj} + \cdots + a_{ip}b_{pj} = \sum_{l=1}^{p} a_{il}b_{lj}$
  - Multiplication is associative but not commutative.
- In the same way that a function of a variable f(x) can be constructed through its Taylor series, functions f(M) of a squared matrix M can be defined through the corresponding Taylor series. Hence for the exponential:

$$\exp(M) = \mathbf{1} + \mathbf{M} + \frac{\mathbf{M}^2}{2} + \cdots = \sum_{n=0}^{\infty} \frac{\mathbf{M}^n}{n!}$$

- **Transpose Matrix**:
  - For $A = (a_{ij})$ a kxn matrix, the transpose matrix of A is the nxk matrix: $A^T = (a_{ji})$.
  - For two matrices A and B with the proper size: $(A^T)^T = A$; $(A + B)^T = A^T + B^T$; $(\alpha A)^T = \alpha A^T$; $(AB)^T = B^T A^T$
- A matrix A is symmetric if $A = A^T$ and anti-symmetric if $A = -A^T$.
- **The trace** of a matrix $A = (a_{ij})$ is the sum of the diagonal terms./ For two square matrices A and B, and α a scalar:
  $tr(A + B) = tr(A) + tr(B); tr(\alpha A) = \alpha tr(A); tr(A) = tr(A^T); tr(AB) = tr(BA)$
- The trace is independent of the basis onto which the operator is defined !

- **Inverse Matrix**: If A is a square matrix (real or complex), B is the inverse of A if AB = BA = I. I is the identity matrix nxn with diagonal coefficients equal to 1, and off-diagonal coefficients equal to 0.
  B is unique! It is also equivalently denoted by A^{-1}. In finite dimensions, it is equivalent to say:
  - A is invertible
  - The equation Ax = b has a unique solution.
  - If A is a square matrix of order n, *rank(A) = n*.
  - The linear application $x \to Ax$ is injective
  - The linear application $x \to Ax$ is surjective

Determinant :

- For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $det(A) = ad - bc$. If $det(A) = 0$, it gives a relation between the two rows or columns that indicates if they are linearly dependent
- Key result: A nxn matrix is invertible if and only if $det(A) \neq 0$
- There are many ways to derive the determinant. A practical one is the Laplace formula:
- Let $A = (a_{ij})$ be a square matrix of order n. Let [A_{ij}] be the submatrix of A obtained by deleting row i and column j. The minor-ij M_{ij} and the cofactor-ij C_{ij} are defined by
  $M_{ij} = det[A_{ij}]$, $C_{ij} = (-1)^{i+j} M_{ij}$, and $det(A) = \sum_{j=1}^{n} a_{ij} C_{ij}$
- For two square matrices: $det(AB) = det(A)\, de\,t(B) = det\,(BA)$
- This is very important as it ensures that the determinant is independent of the basis, so the inversible property is a function of the linear transformation associated to A.

Vector spaces :

$(A^{-1})^{-1} = A;$

$(AB)^{-1} = B^{-1}A^{-1}$

$(A^n)^{-1} = (A^{-1})^n;$

$(\alpha A)^{-1} = \frac{1}{\alpha} A^{-1};$

$(A^T)^{-1} = (A^{-1})^T.$

- A subspace of a vector space V is a subset of V that is also a vector space. To verify that a subset U of V is a subspace you must check that U contains the vector 0 (neutral for addition), and that U is closed under addition and scalar multiplication.
- The **span** of a list of vectors $(v_1, \dots, v_n)$ in V, denoted as $\text{span}(v_1, \dots, v_n)$, is the set of all linear combinations of these vectors:
  $$\text{span}(v_1, \dots, v_n) = \{u \in V,\ \exists (a_1, \dots, a) \in \mathbb{C}^n,\ u = a_1 v_1 + \dots + a_n v_n\}$$
- A vector space V is said to be finite dimensional if it is spanned by some list of vectors in V:

$\exists\ (v_1, \dots, v_n) \in V,\ \forall u \in V, \exists (a_1, \dots, a) \in \mathbb{C}^n,\ u = a_1 v_1 + \dots + a_n v_n$

- If V is not finite dimensional, it is infinite dimensional. In such case, no list of vectors from V can span V.

- A **basis** of V is a list of vectors in V that both spans V and is linearly independent.
  - A list of vectors $(v_1, \dots, v_n)$ is said to be linearly independent if the equation: $a_1 v_1 + \dots + a_n v_n = 0$ has for solution $\forall i,\ a_i = 0$ ie: one cannot express one vector of the set as linear expression of the others.
- A basis of V is a list of vectors in V that both spans V and is linearly independent:
- The dimension of a finite dimensional vector space V is the length of the shortest list of vectors that span V.
- All bases of a finite dimensional vector space have the same length.
- Any list of linearly independent vectors of length n = dimV is a basis of V
- There cannot be a list of n+1 linearly independent vectors in V.

- Let U and V be vector spaces over K ($\mathbb{C}$ *or* $\mathbb{R}$). A function T : V → U is called a linear transformation if, for all u, v ∈ V and α ∈ K:
  - T(u + v) = T(u) + T(v)
  - T(αu) = αT(u)
  - If the image of T is in V, ie if T : V → V, T is called a **linear operator.**
  - To every linear operator T, one can associate a matrix that acts on the vectors of V (finite dimension).
- The notions discussed on matrices above apply to operators:
  - A linear operator T : V → V is said to be injective if Tu = Tv, with u,v ∈ V, implies u = v.
  - T is injective if and only if null(T) = {0}, with the subspace:
  - T is surjective if range(T) = V
  - In infinite dimension, T is bijective if it is injective **and** surjective.
  - In finite dimension, T is bijective and has an inverse if it is injective **or** surjective, just like its associated matrix !
  - dim(range(T)) = rank(T)
    - In finite dimension, it is equivalent to say: The columns (lines) of the associated matrix are linearly independent; the operator is injective; the operator is surjective; The matrix in invertible; $\det\,(A) \neq 0$

- Two operators will commute in terms of their composition, if their associated matrices commute with respect to the multiplication of matrices.
- The commutator [·,·] of two operators X,Y is defined as [X,Y]≡ XY − YX.
- Two operators X, Y commute if [X,Y] = 0.

- The trace and determinant of operators are defined the same way as above, and do not depend on the basis chosen for the associated matrix.
- Eigen values and eigen vectors of operators:
  - An eigen vector u for a linear operator T is a vector that satisfies $Tu = \lambda u$. λ is called an eigen value.
  - For a given eigenvalue λ, there maybe several linearly independent eigen vectors of T. We say that λ generates a sub-space of a given dimension > 1.
    The eigenvalue is then said to be degenerate.
  - The set of eigenvalues of T is called the spectrum of T.
  - Set of eigenvectors of T corresponding to λ = null(T − λI).
  - The eigen values are found solving det(T − λI) = 0.

- A matrix A is **diagonalizable** if it is similar to a diagonal matrix, i.e. there exist an invertible matrix P, and a diagonal matrix D, such that $P^{-1}AP = D$.
  - Equivalently, A is diagonalizable if there exist a basis of its eigen vectors.
  - The associated linear operator T is diagonalizable if there is a basis of the vectorial space V formed by the eigenvectors of T.
  - A matrix nxn with n distinct and non-zero eigenvalues is diagonalizable.
  - If the dimension of the sub-spaces of the eigen values of A (nxn) add up to n, then it is diagonalizable.

- Let T be a linear operator, and assume $\lambda_1,...\lambda_n$ are distinct eigenvalues of T and $u_1,...u_n$ are corresponding nonzero eigenvectors. Then $(u_1,...u_n)$ are linearly independent.
- Inner Product: an inner product on a vector space V over $\mathbb{R}$ *or* $\mathbb{C}$ is a map from an ordered pair (u, v) of vectors in V to a number $\langle u|v\rangle$ in $\mathbb{R}$ *or* $\mathbb{C}$. The axioms for $\langle u|v\rangle$ are inspired by the axioms for the dot product of vectors:
  1. $\langle v|v\rangle \geq 0$, for all vectors $v \in$ V.
  2. $\langle v|v\rangle = 0$ if and only if $v = 0$.
  3. $\langle u|v_1 + v_2\rangle = \langle u|v_1\rangle + \langle u|v_2\rangle$. Additivity in the second entry.
  4. $\langle u|\alpha v\rangle = \alpha \langle u|v\rangle$, $\alpha \in \mathbb{C}$. Homogeneity in the second entry.
  5. $\langle u|v\rangle = \langle v|u\rangle^*$. Conjugate exchange symmetry.
  - The norm of a vector is also noted: $|v|^2 = \langle v|v\rangle \geq 0$
  - Dirac notation: ket $|v\rangle$ is a vector; bra $\langle v|$is a linear operator acting on a vector via the dot product.
  - Two vectors are orthogonal if $\langle u|v\rangle = 0$.
  - Schwartz inequality: $|\langle u|v\rangle| \leq |u||v|$
  - A list of vectors is said to be orthonormal if all vectors have norm one and are pairwise orthogonal. A set of orthonormal vectors are necessarily linearly independent.

Hilbert Spaces

- A Hilbert space H is a real or complex **inner product** space that is also a **complete metric** space with respect to the distance function induced by the inner product. Inner product space is simply a vectorial space with an inner product.
- A complete metric is the property that every Cauchy sequence of H with respect to the metric converges in H.
- We consider a linear operator T on a vector space V that has an inner product. The linear operator $T^\dagger$ on V called the **adjoint** of T, is constructed as follow: for u,v vectors of V:
  - $T^\dagger$ is a linear operator:
  - $\langle u, Tv\rangle = \langle T^\dagger u, v\rangle$
  - For T and S two linear operators: $(ST)^\dagger = T^\dagger S^\dagger$
  - The adjoint of the adjoint is the original operator: $(S^\dagger)^\dagger = S$
  - $(T^\dagger)_{ij} = (T_{ji})^*$: over an orthonormal basis, the adjoint matrix is the transpose and complex conjugate.
- **Self-adjoint** (or Hermitian in finite dimension) operators are linear operators T for which $T = T^\dagger$.
- One can show that: $T = T^\dagger$ if and only if $\forall v \in V, \langle v, Tv\rangle \in \mathbb{R}$
- Two other very important results:
  - The eigenvalues of Hermitian operators are real;
  - Different eigenvalues of a Hermitian operator correspond to orthogonal eigenfunctions:

- An operator U in a complex vector space V is said to be **a unitary operator** if it is surjective and does not change the magnitude of the vector it acts upon.
- A more common definition: $U^\dagger U = UU^\dagger = I$
- Unitary operators preserve inner products in the following sense: $\langle Uu, Uv\rangle = \langle u, v\rangle$

- **Spectral theorem (finite dimension):**
  If A is a Hermitian operator on a Hilbert space V of finite dimension, then there exists an orthonormal basis of V consisting of eigenvectors of A. Each eigenvalue is real.

  - This is equivalent to say that A can be diagonalized;

  - It is also equivalent to the fact that the sub-spaces of the eigenvalues of V are orthogonal, and the sum of their dimension is equal to dim(V).
  - The spectral theorem actually applies to Normal operators, defined as operators for which $[T,T^{\dagger}] = 0$. This includes self-adjoint and unitary matrices.

- **Spectral theorem (infinite dimension)**:
  In infinite dimension, the problem is more complex and the theorem holds only in certain conditions (that are almost always met in QM). It applies to certain types of operators: Compact self-adjoint operators; Bounded self-adjoint operators.

- **Spectral decomposition**: in finite dimension, a self-adjoint operator can be diagonalized, hence possess a set of orthonormal eigenvectors that form a basis. If a_α are its eigenvalues, that can be degenerate, hence span a sub-space of dimension n_α and eigenvectors $|\alpha, r_\alpha\rangle$, one can write: $\hat{A} = \sum_\alpha \sum_{r_\alpha=1}^{n_\alpha} a_\alpha |\alpha, r_\alpha\rangle \langle\alpha, r_\alpha|$
- This is based on the concept of outer product which is an operator $|\psi\rangle \langle\varphi|$. For an orthogonal basis, $\widehat{P_\alpha} = \sum_{r_\alpha=1}^{n_\alpha} |\alpha, r_\alpha\rangle\langle\alpha, r_\alpha|$ is a projector on the sub-space of a_α.
- For an object in state $|\psi\rangle$, the probability to find an eigen value a_α of an observable $\widehat{\boldsymbol{A}}$ is given by:
  $P(a_\alpha) = \langle\psi|\widehat{P_\alpha}|\psi\rangle = \sum_{r_\alpha=1}^{n_\alpha} |\langle\alpha, r_\alpha|\psi\rangle|^2$, where n_α is the dimension of the sub-space generated by a_α, and the $|\alpha, r_\alpha\rangle$ the associated orthonormal eigenvectors.

- **Commuting observables**:
  - If two normal operators commute on a Hilbert space, there exists a basis of common eigenvectors.
  - This is quite powerful and is used for example in the quantum numbers of orbitals in the Hydrogen atom, or to prove the Bloch theorem.

Week 7-9: Functions

- Given two sets of real numbers, a domain (often referred to as the x-values, and interval *I*) and a co-domain (often referred to as the y-values), a real function assigns to each x-value a *unique* y-value.
- *Injective function*: function *f* that maps distinct elements of its domain to distinct elements: $f(x_1) = f(x_2)$ implies $x_1 = x_2$.
- Surjective functions: a function *f* such that every element y can be mapped from element *x* so that $f(x) = y$.
- Composition: if a function f is defined from I to X, and g is defined over X, one can define $\forall x \in I,\ h(x) = gof(x)$.
- f^{-1} is the inverse of f and is defined such that $f^{-1}of = fof^{-1} = I_d$ *(the identity function).*
- A function is even (odd) *if* $\forall x \in I,\ f(x) = f(-x)\ (f(x) = -f(-x))$
- Periodicity: *f* is periodic of period *T* if $\forall x \in I,\ f(x+T) = f(x)$.

Sequences

- Functions are extension of the concept of sequences that can be seen as functions from the domain $\mathbb{N}$ into $\mathbb{R}$ or $\mathbb{C}$.
- Examples: $u_n = u_{n-1} + r = u_0 + rn$ (arithmetic sequence) ; $u_n = ru_{n-1} = u_0 r^n$ (geometric sequence)
- A sequence converges towards a limit $l \in \mathbb{R}\ (or\ \mathbb{C})$ if and only if: $\forall\varepsilon > 0,\ \exists\ N \in \mathbb{N}$ such that $(n \geq N\ \Rightarrow \lceil u_n - l| \leq \varepsilon)$
  - This limit is unique;
  - It is equivalent to say that $(u_n - l)$ converges to 0.
- A sequence **tends to** $+\infty$ if and only if: $\forall\ A \in \mathbb{R}_+^*$, $\exists\ N \in \mathbb{N}$ such that $\forall\ n \in \mathbb{N}$, $(n\ \geq N\ \Rightarrow u_n \geq A)$
  - A sequence **tends to** $-\infty$ if and only if: $\forall\ B \in \mathbb{R}_-^*$, $\exists\ N \in \mathbb{N}$ such that $\forall\ n \in \mathbb{N}$, $(n\ \geq N\ \Rightarrow u_n \leq \mathrm{B})$
  - A sequence $(v_n)_{n\in\mathbb{N}}$ **diverges** if it does not converge nor tend to $\pm\infty$

Other important definitions and results:

- A sequence is increasing if : $\forall\ n \in \mathbb{N}, u_{n+1} \geq u_n$
- A sequence is decreasing if : $\forall\ n \in \mathbb{N},\ u_{n+1} \leq u_n$
- If a sequence $(v_n)_{n\in\mathbb{N}}$ is increasing (decreasing) and has no upper bound (lower bound), then it diverges to $+\infty$ $(-\infty)$. (you showed it in exercises week 1).
- If a sequence $(v_n)_{n\in\mathbb{N}}$ is increasing (decreasing) and has an upper bound l (lower bound), then it converges towards l.
- Squeeze (or sandwich) theorem: If (a_n), (b_n), and (c_n) are three real-valued sequences satisfying $a_n \leq b_n \leq c_n$ for all n, and if furthermore $a_n \to \ell$ and $c_n \to \ell$, then $b_n \to \ell$.

Functions:

- A function is increasing if $\forall(x_1, x_2) \in I^2,\ \ x_1 \geq x_2 \Longrightarrow f(x_1) \geq f(x_2)$
- A function is decreasing if $\forall(x_1, x_2) \in I^2,\ \ x_1 \geq x_2 \Longrightarrow f(x_1) \leq f(x_2)$
- A function $f: I \to \mathbb{R}$ with *I* including $+\infty$, admits l for limit when x goes to infinity if and only if
  $$\forall\varepsilon > 0,\ \exists A > 0,\ \forall x \in I,\ (x \geq A \Longrightarrow |f(x) - l| < \varepsilon)$$
- A function $f: I \to \mathbb{R}$ with *I* including $+\infty$, admits $+\infty$ for limit when x goes to infinity if and only if
  $$\forall A > 0,\ \exists A' > 0,\ \forall x \in I,\ (x \geq A' \Longrightarrow f(x) \geq A)$$
- A function $f: I \to \mathbb{R}$ (or other domain) admits l for limit in a point $a \in I$ if and only if
  For all sequence $(u_n)_{n\in\mathbb{N}}$ such that $\lim_{n\to\infty} u_n = a$, $\lim_{n\to\infty} f(u_n) = l$.

- One can express this without sequences: $\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in I,\ |x-a| < \alpha \Rightarrow |f(x) - l| < \varepsilon$
- Divergence to $+\infty$: $\forall A > 0,\ \exists \alpha > 0,\ \forall x \in I,\ |x-a| < \alpha \Rightarrow f(x) \geq A$
- Divergence to $-\infty$: $\forall B < 0,\ \exists \alpha > 0,\ \forall x \in I,\ |x-a| < \alpha \Rightarrow f(x) \leq B$
- If f is increasing (decreasing) and has an upper bound (lower bound), then it converges.
- If f is increasing (decreasing) and has no upper bound (lower bound), then it tends to $+\infty$ $(-\infty)$.

- $f: I \to \mathbb{R}$ has a right limit l at $a \in I$ if: $\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in I,\ 0 < x - a \leq \alpha \Rightarrow |f(x) - l| < \varepsilon$
  Notation: $\lim\limits_{x \to a^+} f(x) = l$
- $f: I \to \mathbb{R}$ has a left limit l at $a \in I$ if: $\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in I,\ 0 < a - x \leq \alpha \Rightarrow |f(x) - l| < \varepsilon$
  Notation: $\lim\limits_{x \to a^-} f(x) = l$

- For $(\lambda, l, l') \in \mathbb{C}^3,\ f, g: I \to \mathbb{R}$ f,g admit l and l' as limit at a point $a \in I$ respectively:
-

$$f(x) \xrightarrow[x \to a]{} l \Longrightarrow |f(x)| \xrightarrow[x \to a]{} |l|$$

$$f(x) \xrightarrow[x \to a]{} 0 \Longleftrightarrow |f(x)| \xrightarrow[x \to a]{} 0$$

$$\left.\begin{array}{l} f(x) \xrightarrow[x \to a]{} l \\ g(x) \xrightarrow[x \to a]{} l' \end{array}\right\} \Longrightarrow f(x) + g(x) \xrightarrow[x \to a]{} l + l'$$

$$f(x) \xrightarrow[x \to a]{} l \Longrightarrow \lambda f(x) \xrightarrow[x \to a]{} \lambda l$$

$$\left\{\begin{array}{l} f(x) \xrightarrow[x \to a]{} 0 \\ \text{g is bounded around a} \end{array}\right| \Longrightarrow f(x)g(x) \xrightarrow[x \to a]{} 0$$

$$\left.\begin{array}{l} f(x) \xrightarrow[x \to a]{} l \\ g(x) \xrightarrow[x \to a]{} l' \end{array}\right\} \Longrightarrow f(x)g(x) \xrightarrow[x \to a]{} ll'$$

$$\left.\begin{array}{l} g(x) \xrightarrow[x \to a]{} l' \\ l' \neq 0 \end{array}\right\} \Longrightarrow \frac{1}{g(x)} \xrightarrow[x \to a]{} \frac{1}{l'}$$

$$\left.\begin{array}{l} f(x) \xrightarrow[x \to a]{} l \\ g(x) \xrightarrow[x \to a]{} l' \\ l' \neq 0 \end{array}\right\} \Longrightarrow \frac{f(x)}{g(x)} \xrightarrow[x \to a]{} \frac{l}{l'}.$$

- If f is complex, then:

$$f: I \longrightarrow \mathbb{C},\ (\alpha, \beta) \in \mathbb{R}^2$$

$$f(x) \xrightarrow[x \to a]{} \alpha + i\beta \Longleftrightarrow \left\{\begin{array}{l} (\text{Ré } f)(x) \xrightarrow[x \to a]{} \alpha \\ (\text{Im } f)(x) \xrightarrow[x \to a]{} \beta. \end{array}\right.$$

- $f: I \to \mathbb{R}$ with $I \subset \mathbb{R}$, is continuous at the point $x_0 \in I$ if: $\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall x \in I,\ |x - x_0| < \alpha \Rightarrow |f(x) - f(x_0)| < \varepsilon$
- It is equivalent to say that f is continuous at point $x_0 \in I$ if and only if **f has a right and left limit at x_0 and the limits are equal.**
- Definition with sequences: A function $f: I \to \mathbb{R}$ (or other domain) admits l for limit in a point $a \in I$ if and only if :
  For all sequence $(u_n)_{n \in \mathbb{N}}$ such that $\lim\limits_{n \to \infty} u_n = a$, $\lim\limits_{n \to \infty} f(u_n) = f(a)$
- Important results: if *f* and *g* are two continuous functions over an interval I:
  - $|f|$ is continuous ; $f + g$ is also continuous over I; $\lambda f,\ \lambda \in \mathbb{R}\ or\ \mathbb{C}$, is continuous; $f \times g$ is continuous; If $g \neq 0$ over I, f/g is continuous; If g is continuous over $f(I)$, $h(x) = gof(x)$ is continuous; f^{-1}, if defined, is continuous over $f(I)$.
  - If f is complex , it is continuous if and only if its real and imaginary parts are.

- A function $f: I \to \mathbb{R}$ with $I \subset \mathbb{R}$, is differentiable at $x \in I$ if:

$$\forall \varepsilon > 0,\ \exists \alpha > 0,\ \forall h \in I,\ |h| < \alpha \Rightarrow \left|\frac{f(x+h) - f(x)}{h} - l\right| < \varepsilon, \quad l = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = f'(x)$$

Differentiability

- A function f as defined earlier can be right and / or left differentiable if $\frac{f(x+h) - f(x)}{h}$ admits a right and left limit respectively.
- **Corollary**: f is differentiable at $a \in I$ if it is right and left differentiable, and the values are equal.
- If a function is differentiable at point a, it is continuous at a. The reverse is not true ! (i.e. continuity does not imply differentiability)!
- Important immediate results:
  - f is increasing (decreasing) over a domain *I* if and only if $\forall x \in I,\ f'(x) > 0\ (f'(x) < 0)$.
  - If $f: \mathbb{R} \to \mathbb{R}$ is monotonic over a segment $[a, b] \subset \mathbb{R}$, it then takes all the values within $[\inf(f(a), f(b)),\ \sup(f(a), f(b))]$.

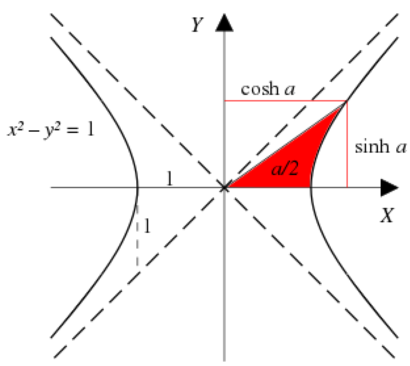
- Operations on derivatives:

General rules	Function $y = f(x)$	Derivative $y' = f'(x)$
1. Constant factor	$y = cf(x)$	$y' = cf'(x)$
2. Sum (algebraic) rule	$y = u(x) + v(x)$	$y' = u'(x) + v'(x)$
3. Product rule	$y = u(x)v(x)$	$y' = u'(x)v(x) + u(x)v'(x)$
4. Quotient rule	$y = \frac{u(x)}{v(x)}$	$y' = \frac{u'(x)v(x) - u(x)v'(x)}{v(x)^2}$
5. Chain rule	$y = f[g(x)]$	$y' = \frac{\mathrm{d}f}{\mathrm{d}g} g'(x)$
6. Inverse functions	$y = f^{-1}(x)$ i.e. $x = f(y)$	$y' = \frac{1}{\mathrm{d}x/\mathrm{d}y} = \frac{1}{f'(y)}$

- Common functions:
  - A *power function* is a function that can be represented in the form $f(x) = kx^{\alpha}$, where k and α are real numbers, it is a continuous functions and can be differentiated until the derivative is null: $\forall \alpha \in \mathbb{R},\ f'(x) = \alpha k x^{\alpha-1}$
  - Exponential functions : function of the form $f\colon \mathbb{R}\ (or\ \mathbb{C}) \to \mathbb{R}\ (or\ \mathbb{C})\quad f(x) = a^x$
- From the fundamental definition of the differentiability of a function, we can find the derivative of exponential functions, and find a number e for which $(e^x)' = e^x$
- e is defined as: $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$
- The Hôpital rule: f and g are two functions, differentiable over an interval I, not necessarily at c;
  - g' is not zero around c (for all $x \neq c$)
  - We have : $\lim_{x\to c} f(x) = \lim_{x\to c} g(x) = 0\ or\ \pm\infty$
  - $\lim_{x\to c} \frac{f'(x)}{g'(x)}$ exists and: $\lim_{x\to c} \frac{f(x)}{g(x)} = \lim_{x\to c} \frac{f'(x)}{g'(x)}$
- The rule also applies for $x \to \infty$
- The Rolle theorem: if f is a function defined over $[a, b] \subset \mathbb{R}$, continuous and differentiable, and if f(a) = f(b), then $\exists c \in]a, b[,\ f'(c) = 0$.
-
- Cauchy's mean value theorem: If f, g are two functions defined over $[a, b] \subset \mathbb{R}$, continuous over $[a, b]$ and differentiable over $]a, b[$, then $\exists c \in]a, b[$, such that:
-

$$\big(f(b) - f(a)\big)g'(c) = \big(g(b) - g(a)\big)f'(c)$$

- Hyperbolic functions:
  - $cosh(x) = \frac{e^x + e^{-x}}{2}$
  - $\sinh(x) = \frac{e^x - e^{-x}}{2}$
  - $tanh(x) = \frac{\sinh(x)}{\cosh(x)}$
  - $cosh^2(x) - sinh^2(x) = 1$
  - $\frac{d}{dx}(cosh^{-1}(x)) = \frac{1}{\sqrt{x^2-1}}$



- Common derivatives:

Derivatives of fundamental functions	Function $y = f(x)$	Derivative $y' = f'(x)$
1. Constant factor	$y = \text{constant}$	$y' = 0$
2. Power function	$y = x^n$	$y' = nx^{n-1}$
3. Trigonometric functions	$y = \sin x$	$y' = \cos x$
	$y = \cos x$	$y' = -\sin x$
	$y = \tan x$	$y' = \frac{1}{\cos^2 x} = 1 + \tan^2 x$
	$y = \cot x$	$y' = \frac{-1}{\sin^2 x} = -1 - \cot^2 x$
4. Inverse trigonometric functions	$y = \sin^{-1} x$	$y' = \frac{1}{\sqrt{1-x^2}}$
	$y = \cos^{-1} x$	$y' = -\frac{1}{\sqrt{1-x^2}}$
	$y = \tan^{-1} x$	$y' = \frac{1}{1+x^2}$
	$y = \cot^{-1} x$	$y' = -\frac{1}{1+x^2}$

Derivatives of fundamental functions	Function $y = f(x)$	Derivative $y' = f'(x)$
5. Exponential function	$y = e^x$	$y' = e^x$
Logarithmic function	$y = \ln x$	$y' = \frac{1}{x}$
6. Hyperbolic trigonometric functions	$y = \sinh x$	$y' = \cosh x$
	$y = \cosh x$	$y' = \sinh x$
	$y = \tanh x$	$y' = \frac{1}{\cosh^2 x} = 1 - \tanh^2 x$
	$y = \coth x$	$y' = \frac{1}{\sinh^2 x} = 1 - \coth^2 x$
7. Inverse hyperbolic trigonometric functions	$y = \sinh^{-1} x$	$y' = \frac{1}{\sqrt{1+x^2}}$
	$y = \cosh^{-1} x$	$y' = \frac{1}{\sqrt{x^2-1}}$ $(x > 1)$
	$y = \tanh^{-1} x$	$y' = \frac{1}{1-x^2}$ $(\lvert x \rvert < 1)$
	$y = \coth^{-1} x$	$y' = -\frac{1}{x^2-1}$ $(\lvert x \rvert > 1)$

Extremums:

- For a function to have an extremum at a point x_0, it is necessary that $f'(x_0) = 0$. It is however not sufficient. It must also be such that $f''(x_0) > 0$ (convex) or $f''(x_0) < 0$ (concave).
- Inflexion point: $f''(x_0) = 0$, marking where the concavity of a function changes. We must also have $f'''(x_0) \neq 0$

Taylor Series:

- Reminder: $f(x + \Delta x) = f(x) + f'(x)\Delta x + \Delta x h(x)$ with $\lim_{\Delta x \to 0} h(x + \Delta x) = 0$
- Taylor-Lagrange: for a function at least n+1 times differentiable ($n \in \mathbb{N}$), defined over an interval $[a, b] \subset \mathbb{R}$, (the (n+1)th derivative needs to exist only in $]a, b[$), then $\exists c \in\,]a, b[$ such that:

$$f(b) = f(a) + (b-a)f'(a) + \frac{(b-a)^2}{2!}f''(a) + \dots + \frac{(b-a)^n}{n!}f^{(n)}(a) + \frac{(b-a)^{n+1}}{(n+1)!}f^{(n+1)}(c)$$

- Let's consider the domain of definition of f, $I \subset \mathbb{R}$, that includes 0, and a arbitrary point x in this interval. We can re-write the Taylor Lagrange polynomial what is called the Maclaurin form (with $c \in]0, x[$):

$$\forall x \in I,\ f(x) = \sum_{k=0}^{n} \frac{x^k}{k!} f^{(k)}(0) + \frac{x^{n+1}}{(n+1)!} f^{(n+1)}(c)$$

$R_n(x) = \frac{x^{n+1}}{(n+1)!} f^{(n+1)}(c)$ is called the remainder of the Taylor polynomial $\sum_{k=0}^{n} \frac{x^k}{k!} f^{(k)}(x)$.

- This remainder is small, and hence the function is well approximated by the Taylor polynomial, in two situations:
  - *Taylor Expansion* : x is close to 0, for all n, the polynomial is a local approximation of the function around 0. The approximation globally improves as the degree of the polynomial increases for small x.
  - *Taylor Series* : n *is large* $(n \to \infty)$, $f$or all x, the polynomial is a global approximation of the function over a certain domain where the series $\sum_{k=0}^{\infty} \frac{x^k}{k!} f^{(k)}(x)$ converges.
- There are different tests that can assess the convergence of a series: Ratio test: one looks at the behavior of the ratio of two following sequence number in the series as n goes to infinity.

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots + a_n x^n + a_{n+1} x^{n+1} + \cdots$$

  - the ratio is : $\frac{a_{n+1} x^{n+1}}{a_n x^n}$
- Taking the limit:

$$\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} x \right| = \frac{|x|}{R} \qquad \text{where} \quad R = \lim_{n \to \infty} \left| \frac{a_n}{a_{n+1}} \right|$$

- The series is absolutely convergent if |x| < R and divergent if |x| > R. Hence a power series is convergent in a definite interval (−R,R) and divergent outside this interval.
- Other convergence tests exist like the Cauchy-Hadamar: $\frac{1}{R} = \lim_{n \to \infty} \sqrt[n]{|a_n|}$

- Examples of Maclaurin series valid over $\mathbb{R}$:

$$\cos(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + \ldots$$

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \ldots$$

$$\arctan(x) = x - \frac{x^3}{3} + \frac{x^5}{5} - \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + \ldots$$

$$\arcsin(x) = x + \frac{1}{2}\frac{x^3}{3} + \frac{1.3}{2.4}\frac{x^5}{5} + \ldots + \frac{1.3\ldots(2n-1)}{2.4\ldots(2n)}\frac{x^{2n+1}}{2n+1} + \ldots$$

$$\exp(x) = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + \ldots$$

$$\cosh(x) = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + \frac{x^{2n}}{(2n)!} + \ldots$$

$$\sinh(x) = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + \frac{x^{2n+1}}{(2n+1)!} + \ldots$$

$$(1+x)^\alpha = 1 + \frac{\alpha}{1!}x + \frac{\alpha(\alpha-1)}{2!}x^2 + \ldots + \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!}x^n + \ldots$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots + (-1)^{n-1}\frac{x^n}{n} + \ldots$$

Taylor expansion **around 0 at the order n**:

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + o(x^n)$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + (-1)^p \frac{x^{2p+1}}{(2p+1)!} + o(x^{2p+2})$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + (-1)^p \frac{x^{2p}}{(2p)!} + o(x^{2p+1})$$

$$\sinh x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{x^{2p+1}}{(2p+1)!} + o(x^{2p+2})$$

$$\cosh x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2p}}{(2p)!} + o(x^{2p+1})$$

Primitives and Integrals

- Given F and f two functions continuous and differentiable over $I \subset \mathbb{R}$, F is a primitive of f if : $\forall x \in I,\ F'(x) = f(x)$
- If F is a primitive of f, $\forall \lambda \in \mathbb{R}$ *or* $\mathbb{C}$, $F + \lambda$ is a primitive of f.
- Fundamental theorem: F and f two functions continuous and differentiable over $[a, b] \subset \mathbb{R}$, F primitive of f. The area under the curve $f(x),\ x \in [a, b]$ is written:

$$F(b) - F(a) = \int_a^b f(x)dx$$

- Let f be a continuous real-value function defined on a closed interval [a, b]. Let F be the function defined, for all x in [a, b], by $F(x) = \int_0^x f(t)dt$. Then F is uniformly continuous on [a, b] and differentiable on the open interval (a, b), and $F'(x) = f(x)$ for all x in (a, b) so F is an antiderivative (or primitive) of f.
- The form expressed above is an indefinite form, also written $\int f(x)dx$. Definite forms is an integral over a defined interval that returns a number.
- Every continuous function has an anti-derivative, actually an infinity of them shifted by a constant.
- Integration by parts : $\int_a^b u(x)v'(x)dx = [u(x)v(x)]_a^b - \int_a^b u'(x)v(x)dx$
- Substitution : $\int_a^b f(g(x))dx = \int_{g(a)}^{g(b)} f(u)\frac{du}{g'}$ with $u = g(x)$

- Common primitives

$f(x)$	$\int f(x)\,dx$	$f(x)$	$\int f(x)\,dx$
c	cx	$\frac{1}{x^2+a^2}$	$\frac{1}{a}\tan^{-1}\frac{x}{a}$ or $\frac{-1}{a}\cot^{-1}\frac{x}{a}$
x^n	$\frac{x^{n+1}}{n+1} \quad (n \neq -1)$	$\frac{1}{x^2+2ax+b}$	$\frac{1}{\sqrt{b-a^2}}\tan^{-1}\left(\frac{x+a}{\sqrt{b-a^2}}\right)$ $(b>a^2)$
$\frac{1}{x}$	$\ln\lvert x\rvert \quad (x \neq 0)$	$\frac{2x+a}{x^2+ax+b}$	$\ln\lvert x^2+ax+b\rvert$
e^x	e^x		
a^x	$\frac{a^x}{\ln a} \quad \begin{pmatrix} a>0 \\ a\neq 1 \end{pmatrix}$	$\sqrt{ax+b}$	$\frac{2}{3a}\sqrt{(ax+b)^3}$
$\ln x$	$x\ln x - x \quad (x>0)$	$\frac{1}{\sqrt{ax+b}}$	$\frac{2}{a}\sqrt{ax+b}$
$\frac{1}{x-a}$	$\ln\lvert x-a\rvert$	$\frac{1}{\sqrt{a^2-x^2}}$	$\sin^{-1}\frac{x}{a}$
$\frac{1}{(x-a)^2}$	$-\frac{1}{x-a}$	$\sqrt{a^2-x^2}$	$\frac{x}{2}\sqrt{a^2-x^2}+\frac{a^2}{2}\sin^{-1}\frac{x}{a}$
$\frac{1}{x^2-a^2}$	$\frac{1}{2a}\ln\left\lvert\frac{x-a}{x+a}\right\rvert = \begin{cases} \frac{-1}{a}\tanh^{-1}\frac{x}{a}, & \lvert x\rvert<\lvert a\rvert \\ \frac{-1}{a}\coth^{-1}\frac{x}{a}, & \lvert x\rvert>\lvert a\rvert \end{cases}$		
$\cos^2 x$	$\frac{1}{2}(x+\sin x\cos x)=\frac{1}{2}\left(x+\frac{\sin 2x}{2}\right)$	$\cot^{-1} x$	$x\cot^{-1}x+\ln\sqrt{1+x^2}$
$\frac{1}{\cos x}$	$\ln\left\lvert\tan\left(\frac{x}{2}+\frac{\pi}{4}\right)\right\rvert$	$\sinh x$	$\cosh x$
$\frac{1}{\cos^2 x}$	$\tan x$	$\cosh x$	$\sinh x$
		$\tanh x$	$\ln\lvert\cosh x\rvert$
$\frac{1}{1+\sin x}$	$\tan\left(\frac{x}{2}-\frac{\pi}{4}\right)$	$\coth x$	$\ln\lvert\sinh x\rvert$
		$\sinh^{-1} x$	$x\sinh^{-1}x-\sqrt{x^2+1}$
		$\cosh^{-1} x$	$x\cosh^{-1}x-\sqrt{x^2-1}$
		$\tanh^{-1} x$	$x\tanh^{-1}x+\ln\sqrt{1-x^2}$
		$\coth^{-1} x$	$x\coth^{-1}x+\ln\sqrt{x^2-1}$
$\frac{1}{\sqrt{x^2+a^2}}$	$\ln\left(\frac{x+\sqrt{x^2+a^2}}{\lvert a\rvert}\right)=\sinh^{-1}\frac{x}{a}$	$\frac{1}{1-\sin x}$	$-\cot\left(\frac{x}{2}-\frac{\pi}{4}\right)=\tan\left(\frac{x}{2}+\frac{\pi}{4}\right)$
$\sqrt{x^2+a^2}$	$\frac{x}{2}\sqrt{x^2+a^2}+\frac{a^2}{2}\ln(x+\sqrt{a^2+x^2})$	$\frac{1}{1+\cos x}$	$\tan\frac{x}{2}$
$\frac{1}{\sqrt{x^2-a^2}}$	$\ln\left\lvert\frac{x+\sqrt{x^2-a^2}}{a}\right\rvert=\cosh^{-1}\frac{x}{a}$	$\frac{1}{1-\cos x}$	$-\cot\frac{x}{2}$
		$\tan x$	$-\ln\lvert\cos x\rvert$
$\sin x$	$-\cos x$	$\tan^2 x$	$\tan x - x$
$\sin^2 x$	$\frac{1}{2}(x-\sin x\cos x)=\frac{1}{2}\left(x-\frac{\sin 2x}{2}\right)$	$\cot x$	$\ln\lvert\sin x\rvert$
$\frac{1}{\sin x}$	$\ln\left\lvert\tan\frac{x}{2}\right\rvert$	$\cot^2 x$	$-\cot x - x$
$\frac{1}{\sin^2 x}$	$-\cot x$	$\sin^{-1} x$	$x\sin^{-1}x+\sqrt{1-x^2}$
		$\cos^{-1} x$	$x\cos^{-1}x-\sqrt{1-x^2}$
$\cos x$	$\sin x$	$\tan^{-1} x$	$x\tan^{-1}x-\ln\sqrt{1+x^2}$

- Calculation of Arc length : $s=\int_a^b\sqrt{1+\left(\frac{dy}{dx}\right)^2}\,dx=\int_a^b\sqrt{1+f'(x)^2}\,dx$

Exact and Inexact differentials

- Partial differentiation: Multi-variable functions will be studied usually by looking at how they vary when changing only one variable at a time:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \frac{d}{dx} f(x, y_0)\Big|_{x=x_0} \qquad \frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}$$

$$\frac{\partial f}{\partial y}(x_0, y_0) = \frac{d}{dy} f(x_0, y)\Big|_{y=y_0}$$

- Higher order Partial differentiation : since partial derivatives of a function are also functions of several variables, they can be differentiated with respect to any variable. For a function of two variables:

$$\frac{\partial f}{\partial x} \longmapsto \frac{\partial}{\partial x}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial x^2}, \quad \frac{\partial}{\partial y}\frac{\partial f}{\partial x} = \frac{\partial^2 f}{\partial y \, \partial x}$$

$$\frac{\partial f}{\partial y} \longmapsto \frac{\partial}{\partial x}\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial x \, \partial y}, \quad \frac{\partial}{\partial y}\frac{\partial f}{\partial y} = \frac{\partial^2 f}{\partial y^2}$$

- Differentiability: a function f defined on an open set I of $\mathbb{R}^n$,f is differentiable in I if all its partial derivatives exist and are continuous.
- Clairaut's theorem: for a function f defined on an open set I of $\mathbb{R}^2$ if all the partial derivatives of f exist and are continuous, then: $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$
- Total differential: a differentiable function (in 2D) has a total differential defined as: $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$
- In some open domain of a space, a differential form $P(x,y)dx + Q(x,y)dy$ where P and Q are continuous two variable functions, is an *exact differential* if it is equal to the total differential of a differentiable function f: $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$ with $\frac{\partial f}{\partial x} = P(x,y)$ and $\frac{\partial f}{\partial y} = Q(x,y)$, in an orthogonal coordinate system.

  Partial derivative test: a differential form $P(x,y)dx + Q(x,y)dy$ is an *exact differential* if and only if $\frac{\partial P}{\partial y} = \frac{\partial Q}{\partial x}$.

Changing the order of limits: For a function of two variables f(x,y), we can invert the order of limits if and only if:

$$\left(\left(\lim_{x \to x_0} f(x,y) = g(y) \text{ uniformly} \right) \wedge \left(\lim_{y \to y_0} f(x,y) = h(x) \right) \right)$$

$$\implies \left(\lim_{x \to x_0} \lim_{y \to y_0} f(x,y) = \lim_{y \to y_0} \lim_{x \to x_0} f(x,y) = \lim_{(x,y) \to (x_0,y_0)} f(x,y) \right)$$

- Uniform convergence :

$$(\forall \epsilon > 0)\,(\exists \delta > 0)\,(\forall x \in D_x)$$
$$(0 < |y - y_0| < \delta \implies |f(x,y) - g(x)| < \epsilon)$$

Changing the order of limit and integral:

If $(f_n)_{n \in \mathbb{N}^*}$is a sequence of Riemann integrable functions defined on a compact interval I, (a close interval in $\mathbb{R}$ for example) which uniformly converge with limit f, then f is Riemann integrable and its integral can be computed as the limit of the integrals of the f_n: $\int_{x \in I} f = \lim_{n \to \infty} \int_{x \in I} f_n$

Changing the order of integrals:

- Fubini's theorem : one may switch the order of integration if the double integral yields a finite answer when the integrand is replaced by its absolute value.

$$\iint_{X \times Y} f(x,y)\, \mathrm{d}(x,y) = \int_X \left(\int_Y f(x,y)\, \mathrm{d}y \right) \mathrm{d}x = \int_Y \left(\int_X f(x,y)\, \mathrm{d}x \right) \mathrm{d}y \quad \text{if} \quad \iint_{X \times Y} |f(x,y)|\, \mathrm{d}(x,y) < +\infty.$$

<u>Weeks 9&10: Fourier and Laplace transforms</u>

- For $x(t)$ a periodic function of period T, and $\omega_0 = \frac{2\pi}{T}$, $\quad x(t) = x(t+T) = \sum_{k=-\infty}^{\infty} a_k e^{j\omega_0 kt}$
- The $(a_k)_{k \in \mathbb{Z}}$ are the Fourier coefficients and the series is called the Fourier series.
- $\forall k, \ a_k = \frac{1}{T} \int_0^T x(t) e^{-jk\omega_0 t}\, dt$
- One can show that all T-periodic functions (and so in particular for regular continuous functions we handle in engineering most of the time), the expression of Fourier series exist, and converges towards the original function uniformly.
- For real functions, one often uses the relation: $\quad f(z) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nz + b_n \sin nz)$

- Parseval equality: $\sum_{-\infty}^{+\infty}|a_k|^2 = \frac{1}{T}\int_0^T |x(t)|^2\, dt$
  If a function f has all its Fourier coefficient equal to zero, the function is zero.
- For a function f that is wholly or piece-wise continuous and integrable (that vanishes at infinity), or in other words that is integrable, in particular: The following integral form exist and is called the Fourier transform:

$$F(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f(t)e^{-j\omega t}dt \quad \text{and we have} \quad f(t) = \int_{-\infty}^{+\infty} F(\omega)e^{j\omega t}d\omega$$

  - Fourier transform of the derivative of f: $\mathcal{F}(f')(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} f'(t)e^{-j\omega t}dt = j\omega\mathcal{F}(f)(\omega)$
  - For a multi-variable function, it is possible to apply the Fourier transform to one variable only: for a function *c(x,t)*, we can apply the Fourier transform to the space variable x, leaving the time variable t unchanged:

$$\mathcal{F}_x(c(x,t)) = \hat{c}(\omega,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} c(x,t)e^{-j\omega x}dx$$

- For a function $f: \mathbb{R}^3 \to \mathbb{R}$, within an orthonormal basis:
  $\mathcal{F}(f)(\boldsymbol{\xi}) = \int_{\mathbb{R}^3} f(\boldsymbol{r})e^{-j\boldsymbol{r}.\boldsymbol{\xi}}d^3r$ Where $\boldsymbol{r}.\boldsymbol{\xi}$ is the dot product between two vectors.

- The Dirac delta function loosely defined is actually an example of a distribution, it is defined by its integration properties.

$$\delta(x) = \begin{cases} 0, & x \neq 0 \\ \infty, & x = 0 \end{cases} \qquad \int_{-\infty}^{\infty} \delta(x)\, dx = 1$$

- The Fourier transform of the delta function is: $\mathcal{F}(\delta)(\omega) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} \delta(t)e^{-j\omega t}dt = \frac{1}{2\pi}e^{-j\omega\times 0} = \frac{1}{2\pi}, \forall\omega$
- From the inverse theorem: $\delta(t) = \int_{-\infty}^{+\infty} \frac{1}{2\pi}e^{+j\omega t}d\omega = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{+j\omega t}d\omega = \mathcal{F}(f = 1)$

- For the function $f(t) = e^{j\omega_0 t}$:

$$\mathcal{F}(e^{j\omega_0 t}) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} e^{-j(\omega-\omega_0)t}dt = \frac{1}{2\pi}2\pi\delta\,(\omega-\omega_0) = \delta(\omega-\omega_0)$$

- Fourier transform of a Gaussian f(t):

$$f(t) = \frac{1}{\sqrt{2\pi}}\cdot e^{-\frac{a}{2}t^2} \qquad\qquad F(\omega) = \frac{1}{\sqrt{a}}\cdot e^{-\frac{\omega^2}{2a}}$$

Laplace Transforms

- The Laplace transform is defined in two ways:
  - Unilateral: $\mathcal{L}[f(t)] = \int_0^{\infty} e^{-st}f(t)dt = \overline{f(s)}$ (causal system that exists for t>0).
  - Bilateral: $\mathcal{L}[f(t)] = \int_{-\infty}^{\infty} e^{-st}f(t)dt = \overline{f(s)}$
- The Laplace transform is said to be defined in the s-domain, where s is a complex number.
- Properties of the Fourier (X(jω)) and Laplace (X(s)) transforms of a function x(t):

		Bilateral Laplace	Fourier
Property	$x(t)$	$X(s)$	$X(j\omega)$
Linearity	$ax_1(t) + bx_2(t)$	$aX_1(s) + bX_2(s)$	$aX_1(j\omega) + bX_2(j\omega)$
Time shift	$x(t-t_0)$	$e^{-st_0}X(s)$	$e^{-j\omega t_0}X(j\omega)$
Time scale	$x(at)$	$\frac{1}{\lvert a\rvert}X\left(\frac{s}{a}\right)$	$\frac{1}{\lvert a\rvert}X\left(\frac{j\omega}{a}\right)$
Differentiation	$\frac{dx(t)}{dt}$	$sX(s)$	$j\omega X(j\omega)$
Multiply by t	$tx(t)$	$-\frac{d}{ds}X(s)$	$-\frac{1}{j}\frac{d}{d\omega}X(j\omega)$
Convolution	$x_1(t) * x_2(t)$	$X_1(s) \times X_2(s)$	$X_1(j\omega) \times X_2(j\omega)$

- Be careful however, for the unilateral Laplace transform, a limit at the 0 boundary must be considered:

$$\mathcal{L}[f'(t)] = \int_0^{\infty} e^{-st}f'(t)dt = [e^{-st}f(t)]_0^{\infty} - (-s)\int_0^{\infty} e^{-st}f(t)dt$$
$$\mathcal{L}[f'(t)] = s\mathcal{L}[f(t)] - f(0)$$

- Convolution: operation on two functions that reflect how the shape of one is modified by the other:

$$f * g(t) = \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau = \int_{-\infty}^{+\infty} f(t-\tau)g(\tau)d\tau$$

- Laplace transform of common functions :

$f(t) = \mathcal{L}^{-1}\{F(s)\}$	$F(s) = \mathcal{L}\{f(t)\}$
1	$\frac{1}{s}$
e^{at}	$\frac{1}{s-a}$
$t^n, \quad n = 1, 2, 3, \ldots$	$\frac{n!}{s^{n+1}}$
$t^p, p > -1$	$\frac{\Gamma(p+1)}{s^{p+1}}$
$\sqrt{t}$	$\frac{\sqrt{\pi}}{2s^{\frac{3}{2}}}$
$t^{n-\frac{1}{2}}, \quad n = 1, 2, 3, \ldots$	$\frac{1 \cdot 3 \cdot 5 \cdots (2n-1)\sqrt{\pi}}{2^n s^{n+\frac{1}{2}}}$
$\sin(at)$	$\frac{a}{s^2+a^2}$
$\cos(at)$	$\frac{s}{s^2+a^2}$
$t\sin(at)$	$\frac{2as}{(s^2+a^2)^2}$
$t\cos(at)$	$\frac{s^2-a^2}{(s^2+a^2)^2}$

$\cos(at+b)$	$\frac{s\cos(b) - a\sin(b)}{s^2+a^2}$
$\sinh(at)$	$\frac{a}{s^2-a^2}$
$\cosh(at)$	$\frac{s}{s^2-a^2}$
$e^{at}\sin(bt)$	$\frac{b}{(s-a)^2+b^2}$
$e^{at}\cos(bt)$	$\frac{s-a}{(s-a)^2+b^2}$
$e^{at}\sinh(bt)$	$\frac{b}{(s-a)^2-b^2}$
$e^{at}\cosh(bt)$	$\frac{s-a}{(s-a)^2-b^2}$
$t^n e^{at}, \quad n = 1, 2, 3, \ldots$	$\frac{n!}{(s-a)^{n+1}}$
$f(ct)$	$\frac{1}{c}F\left(\frac{s}{c}\right)$
$u_c(t) = u(t-c)$ Heaviside Function	$\frac{e^{-cs}}{s}$
$\delta(t-c)$ Dirac Delta Function	e^{-cs}

Week 10&11: Differential Equations

- The general form of a second order linear ODE is: $a_2y'' + a_1y' + a_0y = f(x)$
- The solution is the sum of a particular solution of the non-homogeneous function, and general solution of the homogeneous equation.
- In the general case for the homogeneous function:

Systematic procedure for the solution of the homogeneous second-order DE	Example
Let the equation be $a_2y'' + a_1y' + a_0y = 0$	$y'' + 3y' + 2y = 0$
Let $y = e^{rx}$ be a solution of the DE. Substituting for $y' = re^{rx}$	$y = e^{rx}, \quad y' = \frac{dy}{dx} = re^{rx}$
and $y'' = r^2e^{rx}$	$y'' = \frac{d^2y}{dx^2} = r^2e^{rx}$
gives $a_2r^2e^{rx} + a_1re^{rx} + a_0e^{rx} = 0$	
We can factorise e^{rx}: $e^{rx}(a_2r^2 + a_1r + a_0) = 0$	$e^{rx}(r^2 + 3r + 2) = 0$
Since $e^{rx} \neq 0$, the expression in the bracket must be zero: $a_2r^2 + a_1r + a_0 = 0$	$r^2 + 3r + 2 = 0$
This is a quadratic in r. It is called the *auxiliary equation* of the DE. Its roots are $r_{1,2} = \frac{-a_1 \pm \sqrt{a_1{}^2 - 4a_2a_0}}{2a_2}$	$r_1 = -1, \quad r_2 = -2$
Provided that r_1 and r_2 are different, the general solution of the DE is $y = C_1e^{r_1x} + C_2e^{r_2x}$	$y = C_1e^{-x} + C_2e^{-2x}$

- The general form of a first order ODE is as follow: $p(x)y' + q(x)y = f(x)$
- The general solution of the non-homogeneous function above is: $y((x) = \frac{1}{I(x)}\int \frac{I(x)}{p(x)} f(x)dx$ where $I(x) = e^{\int \frac{q(x)}{p(x)}dx}$

Partial differential Equations

- The general form of a PDE is : $F(x_1,\ldots,x_n,\partial_{x_1}u,\ldots,\partial_{x_n}u,\partial^2_{x_1}u,\partial^2_{x_1x_2}u,\ldots,\partial^n_{x_1\ldots x_n}u)=0$
  - A linear equation is one in which the equation and any boundary or initial conditions do not include any product of the dependent variables or their derivatives;
  - Semi-linear equation: coefficients of the highest derivative are functions of the independent variables only.
  - Quasilinear equation: the coefficients of the higher order terms are functions of merely lower-order derivatives of the dependent variables.
  - The superposition principle does not apply to non-linear equations.

Week 12-14 : Probability and Statistics

- The set of functions square integrable over an interval (i.e. $\int |f(x)|^2$ exists and is finite), often called $\mathcal{L}^2$, is an Hilbert space, with the inner product:
$$\langle f,g\rangle = \int f^*(x)g(x)dx \text{ and } \|f\|^2 = \langle f,f\rangle = \int |f(x)|^2 dx$$
- The common operators discussed in wave mechanics are self-adjoint (position, momentum, most Hamiltonians).

Probability :

- Consider performing N times an experiment that returns events in a set Ω .We can consider an event α, which happens N_α times. The empirical frequency is given by $f_\alpha(N) = \frac{N_\alpha}{N}$
When N becomes large and the experiments are done independently from each other (they don't influence each other), $f_\alpha(N)$ converges to a well defined and finite limit called the event probability:
$$P(\alpha) = \lim_{N\to\infty} f_\alpha(N) \geq 0$$
- From the definition, it is straightforward that (it is what defines a probability function):
  - $P(\Omega) = 1$ and $P(\emptyset) = 0$.
  - If (A_i) is a family of events that are not overlapping (i.e. $\forall i \neq j, A_i \cap A_j = \emptyset$): $P(\bigcup_i A_i) = \sum_i P(A_i)$
- **Normalization:** $\sum p_i = 1$
- For two independent events A and B, $P(A \cap B) = P(A) \times P(B)$
- Random variables: If one considers a game of drawing balls, with the probabilities p_i, to which is associated not just an event (a ball with i on it), but a number, such as an amount of money won x_i. x_i is a variable associated with a random event, it is a called a random variable.
  - The set $\{x_i, p_i\}$ defines the law for the random variable x_i. In this case, it is a discrete law.
  - Continuous random variables: a continuous random variable x can take values in an interval $[a,b]$, the probability density $p(x) \geq 0$ defines the law of the random variable: the probability to find a value between x & $x+dx$ is $p(x)dx$. Normalization: $\int_a^b p(x)dx = 1$

- **Conditional probability**:
For two events A and B, the probability that the event B occurs, knowing event A, is given by: $P(B/A) = \frac{P(B\cap A)}{P(A)}$
If A and B are independent, knowing A should not affect the probability of B, and we see that $P(B/A) = P(B)$.
- **Average:**
We consider a random variable x, with outcome values (x_α, p_α) if discrete, and if continuous, $x \in [a,b]$ with a probability density $p(x)$. A function $f(x)$ is also a random variable. We can define its average:
  - $\sum_\alpha f(x_\alpha)\, p_\alpha$ if discrete;
  - $\int_a^b f(x)p(x)dx$ if continuous.
- In particular, the average for the random variable x is given by: $\langle x\rangle = \int_a^b xp(x)dx$

- **Standard deviation:**
The standard deviation of the random variable is defined as: $(\Delta x)^2 = \sigma^2 = \langle (x - \langle x\rangle)^2\rangle$
  - Δx (σ) is called standard deviation, $(\Delta x)^2$ (σ^2) is called variance.
  - We also have: $\sigma^2 = \langle x^2\rangle - \langle x\rangle^2$

- **The binomial distribution** is a discrete probability distribution of the number of successes in a sequence of n independent experiments, each with a probability of success p or failure 1 − p.
  - For a single trial, the binomial distribution is a Bernoulli distribution
  - For n trials, the probability of having a certain sequence of outcomes with k successes is: $p^k(1-p)^{n-k}$
  - If we associate a random variable for each trial X_k that is 1 for success, and 0 for failure, we can also define the random variable X that is the sum of the variable of each trial: $X = X_1 + \ldots + X_n$,
  - The probability to have k successes is P(X = k), which is also the number of configurations that returns X = k. This is called the Binomial distribution of parameters n and p: $P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

Statistical Physics :

- Consider a system of Hamiltonian $\hat{H}$ with eigenvectors $|\psi_n\rangle$. An observable $\hat{O}$ with eigenvalues o_α and eigenvectors $|\alpha\rangle$, will have outcomes o_α of probability $|\langle\alpha|\psi_n\rangle|^2$. The average measure will be: $\langle\psi_n|\hat{O}|\psi_n\rangle = \sum_\alpha o_\alpha |\langle\alpha|\psi_n\rangle|^2$
- The statistical description of the macroscopic system attributes a certain probability p_n of having the system in the state $|\psi_n\rangle$.We hence have a two-level statistics: $\langle\hat{O}\rangle = \sum_n p_n \langle\psi_n|\hat{O}|\psi_n\rangle$
  One level is linked to the quantum physics that governs nature at the microscopic level. The second one arises from an impossibility to know exactly the Hamiltonian and other parameters of a macroscopic system of 10^{23} particles: a true statistical approach.

Fundamental postulate of statistical physics:

For a closed system of fixed energy E (microcanonical ensemble), the microstates are equiprobable.

- If $\mathcal{W}(E)$ is the number of eigenstates of an Hamiltonian $\hat{H}$ for which $E \leq E_n \leq E + \delta E$, we can assign the probability for each eigenstate to be: $p_n = \frac{1}{\mathcal{W}(E)}$
- One can construct a new operator called density operator $\hat{D}$, as: $\hat{D} = \sum_n p_n |\psi_n\rangle\langle\psi_n|$
- $\hat{D}$ is Hermitian, positive and $\mathrm{Tr}(\hat{D}) = 1$ since the probabilities are normalized. We have: $\langle\hat{O}\rangle = \mathrm{Tr}(\hat{D}\hat{O})$
- p_n can be described as the temporal average of $|\langle\psi_n|\psi(t)\rangle|^2$, i.e. the measure of the overlap of the real wave function of the system, and the eigenstate $|\psi_n\rangle$.

- The concept of **entropy** in thermodynamics is linked to the number of accessible microstates : $S = k ln(\mathcal{W}(E))$
  Where k is the Boltzmann constant $k = 1.380658 \times 10^{-23}\, J.K^{-1}$
- If we consider two isolated systems of energies E_1 and E_2 brought in contact. The total energy $E = E_1 + E_2$ is conserved but E_1 and E_2can vary.
  - The probability to have the system 1 at energy E_1 is: $p(E_1) = \frac{\mathcal{W}_1(E_1)\mathcal{W}_2(E-E_1)}{\sum_{E_1'} \mathcal{W}_1(E_1')\mathcal{W}_2(E-E_1')}$
  - This probability is peaked for $\left.\frac{\partial S_1(E_1)}{\partial E_1}\right)_{E_1=E_1^{eq}} = \left.\frac{\partial S_2(E-E_1)}{\partial E_1}\right)_{E_2=E-E_1^{eq}}$, where E_1^{eq} is the energy of system 1 at equilibrium. We can then define: $\frac{1}{T} \equiv \frac{\partial S}{\partial E}$

- For a grand canonical ensemble, the probability of finding the small system 1 *in one micro-state* of energy E_1 is $p = \frac{\mathcal{W}_2(E-E_1)}{\sum_{E_1'} \mathcal{W}_1(E_1')\mathcal{W}_2(E-E_1')}$, which can be rewritten as: $p = \frac{1}{Z} e^{-\frac{E_1}{kT_2}}$ with $Z = \sum_n e^{-\frac{E_1^{(n)}}{kT_2}}$ the partition function.
- A grand canonical ensemble will allow not only an exchange of energy for a small system with a thermostat, but also an exchange of matter (atoms, molecules, particles etc…)
- The probability to have a microstate of energy E_1 with N_1 particles is: $p = \frac{\mathcal{W}_2(E-E_1, N-N_1)}{\sum_{E_1',N_1'} \mathcal{W}_1(E_1',N_1')\mathcal{W}_2(E-E_1',N-N_1')}$
- With the definition of the differential of multiple variables, the entropy being an exact differential, we can do a very similar development as before and obtain a relationship for the temperature and for the chemical potential: $\mu = -T\frac{\partial S}{\partial N}$
- We then obtain: $p_n = \frac{1}{Z_G} e^{-\beta E_n + \alpha N_n}$ with $\alpha = \beta\mu$ and $Z_G = \sum_n e^{-\beta E_n + \alpha N_n} = \sum_{N=0}^{\infty} e^{\alpha N} Z_N(\beta)$