Summaries of mathematics and physics

Version: May 8, 2021

Contents

Part I

Mathematics

Chapter 1

First year

1.1 Fonaments de les matemàtiques

1.1.1 | Conjunts i aplicacions

Definition 1.1. Un conjunt és una col · lecció d'objectes units per una propietat comuna.

Definition 1.2. Sigui E un conjunt. Diem que un conjunt F és un subconjunt de E ($F \subseteq E$) si, i només si, tot element de F és element de E.

Definition 1.3 (Axioma d'extensionalitat). Siguin E, F dos conjunts. Diem que E i F són iguals si, i només si, $E \subseteq F$ i $F \subseteq E$.

Definition 1.4. Sigui E un conjunt. El subconjunt $\mathcal{P}(E)$ és el conjunt dels subconjunts de E (parts de E).

Definition 1.5. Siguin A, B dos conjunts. La intersecció de A i B ($A \cap B$) és el conjunt format pels elements comuns a A i B.

Definition 1.6. Siguin A, B dos conjunts. La unió de A i B ($A \cup B$) és el conjunt format pels elements de A i de B.

Definition 1.7. Sigui E un conjunt i $A \subseteq B$. El complementari de A en E és el conjunt $A^c = \{x \in E \mid x \notin A\}$.

Proposition 1.8 (Lleis de Morgan). Sigui E un conjunt i siguin A, B dos subconjunts:

1. $(A \cup B)^c = A^c \cap B^c$
2. $(A \cap B)^c = A^c \cup B^c$

Definition 1.9. Siguin A, B dos conjunts. El producte cartesià $A \times B$ és el conjunt $A \times B = \{(a, b) \mid a \in A \text{ i } b \in B\}$.

Definition 1.10. Siguin E, F dos conjunts. Una aplicació de E en F és una regla que a cada element de E se li associa un únic element de F.

Definition 1.11. Sigui $f : E \to F$ una aplicació i $A \subseteq E$ un subconjunt. La imatge de A és el subconjunt de F definit per $f(A) = \{f(a) \mid a \in A\}$. Si $A = E$, $f(E) = \mathrm{Im}(f)$ és la imatge de f.

Definition 1.12. Sigui $f : E \to F$ una aplicació i $b \in F$. Un antecedent de b és un element $a \in E$ tal que $f(a) = b$.

Definition 1.13. Sigui $f : E \to F$ una aplicació i $B \subseteq F$ un subconjunt. La preimatge de B és el subconjunt de E definit per $f^{-1}(B) = \{a \in E \mid f(a) = b, b \in B\}$.

Definition 1.14. Sigui $f : E \to F$ una aplicació. Les següents tres assercions són equivalents:

1. $\forall b \in F$, $f^{-1}(b)$ té com a màxim un element.
2. $\forall \alpha, \beta \in E$, si $\alpha \neq \beta$, llavors $f(\alpha) \neq f(\beta)$.
3. $\forall \alpha, \beta \in E$, si $f(\alpha) = f(\beta)$, llavors $\alpha = \beta$.

Si f compleix una d'aquestes tres assercions, les compleix totes i diem que f és injectiva.

Proposition 1.15. Siguin $f : E \to F, g : F \to G$:

1. Si f i g són injectives, llavors $g \circ f$ també ho és.
2. Si $g \circ f$ és injectiva, llavors f també ho és.

Definition 1.16. Sigui $f : E \to F$ una aplicació. Les següents tres assercions són equivalents:

1. Tot element de F té almenys un antecedent de E per f.
2. $\forall y \in F$, $\exists \alpha \in E$ tal que $f(\alpha) = y$.
3. $\mathrm{Im}(f) = F$.

Si f compleix una d'aquestes tres assercions, les compleix totes i diem que f és exhaustiva.

Proposition 1.17. Siguin $f : E \to F, g : F \to G$:

1. Si f i g són exhaustives, llavors $g \circ f$ també ho és.
2. Si $g \circ f$ és exhaustiva, llavors g també ho és.

Definition 1.18. Sigui $f : E \to F$ una aplicació. Diem que f és bijectiva si és injectiva i exhaustiva. Les aplicacions bijectives tenen una aplicació inversa associada $f^{-1} : F \to E$.

1.1.2 | Grup simètric

Definition 1.19. Sigui $n \in \mathbb{N}$. Denotem per S_n el conjunt de les bijeccions de $\{1, 2, \ldots, n\}$ en ell mateix. Un element de S_n és una permutació de $\{1, \ldots, n\}$.

Theorem 1.20. El cardinal de S_n és $n!$.

Definition 1.21. Sigui $f \in S_n$. El conjunt $E_f = \{m \in \mathbb{N}^* \mid f^m = \mathrm{id}\}$ té un element minimal $o(f)$. L'enter $o(f)$ és l'ordre de f.

Definition 1.22. Sigui $f \in S_n$. El suport de f és $\mathrm{sup}(f) = \{k \in \{1, \ldots, n\} \mid f(k) \neq k\}$.

Lemma 1.23. Sigui $f \in S_n$, llavors:

1. $p \in \mathrm{sup}(f) \implies f(p) \in \mathrm{sup}(f)$.
2. $\mathrm{sup}(f) = \mathrm{sup}(f^{-1})$.

Lemma 1.24. Siguin $f, g \in S_n$. Si $\mathrm{sup}(f) \cap \mathrm{sup}(g) = \emptyset$, llavors $f \circ g = g \circ f$.

Definition 1.25. Sigui $f \in S_n$ i $k \in \{1, \ldots, n\}$. L'òrbita de k és el conjunt finit $\{k, f(k), f^2(k), \ldots\}$.

Theorem 1.26 (Estructura de l'òrbita). Siguin $f \in S_n$ i $\Omega = \{\omega_1, \ldots, \omega_k\}$ el conjunt de les òrbites de f. Llavors:

1. $\bigcup_{j=1}^{k} \omega_j = \{1, \ldots, n\}$.
2. Si $\omega_i, \omega_j \in \Omega$ i $\omega_i \cap \omega_j \neq \emptyset$, llavors $\omega_i = \omega_j$.
3. Cap òrbita és buida.

Theorem 1.27 (Estructura lineal de les òrbites). Sigui $f \in S_n$ i ω una de les seves òrbites. Sigui $a \in \{1, \ldots, n\}$ un element de ω i sigui k el cardinal de ω. Llavors $\omega = \{a, f(a), \ldots, f^{k-1}(a)\}$ i $f^k(a) = a$.

Definition 1.28. Si $f \in S_n$ té una única òrbita no reduïda a un element, llavors diem que f és un cicle de longitud el cardinal del cicle.

Theorem 1.29. Sigui $f \in S_n$, llavors f s'escriu de manera única, llevat de l'ordre, com a producte de cicles amb suports dos a dos disjunts.

Corollary 1.30. Sigui $f \in S_n$ i $f = \sigma_1 \cdots \sigma_l$ la seva descomposició en producte de cicles disjunts. Aleshores $o(f) = \text{mcm}(\sigma_1, \ldots, \sigma_l)$.

Corollary 1.31. Sigui $f \in S_n$, llavors f és producte de transposicions.

Definition 1.32. Sigui $\sigma \in S_n$. El signe de σ és $\varepsilon(\sigma) = (-1)^{n-r}$ on r és el nombre d'òrbites de σ (incloent les trivials).

Theorem 1.33. Sigui $f \in S_n$ arbitrària i $\tau \in S_n$ una transposició. Llavors $\varepsilon(f\tau) = \varepsilon(f)\varepsilon(\tau) = -\varepsilon(f)$.

Corollary 1.34. Sigui $f \in S_n$ i $f = \tau_1 \cdots \tau_l$ una escriptura de f com a producte de transposicions. Llavors $\varepsilon(f) = (-1)^l$.

Corollary 1.35. A les escriptures de f com a producte de transposicions, la paritat del nombre de transposicions no varia.

1.1.3 | Relacions d'equivalència i d'ordre

Definition 1.36. Sigui E un conjunt i $\sim_R$ una relació sobre E. Direm que $\sim_R$ és una relació d'equivalència si, i només si, es compleixen els següents axiomes:

1. La relació és reflexiva:
$$\forall a \in E,\ a \sim_R a.$$
2. La relació és simètrica:
$$\forall a, b \in E \text{ si } a \sim_R b,\ \text{llavors } b \sim_R a.$$
3. La relació és transitiva:
$$\forall a, b, c \in E \text{ si } a \sim_R b \text{ i } b \sim_R c,\ \text{llavors } a \sim_R c.$$

Definition 1.37. Sigui $(E, \sim_R)$ una relació d'equivalència i sigui $a \in E$. La classe d'equivalència de a és el subconjunt de E: $[a] = \{b \in E \mid a \sim_R b\}$.

Theorem 1.38. Sigui $(E, \sim_R)$ una relació d'equivalència. Les classes d'equivalència de E formen una partició de E. És a dir, siguin $\{\omega_i\}$ les classes d'equivalència, llavors:

1. $\bigcup_{i \in I} \omega_i = E$.
2. Si $i, j \in I$ i $\omega_i \cap \omega_j \neq \emptyset$, llavors $\omega_i = \omega_j$.
3. Si $i \in I \implies \omega_i \neq \emptyset$.

Definition 1.39. El conjunt de les classes d'equivalència de $(E, \sim_R)$ es denota $E/\sim_R$ i es diu conjunt quocient.

Definition 1.40. Sigui E un conjunt i $\leq$ una relació sobre E. Direm que $\leq$ és una relació d'ordre si, i només si, es compleixen els següents axiomes:

1. La relació és reflexiva:
$$\forall a \in E,\ a \leq a.$$
2. La relació és antisimètrica:
$$\forall a, b \in E \text{ si } a \leq b \text{ i } b \leq a,\ \text{llavors } a = b.$$
3. La relació és transitiva:
$$\forall a, b, c \in E \text{ si } a \leq b \text{ i } b \leq c,\ \text{llavors } a \leq c.$$

Definition 1.41. Sigui $(E, \leq)$ un conjunt ordenat. Direm que $a \in E$ és minimal (respectivament maximal) si, i només si, $\forall b \in E$ si $b \leq a$ (respectivament $b \geq a$), llavors $b = a$. Direm que $a \in E$ és un mínim (respectivament màxim) si, i només si, $\forall b \in E$, $a \leq b$ (respectivament $a \geq b$).

Lemma 1.42. Sigui $(E, \leq)$ un conjunt ordenat. Si $(E, \leq)$ admet un mínim, llavors aquest és únic.

Definition 1.43. Un conjunt està totalment ordenat si tot parell d'elements és comparable. Un conjunt ordenat està ben ordenat si tot subconjunt admet un element minimal.

Theorem 1.44. Tot conjunt pot ser ben ordenat.

1.1.4 | Cardinalitat i combinatòria

Definition 1.45. Siguin E, F dos conjunts. Direm que E i F tenen el mateix cardinal si, i només si, existeix una bijecció de $E \to F$.

Definition 1.46. Siguin E, F dos conjunts. Direm que $|E| \leq |F|$ si, i només si, existeix una aplicació injectiva $E \to F$.

Theorem 1.47 (Teorema de Cantor-Bernstein). Siguin E, F dos conjunts. Si existeix una injecció $E \to F$ i una injecció $F \to E$, llavors existeix una bijecció $E \to F$. La comparació de cardinals és una relació d'ordre.

Proposition 1.48. Càlcul de cardinals:
Siguin $A, B \subseteq E$ dos subconjunts finits.

1. Principi d'inclusió−exclusió: $|A \cup B| = |A| + |B| - |A \cap B|$
2. Producte cartesià: $|A \times B| = |A||B|$
3. $|A^c| + |A| = |E|$
4. $|\mathcal{P}(E)| = 2^{|E|}$

Theorem 1.49 (Teorema de Cantor). Sigui E un conjunt, llavors $|\mathcal{P}(E)| > |E|$.

Proposition 1.50. Siguin E, F dos conjunts finits. El conjunt de les aplicacions $E \to F$ té cardinal $|F|^{|E|}$.

Definition 1.51. Sigui E un conjunt i $A \in \mathcal{P}(E)$. Definim la funció característica de A com:

$$\begin{aligned} \chi_A : E &\to \{0,1\} \\ r &\mapsto \begin{cases} 1 & \text{si} \quad r \in A \\ 0 & \text{si} \quad r \notin A \end{cases} \end{aligned}$$

Proposition 1.52. Propietats del coeficient binomial:

1. $\binom{n}{k} = \frac{n!}{(n-k)!k!}$
2. $\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}$
3. $\sum_{k=0}^{n} \binom{n}{k} = 2^n$
4. $k\binom{n}{k} = n\binom{n-1}{k-1}$

Proposition 1.53. Sigui $f : E \to F$ una aplicació entre conjunts del mateix cardinal finit. Les següents assercions són equivalents:

1. f és injectiva.
2. f és exhaustiva.
3. f és bijectiva.

Corollary 1.54. Sigui $f : E \to F$ una aplicació entre conjunts finits. Aleshores:

1. f és injectiva $\implies |E| \leq |F|$.
2. f és exhaustiva $\implies |E| \geq |F|$.

Theorem 1.55 (Principi del colomar). Siguin E, F dos conjunts amb n i k elements, respectivament, i $f : E \to F$ una aplicació. Si $n > k$, llavors existeixen elements de E amb $f(a) = f(b)$ i $a \neq b$.

Proposition 1.56 (Variacions sense repetició). El nombre de variacions sense repetició de conjunts amb m elements agafats amb tuples de n elements sense repetir-los és $\frac{n!}{(n-k)!}$.

Proposition 1.57 (Variacions amb repetició). El nombre de variacions amb repetició de conjunts amb n elements agafats amb tuples de k, els quals poden ser repetits, és n^k.

Proposition 1.58 (Combinacions sense repetició). El coeficient binomial $\binom{n}{k}$ és el nombre de subconjunts de k elements entre un conjunt amb n elements.

Proposition 1.59 (Combinacions amb repetició). El coeficient binomial $\binom{n+k-1}{k}$ és el nombre de combinacions amb repetició de k elements escollits entre un conjunt amb n elements.

1.1.5 | Nombres enters i congruències

Definition 1.60. Siguin $m, n \in \mathbb{Z}$. Diem que m és múltiple de n si existeix $k \in \mathbb{Z}$ tal que $m = kn$.

Theorem 1.61. Siguin $D, d \in \mathbb{Z}$, $d \neq 0$. Llavors existeixen $q, r \in \mathbb{Z}$ únics tals que $D = qd + r$ i $0 \leq r \leq |d|$.

Definition 1.62. En un anell $(A, +, \cdot)$, un subconjunt $I \subseteq A$ és un ideal si, i només si:

1. $\forall a, b \in I$, $a + b \in I$.
2. $\forall a \in I$ i $\forall n \in A$, $na \in I$.

Lemma 1.63. $\forall n \in \mathbb{Z}$, $I = n\mathbb{Z} = \{na \mid a \in \mathbb{Z}\}$ és un ideal de $\mathbb{Z}$. Diem que n és un generador de I.

Lemma 1.64. Siguin I, J dos ideals de A. Llavors el conjunt $I \cap J$ és un ideal.

Lemma 1.65. Siguin I, J dos ideals. L'ideal generat per I i J és l'ideal $I + J = \{a + b \mid a \in I, b \in J\}$. A més, aquest ideal és el més petit que conté I i J.

Proposition 1.66. Siguin $a, b \in \mathbb{Z}$. $a\mathbb{Z} \subseteq b\mathbb{Z} \iff b \mid a$.

Corollary 1.67. Siguin $a, b \in \mathbb{Z}$. $a\mathbb{Z} = b\mathbb{Z} \iff a = \pm b$.

Proposition 1.68. Sigui A un anell i siguin $I = a\mathbb{Z}, J = b\mathbb{Z}$ dos ideals. Aleshores $\exists!\, m \in \mathbb{N}^*$ tal que $a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$. Aquest enter m és el mínim comú múltiple de a i b.

Proposition 1.69. Sigui A un anell i siguin $I = a\mathbb{Z}$ $J = b\mathbb{Z}$ dos ideals. Aleshores $\exists!\, d \in \mathbb{N}^*$ tal que $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$. Aquest enter d és el màxim comú divisor de a i b.

Definition 1.70. Siguin $a, b \in \mathbb{Z}$. Diem que a i b són coprimers si, i només si, $\text{mcd}(a, b) = 1$.

Definition 1.71. Diem que a és primer ($a \in \mathbb{P}$) si, i només si, $a\mathbb{Z}$ és maximal per a la inclusió d'ideals propis.

Theorem 1.72 (Teorema de Bézout). Sigui $a, b \in \mathbb{Z}$, llavors existeixen $u, v \in \mathbb{Z}$ tals que $au + bv = \text{mcd}(a, b)$. A més, $\text{mcd}(a, b) = 1 \iff \exists u, v \in \mathbb{Z}$ tals que $au + bv = 1$.

Theorem 1.73 (Teorema de Gauß). Sigui $a, b \in \mathbb{Z}$. Si $a \mid bc$ i $\text{mcd}(a, b) = 1$ llavors $a \mid c$.

Corollary 1.74. Sigui $a_1, a_2 \in \mathbb{Z}$ coprimers. Si $a_1 \mid b$ i $a_2 \mid b$, llavors $a_1 a_2 \mid b$.

Theorem 1.75 (Teorema de La Vallée Poussin-Hadamard). Sigui $x \in \mathbb{R}$, aleshores $\pi(x) \approx \frac{x}{\log(x)}$ on $\pi(x)$ és el nombre de nombres primers $\leq x$.

Theorem 1.76. Siguin $a, b \in \mathbb{Z}$. Llavors

$$\text{mcd}(a, b)\text{mcm}(a, b) = |ab|.$$

Lemma 1.77. Sigui $p \in \mathbb{P}$ i $a \in \mathbb{Z}$. Llavors $p \mid a$ o $\text{mcd}(a, p) = 1$.

Theorem 1.78 (Teorema fonamental de l'aritmètica). Sigui $n \in \mathbb{N}^*$, llavors n s'escriu com a producte de nombres primers únics, llevat de l'ordre.

Theorem 1.79 (Teorema d'Euclides). Sigui $\mathbb{P}$ el conjunt dels nombres primers positius. $\mathbb{P}$ és infinit.

Theorem 1.80. L'equació $ax + by = c$ admet almenys una solució si, i només si, $\text{mcd}(a, b) \mid c$.

Definition 1.81. $x \equiv y \mod n \iff x - y \in n\mathbb{Z}$.

Lemma 1.82. $\mathbb{Z}/n\mathbb{Z}$ té n elements.

Theorem 1.83. Com que $(\mathbb{Z}, +, \cdot)$ és un anell commutatiu, $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ és un anell commutatiu i, per construcció, la projecció canònica

$$f : \mathbb{Z} \to \mathbb{Z}/n\mathbb{Z}$$
$$a \mapsto [a]$$

és un morfisme d'anells.

Lemma 1.84. Sigui $n \in \mathbb{Z}$. Llavors $[a] \in \mathbb{Z}/n\mathbb{Z}$ és invertible per a la multiplicació si, i només si, $\text{mcd}(a, n) = 1$.

Corollary 1.85. Un anell $(A, +, \cdot)$ en el qual tot element no nul és invertible és un cos. $(\mathbb{Z}/n\mathbb{Z}, +, \cdot)$ és un cos si, i només si, $n \in \mathbb{P}$.

Theorem 1.86 (Teorema xinès del residu). Siguin m, n coprimers, aleshores l'aplicació:

$$\psi : \mathbb{Z}/mn\mathbb{Z} \to \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$$
$$\overline{a}^{mn} \mapsto (\overline{a}^m, \overline{a}^n)$$

és un isomorfisme d'anells.

Definition 1.87 (Funció indicatriu d'Euler). Sigui $n \in \mathbb{N}^*$. $\varphi(n) = |\{\alpha \in \mathbb{Z}/n\mathbb{Z} \mid \alpha \text{ és invertible}\}| = |\{0 \leq r \leq n \mid \text{mcd}(r, n) = 1\}|$.

Theorem 1.88 (Teorema d'Euler). Sigui $a \in \mathbb{Z}$ i $n \in \mathbb{N}$ tal que $\text{mcd}(a, n) = 1$, llavors $a^{\varphi(n)} \equiv 1 \mod n$. En particular, $a^{-1} \equiv a^{\varphi(n)-1} \mod n$.

Theorem 1.89 (Petit teorema de Fermat). Si p és primer, $\varphi(p) = p - 1$. Aleshores $a^p \equiv a \mod p$ i, en particular, si $\text{mcd}(a, p) = 1$, $a^{p-1} \equiv 1 \mod p$.

1.1.6 | Polinomis

Proposition 1.90. Sigui A un cos. Si $P, Q \in A[x]$ i $P, Q \neq 0$, llavors $PQ \neq 0$.

Theorem 1.91 (Teorema de divisió euclidiana). Sigui A un cos. Siguin $P, S \in A[x]$ amb $S \neq 0$. Llavors $\exists!\, Q, R \in A[x]$ tals que $P = QS + R$ i $0 \leq \deg(R) < \deg(S)$.

Theorem 1.92. Sigui A un cos. Llavors $A[x]$ és un anell principal, és a dir, si $I \subset A[x]$ és un ideal, llavors $\exists P \in A[x]$ tal que $I = PA[x]$.

Definition 1.93. Sigui $P, Q \in A[x]$. Aleshores $\text{mcd}(P, Q)$ és un generador de $PA[x] + QA[x]$ i $\text{mcm}(P, Q)$ és un generador de $PA[x] \cap QA[x]$.

Definition 1.94. Diem que un polinomi $P = a_0 + a_1 x + \cdots + a_n x^n$ és mònic si $a_n = 1$.

Theorem 1.95 (Teorema de Bézout). Siguin $P, Q \in K[x]$, llavors $\exists U, V \in K[x]$ tals que $PU + QV = \text{mcd}(P, Q)$.

Definition 1.96. Dos polinomis P, Q són coprimers si, i només si, $\text{mcd}(P, Q) = 1$.

Theorem 1.97 (Teorema de Gauß). Siguin $P, A, B \in K[x]$. Si $P \mid AB$ i $\text{mcd}(A, P) = 1$, llavors $P \mid B$.

Definition 1.98. Un polinomi $P \in K[x]$ és primer si, i només si, el seu ideal $PK[x]$ és maximal per a la inclusió d'ideals de $K[x]$, és a dir, $\forall I$ ideal si $PK[x] \subsetneq I$, llavors $K[x] = I$.

Theorem 1.99 (Teorema de Ruffini). Sigui K un cos i sigui $P \in K[x]$ i $\lambda \in K$. Llavors $x - \lambda \mid P \iff P(\lambda) = 0$.

Definition 1.100. Sigui $P \in K[x]$, llavors P és irreductible si, i només si, $PK[x]$ és maximal.

Theorem 1.101. Sigui $P \in K[x]$, llavors P té com a màxim $\deg(P)$ arrels.

Theorem 1.102 (Teorema d'Alembert). Tot polinomi no constant $P \in \mathbb{C}[x]$ té exactament $\deg(P)$ arrels.

Corollary 1.103. Sigui $P \in \mathbb{C}[x]$ i $\deg(P) > 1$. Llavors existeixen $\alpha, r_1, \ldots, r_n \in \mathbb{C}$ únics, llevat de l'ordre, tal que $P = \alpha(x - r_1) \cdots (x - r_n)$ on r_i són les arrels de P i α el coeficient dominant.

Corollary 1.104. Les arrels a $\mathbb{C} \setminus \mathbb{R}$ es presenten en parells $(r_s, \overline{r_s})$.

Theorem 1.105. A $\mathbb{R}[x]$ els polinomis irreductibles són de grau 1 (arrels reals) o de grau 2 (arrels complexes no reals).

1.2 Linear algebra

1.2.1 | Matrices

Definition 2.1. We say a matrix $A \in \mathcal{M}_n(\mathbb{R})$ is *invertible* if there is a matrix $B \in \mathcal{M}_n(\mathbb{R})$ satisfying $AB = BA = I_n$.

Lemma 2.2. El producte de matrius invertibles és invertible.

Theorem 2.3 (Teorema de Gauß). Donada una matriu qualsevol $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, existeix una matriu invertible $P \in \mathcal{M}_m(\mathbb{R})$ tal que $A' = PA$ és esglaonada i reduïda. A més, A' està únicament determinada per A.

Theorem 2.4 (Teorema de la PAQ reducció). Donada una matriu $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, existeixen matrius invertibles $P \in \mathcal{M}_m(\mathbb{R})$ i $Q \in \mathcal{M}_n(\mathbb{R})$ tals que

$$PAQ = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right).$$

El nombre r és el rang de les matrius PAQ i A.

Proposition 2.5. $\forall A \in \mathcal{M}_{m\times n}(\mathbb{R})$ tenim que rang $A =$ rang A^t.

Theorem 2.6 (Teorema de Rouché). Donat un sistema d'equacions lineals $Ax = b$ amb n incògnites, el sistema és:

- Compatible determinat $\iff$ rang $A =$ rang$(A \mid b) = n$.
- Compatible indeterminat amb s variables lliures $\iff$ rang $A =$ rang$(A \mid b) = n - s$.
- Incompatible $\iff$ rang $A \neq$ rang$(A \mid b)$.

Definition 2.7. Un determinant és una aplicació $\mathcal{M}_n(\mathbb{R}) \to \mathbb{R}$ que compleixi:

1. L'aplicació ha de ser lineal en cada fila i columna. És a dir, si $a_1, \ldots, a_n$ són les columnes d'una matriu $A \in \mathcal{M}_n(\mathbb{R})$ i $a_j = \lambda u + \mu v$, aleshores:

$$\begin{aligned}\det A &= \det([a_1|\cdots|a_j|\cdots|a_n]) = \\ &= \det([a_1|\cdots|\lambda u + \mu v|\cdots|a_n]) = \\ &= \lambda \det([a_1|\cdots|u|\cdots|a_n]) + \\ &\quad + \mu \det([a_1|\cdots|v|\cdots|a_n])\end{aligned}$$

2. El determinant canvia de signe si s'intercanvien dues columnes.
3. $\det I_n = 1$.

Proposition 2.8. Donada una matriu $A \in \mathcal{M}_n(\mathbb{R})$, A no és invertible $\iff$ rang $A < n \iff \det A = 0$.

Theorem 2.9. $\forall A, B \in \mathcal{M}_n(\mathbb{R})$ tenim que $\det(AB) = \det A \det B$.

Proposition 2.10. Denotem S_n el grup de permutacions de $\{1, \ldots, n\}$ i sigui $A \in \mathcal{M}_n(\mathbb{R})$. Llavors:

$$\det A = \sum_{\sigma \in S_n} \varepsilon(\sigma) \prod_{i=1}^{n} a_{i\sigma(i)}$$

Proposition 2.11. $\forall A \in \mathcal{M}_n(\mathbb{R})$ tenim que $\det A = \det A^t$.

Theorem 2.12. $\forall A \in \mathcal{M}_n(\mathbb{R})$ es compleix $A(\text{adj } A)^t = (\det A) I_n$ i si $\det A \neq 0$ aleshores,

$$A^{-1} = \frac{1}{\det A}(\text{adj } A)^t$$

1.2.2 | Espais vectorials

Definition 2.13. Sigui E un K-espai vectorial i F un subconjunt de E, llavors $(F, +_E, \cdot_E)$ és un K-espai vectorial si es verifica $\lambda v_1 + \mu v_2 \in F$ $\forall v_1, v_2 \in F$ i $\forall \lambda, \mu \in K$.

Lemma 2.14. La intersecció de subespais vectorials és un subespai vectorial.

Definition 2.15. Donat un K-espai vectorial E, una base de E és un conjunt ordenat B de vectors de E que és:

1. Sistema de generadors de E.
2. Linealment independent.

Theorem 2.16 (Teorema de Steinitz). Donat un K-espai vectorial E, B una base de E i $(v_1, \ldots, v_k)$ vectors linealment independents de E, aleshores podem substituir k vectors apropiats de B per $(v_1, \ldots, v_k)$ i definir una nova base.

Definition 2.17. La suma de dos subespais F, G dins d'un K-espai vectorial E és: $F + G = \langle F \cup G \rangle = \{u + v \mid u \in F, v \in G\}$.

Proposition 2.18 (Fórmula de Graßmann). $\dim(F + G) + \dim(F \cap G) = \dim F + \dim G$.

Definition 2.19. Sigui E un K-espai vectorial i siguin $F, G \subset E$ dos subespais vectorials. Llavors la suma $F + G$ és directe $(F \oplus G) \iff F \cap G = \{0\}$.

Definition 2.20. Donada una matriu $A \in \mathcal{M}_n(\mathbb{R})$, un menor d'ordre r de A és una submatriu $A' \in \mathcal{M}_r(\mathbb{R})$ obtinguda seleccionant r files i r columnes de A.

Definition 2.21. Sigui E un K-espai vectorial i $F \subset E$ un subespai vectorial. Anomenarem subespai complementari de F a tot subespai $G \subset E$ que compleixi $F \oplus G = E$.

Definition 2.22. Direm que dos vectors $u, v \in E$ són equivalents mòdul F (on $F \subset E$ un subespai vectorial) si $u - v \in F$ i escriurem $u \sim_F v$. $\sim_F$ és una relació d'equivalència.

Definition 2.23. L'espai quocient E/F és el conjunt de les classes d'equivalència $u + F = [u]$ amb les operacions:

$$[u] + [v] = [u + v] \qquad a[u] = [au]$$

E/F és un K-espai vectorial.

Proposition 2.24. Sigui E un K-espai vectorial de dimensió $n < \infty$ i $F \subset E$ un subespai vectorial, llavors $\dim(E/F) = \dim E - \dim F$.

1.2.3 | Aplicacions lineals

Definition 2.25. Sigui E, F dos K-espais vectorials. Una aplicació $f : E \to F$ és lineal si es compleix $f(\lambda v_1 + \mu v_2) = \lambda f(v_1) + \mu f(v_2)$ $\forall v_1, v_2 \in E$ i $\forall \lambda, \mu \in K$.

Proposition 2.26. Si $f : E \to F$ i $g : F \to G$ són aplicacions lineals, llavors la composició $g \circ f : E \to G$ és lineal.

Proposition 2.27. Si $f : E \to F$ és una aplicació lineal, llavors $f^{-1} : F \to E$ és lineal.

Proposition 2.28. Sigui $f : E \to F$ una aplicació lineal entre K-espais vectorials i siguin $G \subset E$ i $H \subset F$ subespais vectorials. Aleshores:

1. $f(G) = \{f(u) \mid u \in G\} \subset F$ és un subespai vectorial.
2. $f^{-1}(H) = \{u \in E \mid f(u) \in H\} \subset E$ és un subespai vectorial.

Si $G = E$ i $H = \{0\}$, aleshores:

1. $f(E)$ és la imatge de f i es denota Im f.
2. $f^{-1}(0)$ és el nucli de f i es denota Ker f.

Proposition 2.29. Si E, F són K-espais vectorials de dimensió finita i $f : E \to F$, aleshores f és injectiva $\iff \dim(\text{Ker } f) = 0$ i f és exhaustiva $\iff \dim(\text{Im } f) = \dim F$.

Definition 2.30.

1. Un monomorfisme és una aplicació lineal injectiva.
2. Un epimorfisme és una aplicació lineal exhaustiva.
3. Un isomorfisme és una aplicació lineal bijectiva.
4. Un endomorfisme és una aplicació lineal d'un espai vectorial en ell mateix.
5. Un automorfisme és un endomorfisme bijectiu.

Lemma 2.31. Donada $f : E \to F$ una aplicació lineal, on E, F són K-espais vectorials, i $u_1, \dots, u_k \in E$ es compleix $\langle f(u_1), \dots, f(u_k)\rangle = f(\langle u_1, \dots, u_k\rangle)$.

Theorem 2.32 (Isomorfisme de coordenació). Sigui E un K-espai vectorial de dimensió n i $B = (u_1, \dots, u_n)$ una base de E. Llavors l'aplicació $f : K^n \to E$, $f(a_1, \dots, a_n) = a_1u_1 + \cdots a_nu_n$ és un isomorfisme.

Theorem 2.33 (Teorema de l'isomorfisme). Sigui $f : E \to F$ una aplicació lineal. Existeix un isomorfisme $\tilde{f} : E/\text{Ker } f \to \text{Im } f$ complint $f = i \circ \tilde{f} \circ \pi$ on $\pi : E \to E/\text{Ker } f$ i $i : \text{Im } f \to F$.

Corollary 2.34. Sigui $f : E \to F$ una aplicació lineal i suposem que $\dim E = n < \infty$. Llavors $n = \dim(\text{Ker } f) + \dim(\text{Im } f)$.

Corollary 2.35. Siguin E, F dos K-espais vectorials de dimensió $n < \infty$ i $f : E \to F$ una aplicació lineal. Llavors f és injectiva $\iff$ f és exhaustiva $\iff$ f és bijectiva.

Theorem 2.36 (Teoremes d'Emmy Noether). Sigui E un K-espai vectorial, $F, G \subset E$ dos subespais:

1. Existeix un isomorfisme $F/(F \cap G) \cong (F + G)/G$.
2. Si $G \subset F \subset E$, existeix un isomorfisme $(E/G)/(F/G) \cong E/F$.

Theorem 2.37. Siguin E, F dos K-espais vectorials i siguin $B = (u_1, \dots, u_n)$ una base de E i $v_1, \dots, v_n \in F$ vectors qualssevol. Llavors existeix una única aplicació lineal $f : E \to F$ tal que $f(u_i) = v_i$, $i = 1, \dots, n$.

Proposition 2.38. Siguin $f : E \to F$ i $g : F \to G$ dues aplicacions lineals entre K-espais vectorials i siguin B, B', B'' bases de E, F i G respectivament. Llavors $g \circ f : E \to G$ té matriu $[g \circ f]_{B,B''} = [g]_{B',B''}[f]_{B,B'}$ en les bases B, B''.

Corollary 2.39. Les matrius de canvi de base són invertibles i $[id]^{-1}_{B,B'} = [id]_{B',B}$.

Proposition 2.40. Donada $f : E \to F$ una aplicació lineal i bases B_1, B_2 de E i B_1', B_2' de F es compleix: $[f]_{B_2,B_2'} = [id]_{B_1',B_2'}[f]_{B_1,B_1'}[id]_{B_2,B_1}$.

Theorem 2.41. Donada qualsevol aplicació lineal $f : E \to F$ on $\dim E = n, \dim F = m$. Existeixen bases B_0 de E i B_0' de F en les quals $[f]_{B_0,B_0'} = \left(\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right)$ on $r = \dim \text{Im } f$ i $[f]_{B_0,B_0'} = [id]_{B',B_0'}[f]_{B,B'}[id]_{B_0,B}$ $\forall B, B'$ base de E i F respectivament.

Lemma 2.42. Donats dos K espais vectorials E, F el conjunt $\mathcal{L}(E, F) = \{f \mid f$ és una aplicació lineal de E a $F\}$ és un K-espai vectorial i es compleix que $(\lambda f + \mu g)(v) = \lambda f(v) + \mu g(v)$ $\forall f, g \in \mathcal{L}(E, F)$ i $\lambda, \mu \in K$.

Proposition 2.43. Siguin E, F dos K-espais vectorials amb $\dim E = n < \infty$, $\dim F = m < \infty$. Per a tota base B de E i B' de F, l'aplicació $\varphi : \mathcal{L}(E, F) \to \mathcal{M}_{m\times n}(K)$ definida per $\varphi(f) = [f]_{B,B'}$ és un isomorfisme.

Corollary 2.44. Si $\dim E = n$ i $\dim F = m$ aleshores $\dim \mathcal{L}(E, F) = mn$.

Definition 2.45. L'espai dual d'un K-espai vectorial és $\mathcal{L}(E,K) = E^*$. En el cas $\dim E = n < \infty$, escollir una base B determina un isomorfisme $\varphi : E^* \to \mathcal{M}_{1\times n}(K)$, definit per $\varphi(\omega) = [\omega]_{B,B'}$. Deduïm, doncs, que $\dim E^* = \dim E$.

Definition 2.46. Donats E un K-espai vectorial de dimensió finita i $B = (u_1, \dots, u_n)$ una base de E. La base dual de B és la base de E^* formada per $(\eta_1, \dots, \eta_n)$ on $\eta_i(u_j) = \delta_{ij}$.

Lemma 2.47. Sigui E un K-espai vectorial de dimensió $n < \infty$ i sigui $B = (u_1, \dots, u_n)$ una base de E. $\forall u \in E$, $(u_1^*(u), \dots, u_n^*(u)) = [u]_B \in K^n$ on $(u_1^*, \dots, u_n^*)$ és la base dual de B.

Lemma 2.48. Sigui E un K-espai vectorial de dimensió $n < \infty$ i sigui $B = (u_1, \dots, u_n)$ una base de E. $\forall \omega \in E^*$, $(\omega(u_1), \dots, \omega(u)) = [\omega]_{B^*}$.

Definition 2.49. Si $f \in \mathcal{L}(E,F)$ l'aplicació $f^* : F^* \to E^*$ definida per $f^*(\omega) = \omega \circ f$ s'anomena aplicació dual de f. Aquesta aplicació és lineal.

Theorem 2.50. Siguin E, F dos K-espais vectorials de dimensió finita i siguin B, B' bases de E, F respectivament. Llavors $[f^*]_{B'^*,B^*} = ([f]_{B,B'})^t$.

Definition 2.51. Donat un K-espai vectorial E, el bidual és el K-espai vectorial definit per $(E^*)^* = \mathcal{L}(E^*, K)$. A més, si $\dim E = n < \infty$, l'aplicació $f : E \to (E^*)^*$ definida per $f(v) = \phi_v$ ($\phi_v : E^* \to K$, $\phi_v(\omega) = \omega(v)$) és un isomorfisme natural.

Definition 2.52. Sigui E un K-espai vectorial i F un subespai vectorial de E^*. El subespai de E incident de F és $F^{inc} = \{v \in E \mid \omega(v) = 0 \ \forall \omega \in F\}$.

Theorem 2.53. Sigui E un K-espai vectorial de dimensió $n < \infty$ i $F \subset E^*$ un subespai amb $\dim F = m$. Llavors $\dim F^{inc} = n - m$.

Definition 2.54. Donat un subespai vectorial i $F \subset E$, el seu subespai incident és $F^{inc} = \{\omega \in E^* \mid \omega(v) = 0 \ \forall v \in F\}$.

Proposition 2.55. $(F^{inc})^{inc} = F$ tant si $F \subset E$ com si $F \subset E^*$.

1.2.4 | Classificació d'endomorfismes

Definition 2.56. Dues matrius $M, N \in \mathcal{M}_n(K)$ s'anomenen similars si existeix $P \in \mathcal{M}_n(K)$ invertible tal que $M = P^{-1}NP$.

Proposition 2.57. Donats $f, g \in \mathcal{L}(E)$ on E és un K espai vectorial de dimensió $n < \infty$:

1. f i g són similars $\iff \forall B$ base de E les matrius $[f]_B$ i $[g]_B$ són similars.
2. f i g són similars $\iff \exists h$ automorfisme tal que $g = h^{-1}fh$.

Definition 2.58. Una matriu $A \in \mathcal{M}_n(K)$ és diagonalitzable si és similar a una matriu diagonal. Un endomorfisme és diagonalitzable si la seva matriu en alguna base és diagonalitzable.

Definition 2.59. Donat $f \in \mathcal{L}(E)$ diem que un vector $u \in E$, $u \neq 0$ és vector propi de f de valor propi λ si $f(u) = \lambda u$.

Lemma 2.60. Donats $f \in \mathcal{L}(E)$ i $\lambda \in K$, els vectors propis de f de valor propi λ són els vectors no nuls del subespai $Ker(f - \lambda id)$ (subespai propi de valor propi λ).

Definition 2.61. Donada una matriu $A \in \mathcal{M}_n(K)$, el polinomi $p_A(\lambda) = \det(A - \lambda I_n)$ s'anomena polinomi característic de A.

Proposition 2.62. Donat $f \in \mathcal{L}(E)$, vectors propis de f de valors propis diferents són linealment independents.

Proposition 2.63. Sigui E un K-espai vectorial de dimensió $n < \infty$ i sigui λ una arrel del polinomi característic $p_f(x)$ de multiplicitat m. Llavors $1 \leq \dim(\mathrm{Ker}(f - \lambda id)) \leq m$.

Theorem 2.64 (Teorema de diagonalització). Sigui $f \in \mathcal{L}(E)$, f és diagonalitzable si i només si:

1. $p_f(x) = (-1)^n(x - \lambda_1)^{m_1} \cdots (x - \lambda_k)^{m_k}$ amb $\lambda_1, \dots, \lambda_k \in K$ diferents.
2. $\dim(\mathrm{Ker}(f - \lambda_i id)) = m_i$.

Corollary 2.65. Si $n = \dim E$ i f té n valors propis diferents, f és diagonalitzable.

Definition 2.66. El polinomi mínim de f és un polinomi $P \in K[x]$ tal que:

- $P(f) = 0$.
- P és mònic.
- P és de grau mínim.

Theorem 2.67 (Teorema de Cayley-Hamilton). Sigui K un cos, $n \geq 1$ i $A \in \mathcal{M}_n(K)$. Llavors $m_A(x) \mid p_A(x) \mid m_A(x)^n$. Sigui K un cos i $f \in \mathcal{L}(E)$, $\dim_K E = n$. Llavors $m_f(x) \mid p_f(x) \mid m_f(x)^n$.

Definition 2.68. Un cos satisfent que tot polinomi de grau ≥ 1 factoritzi completament en factors lineals s'anomena algebraicament tancat.

Definition 2.69. Donat $f \in \mathcal{L}(E)$ diem que $W \subseteq E$ és un subespai invariant de E per f si $f(W) \subseteq W$.

Lemma 2.70. Donat $f \in \mathcal{L}(E)$ si $E = W_1 \oplus W_2$ amb W_1, W_2 subespais invariants, aleshores $p_f(x) = p_{f|W_1}(x)p_{f|W_2}(x)$ i $m_f(x) = \mathrm{mcm}(m_{f|W_1}(x), m_{f|W_2}(x))$.

Theorem 2.71. Sigui $f \in \mathcal{L}(E)$, $\dim E = n < \infty$. Si $p_f(x) = q_1(x)^{n_1} \cdots q_r(x)^{n_r}$ i $m_f(x) = q_1(x)^{m_1} \cdots q_r(x)^{m_r}$ amb $q_i(x)$ factors irreductibles diferents. Aleshores $E = \mathrm{Ker}(q_1(f)^{m_1}) \oplus \cdots \oplus \mathrm{Ker}(q_r(f)^{m_r})$. A més $\mathrm{Ker}(q_i(f)^{m_i}) = n_i \deg(q_i(x))$.

Theorem 2.72. Si $p_f(x) = (x-\lambda_1)^{n_1} \cdots (x-\lambda_k)^{n_k}$ i g compleix:

1. $p_f(x) = p_g(x)$
2. $m_f(x) = m_g(x)$
3. $\dim(\mathrm{Ker}((f-\lambda id)^r)) = \dim(\mathrm{Ker}((g-\lambda id)^r)) \; \forall \lambda \in K \; \forall r \geq 1$,

llavors $f \sim g$.

Proposition 2.73. Donats E un K-espai vectorial de dimensió $n < \infty$, $A \in \mathcal{M}_n(K)$. Si $p_A(x) = \pm(x-\lambda_1)^{n_1} \cdots (x-\lambda_k)^{n_k}$, existeix una matriu invertible P complint:

$$P^{-1}AP = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_e \end{pmatrix}$$

on $J_1, \ldots, J_e$ són blocs de Jordan de valors propis $\lambda_1, \ldots, \lambda_k$ complint:

1. Per a cada λ_i la suma de les mides dels blocs de Jordan de valor propi λ_i és n_i.
2. Les mides dels blocs de Jordan estan determinades per $\dim(\mathrm{Ker}((f-\lambda id)^r))$.

1.2.5 | Formes bilineals simètriques

Definition 2.74. Siguin K un cos i E, F, G tres K-espais vectorials. Diem que una aplicació $\varphi : E \times F \to G$ és bilineal si:

1. $\varphi(\lambda u_1 + \mu u_2, v) = \lambda\varphi(u_1, v) + \mu\varphi(u_2, v) \; \forall u_1, u_2 \in E$, $\forall v \in F$ i $\forall \lambda, \mu \in K$.
2. $\varphi(u, \lambda v_1 + \mu v_2) = \lambda\varphi(u, v_1) + \mu\varphi(u, v_2) \; \forall v_1, v_2 \in F$, $\forall u \in E$ i $\forall \lambda, \mu \in K$.

Definition 2.75. Una forma bilineal sobre el K-espai vectorial E és una aplicació lineal $\varphi : E \times E \to K$.

Definition 2.76. Una forma bilineal $\varphi : E \times E \to K$ és simètrica si $\varphi(u, v) = \varphi(v, u) \; \forall u, v \in E$.

Definition 2.77. Sigui $\varphi : E \times E \to K$ una forma bilineal i sigui $B = (v_1, \ldots, v_n)$ una base de E. La matriu de la forma bilineal φ respecte de la base B és la matriu:

$$[\varphi]_B = \begin{pmatrix} \varphi(v_1, v_1) & \varphi(v_1, v_2) & \cdots & \varphi(v_1, v_n) \\ \varphi(v_2, v_1) & \varphi(v_2, v_2) & \cdots & \varphi(v_2, v_n) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi(v_n, v_1) & \varphi(v_n, v_2) & \cdots & \varphi(v_n, v_n) \end{pmatrix}.$$

Proposition 2.78. Sigui $B = (u_1, \ldots, u_n)$ una base de E. Una forma bilineal φ sobre E és simètrica si, i només si, $[\varphi]_B$ és una matriu simètrica.

Proposition 2.79. Sigui $\varphi : E \times E \to K$ una forma bilineal. Siguin B, B' bases de E. Aleshores es compleix $[\varphi]_{B'} = ([id]_{B',B})^t [\varphi]_B [id]_{B',B}$.

Definition 2.80. Sigui $\varphi : E \times E \to \mathbb{R}$ una forma bilineal. Es diu que $u, v \in E$ són ortogonals si $\varphi(u, v) = 0$. Un vector no nul $v \in E$ és isòtrop si és ortogonal a ell mateix, és a dir, si $\varphi(v, v) = 0$. Sigui $B = (v_1, \ldots, v_n)$ una base de E. Es diu que B és una base ortogonal respecte de φ si $\varphi(v_i, v_j) = 0 \; \forall i \neq j$.

Theorem 2.81. Sigui $\varphi : E \times E \to \mathbb{R}$ una forma bilineal simètrica. Llavors E té una base ortogonal respecte de φ.

Definition 2.82. Sigui φ una forma bilineal simètrica sobre E. Sigui W un subespai vectorial de E. Definim l'ortogonal de W respecte de φ per

$$W^\perp = \{u \in E : \varphi(u, \omega) = 0 \; \forall \omega \in W\}$$

Definim el radical de φ per

$$\mathrm{rad}\ \varphi = E^\perp$$

Direm que φ és no singular si $\mathrm{rad}(\varphi) = \{0\}$.

Definition 2.83. Sigui φ una forma bilineal simètrica no singular sobre E. Donat un $u \in E$ definim $\varphi_u : E \to \mathbb{R}$, $\varphi_u(v) = \varphi(u, v)$. Llavors l'aplicació

$$\begin{aligned} \Phi : E &\to E^* \\ u &\mapsto \varphi_u \end{aligned}$$

és un isomorfisme.

Definition 2.84. Sigui φ una forma bilineal simètrica no singular sobre E. Sigui W un subespai vectorial de E. Llavors:

1. $\dim E = \dim W + \dim W^\perp$.
2. $(W^\perp)^\perp = W$.
3. Si la restricció de φ sobre W és no singular, llavors $E = W \oplus W^\perp$.

Definition 2.85. Sigui φ una forma bilineal simètrica sobre E. Direm que la suma $W_1 + W_2$ de dos subespais vectorials W_1 i W_2 de E és una suma ortogonal si és directa i $\varphi(u, v) = 0 \; \forall u \in W_1$ i $v \in W_2$. Escriurem $W_1 \perp W_2$ per denotar que la suma $W_1 + W_2$ és ortogonal.

Definition 2.86. Sigui φ una forma bilineal simètrica sobre E. Siguin W_1 i W_2 dos subespais vectorials de E tals que $E = W_1 \perp W_2$. Llavors per a cada vector $v \in E$ existeixen $v_1 \in W_1$ i $v_2 \in W_2$ únics tals que $v = v_1 + v_2$. L'aplicació $\pi : E \to W_i$ definida per $\pi(v) = v_i$ amb $v_i \in W_i$ es diu que és la projecció ortogonal de E sobre W_i segons la descomposició $E = W_1 \perp W_2$.

Definition 2.87. Una geometria ortogonal sobre $\mathbb{R}$ és un parell (E, φ), on E és un $\mathbb{R}$-espai vectorial i φ és una forma bilineal simètrica sobre E.

Definition 2.88. Siguin (E_1, φ_1) i (E_2, φ_2) dues geometries ortogonals sobre $\mathbb{R}$. Una isometria de (E_1, φ_1) a (E_2, φ_2) és un isomorfisme $f : E_1 \to E_2$ tal que

$$\varphi_2(f(u), f(v)) = \varphi_1(u, v)$$

per a tot $u, v \in E_1$. Direm que (E_1, φ_1) i (E_2, φ_2) són isomètriques si existeix una isometria de (E_1, φ_1) a (E_2, φ_2).

Definition 2.89. Sigui E un $\mathbb{R}$-espai vectorial. Direm que dues formes bilineals simètriques φ_1, φ_2 sobre E són equivalents si, i només si, (E, φ_1) i (E, φ_2) són isomètriques.

Definition 2.90. Siguin $A, B \in \mathcal{M}_n(\mathbb{R})$. Direm que A i B són congruents si existeix una matriu $P \in \mathcal{M}_n(\mathbb{R})$ invertible tal que $A = P^t B P$.

Proposition 2.91. Sigui E un $\mathbb{R}$-espai vectorial de dimensió $n < \infty$ i φ_1, φ_2 formes bilineals simètriques sobre E. Sigui B_1 una base de V. Llavors les condicions següents són equivalents:

1. Les geometries ortogonals (E, φ_1) i (E, φ_2) φ i ψ són isomètriques.
2. Existeix una bases B_2 de E tal que $[\varphi_1]_{B_1} = [\varphi_2]_{B_2}$.
3. Les matrius $[\varphi_1]_{B_1}$ i $[\varphi_2]_{B_2}$ són congruents.

Theorem 2.92 (Teorema de Sylvester o Llei d'inèrcia). Sigui E un $\mathbb{R}$-espai vectorial de dimensió $n < \infty$. Sigui φ una forma bilineal simètrica sobre E. Llavors existeix una base B de E tal que

$$[\varphi]_B = \begin{pmatrix} 0 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 0 & & & & & & \\ & & & 1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & 1 & & & \\ & & & & & & -1 & & \\ & & & & & & & \ddots & \\ & & & & & & & & -1 \end{pmatrix}$$

on a la diagonal hi ha r_0 $0's$, r_+ $1's$ i r_- $-1's$ i r, r_+, r_- no depenen de la base B.

Definition 2.93. Sigui φ una forma bilineal simètrica sobre un $\mathbb{R}$-espai vectorial E de dimensió $n < \infty$. Sigui B una base ortogonal de E respecte de φ. Definim el rang de φ com $\operatorname{rang}(\varphi) = \operatorname{rang}([\varphi]_B)$. Definim la signatura de φ com $\operatorname{sig}(\varphi) = (r_+, r_-)$, on r_+ és el nombre de reals positius que hi ha a la diagonal de $[\varphi]_B$ i r_- és el nombre de reals negatius que hi ha a la diagonal de $[\varphi]_B$.

Theorem 2.94. Siguin (E_1, φ_1), (E_2, φ_2) dues geometries ortogonals sobre $\mathbb{R}$ de dimensió finita. Llavors (E_1, φ_1) i (E_2, φ_2) són isomètriques si, i només si, $\dim E_1 = \dim E_2$ i $\operatorname{sig}(\varphi_1) = \operatorname{sig}(\varphi_2)$.

Definition 2.95. Sigui φ una forma bilineal simètrica sobre un $\mathbb{R}$-espai vectorial E. Es diu que φ és definida positiva si $\forall v \in E$, $v \neq 0$, tenim $\varphi(v, v) > 0$. Es diu que φ és definida negativa si $\forall v \in E$, $v \neq 0$, tenim $\varphi(v, v) < 0$.

Definition 2.96. Un producte escalar sobre un $\mathbb{R}$-espai vectorial E és una forma bilineal simètrica definida positiva sobre E. Un espai vectorial euclidià és un parell (E, φ), on E és un $\mathbb{R}$-espai vectorial i φ és un producte escalar sobre E.

Theorem 2.97 (Desigualtat de Cauchy-Schwartz). Sigui φ un producte escalar sobre un $\mathbb{R}$-espai vectorial E, llavors:

$$\varphi(u, v)^2 \leq \varphi(u, u)\varphi(v, v) \quad \forall u, v \in E$$

Definition 2.98. Sigui E un $\mathbb{R}$-espai vectorial. Una norma sobre E és una aplicació

$$\begin{aligned} \| \| : E &\to \mathbb{R} \\ u &\mapsto \|u\| \end{aligned}$$

tal que

1. $\|u\| = 0 \iff u = 0$.
2. $\|\lambda u\| = |\lambda| \|u\|$, $\forall u \in E$, $\lambda \in \mathbb{R}$.
3. $\|u + v\| \leq \|u\| + \|v\|$, $\forall u, v \in E$.

Proposition 2.99. Sigui (E, φ) un espai euclidià. Llavors l'aplicació

$$\begin{aligned} \| \|_\varphi : E &\to \mathbb{R} \\ u &\mapsto \|u\|_\varphi = \sqrt{\varphi(u, u)} \end{aligned}$$

és una norma, que es diu norma associada al producte escalar φ.

Definition 2.100. Sigui (E, φ) un espai vectorial euclidià de dimensió finita. Diem que una base $B = (v_1, \ldots, v_n)$ és ortonormal respecte de φ si és ortogonal respecte de φ i $\|v_i\|_\varphi = 1$ per a $i = 1, \ldots, n$.

Corollary 2.101. Sigui (E, φ) un espai vectorial euclidià de dimensió finita. Llavors E té una base ortonormal respecte φ.

Definition 2.102. Sigui (E, φ) un espai vectorial euclidià. Siguin $u, v \in E \setminus \{0\}$. Definim l'angle respecte φ entre u i v com l'únic $\alpha \in [0, \pi]$ tal que:

$$\cos \alpha = \frac{\varphi(u, v)}{\|u\|_\varphi \|v\|_\varphi}$$

Definition 2.103. Sigui (E, φ) un espai vectorial euclidià. Sigui $f \in \mathcal{L}(E)$. Definim l'adjunt de f respecte de φ com l'únic endomorfisme f' de E tal que $\varphi(f(u), v) = \varphi(u, f'(v))$ $\forall u, v \in E$. Si $f = f'$ diem que f és autoadjunt.

Lemma 2.104. Sigui (E, φ) un espai vectorial euclidià de dimensió $n < \infty$. Sigui f un endomorfisme autoadut de E. Llavors existeixen $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ tals que

$$p_f(x) = (x - \lambda_1) \cdots (x - \lambda_n)$$

Definition 2.105. Sigui $A \in \mathcal{M}_n(K)$. Aleshores A és ortogonal si, i només si, $PP^t = P^tP = I_n$.

Theorem 2.106 (Teorema espectral). Sigui (E, φ) un espai vectorial euclidià de dimensió $n < \infty$. Sigui $f \in \mathcal{L}(E)$ autoadjunt de E. Llavors E té una base ortonormal de vectors propis de f. En particular, l'endomorfisme f diagonalitza.

Corollary 2.107. Tota matriu simètrica $A \in \mathcal{M}_n(\mathbb{R})$ és diagonalitzable.

Definition 2.108. Donada una matriu $A \in \mathcal{M}_{m\times n}(\mathbb{C})$, $A = (a_{ij})$ denotem per $\overline{A} = (\overline{a_{ij}})$ la conjugada de A. $\forall A_1, A_2, A, B$ matrius de mides adequades i $\forall \lambda \in \mathbb{C}$ es satisfan les següents propietats:

1. $\overline{A_1 + A_2} = \overline{A_1} + \overline{A_2}$
2. $\overline{AB} = \overline{AB}$
3. $\overline{\lambda A} = \overline{\lambda A}$

Theorem 2.109 (Regla dels signes de Descartes). Donat un polinomi $P(x) = a_d x^d + \cdots + a_0$:

1. El nombre d'arrels positives de $P(x)$ és com a molt igual al nombre de canvis de signe en $[a_d, a_{d-1}, \ldots, a_1, a_0]$.
2. SI $P(x) = a_d(x - \alpha_1)^{n_1} \cdots (x - \alpha_r)^{n_r}$, aleshores el nombre d'arrels positives és igual al nombre de canvis de signe (comptant les arrels amb multiplicitat).

1.3 Funcions de variable real

1.3.1 | La recta real

Definition 3.1. Sigui $(K, +, \cdot)$ un cos. Diem que K amb una relació d'ordre total $(\leq)$ és un cos ordenat si es verifiquen les següents propietats:

1. Si $x, y, z \in K$ i $x \leq y$, aleshores $x + z \leq y + z$.
2. Si $x, y \in K$ i $x \geq 0$ i $y \geq 0$, aleshores $x \cdot y \geq 0$.

Definition 3.2. Sigui K un cos ordenat i $A \subset K$. Diem que A està acotat superiorment (respectivament inferiorment) si existeix $M \in K$, que anomenarem cota superior (respectivament cota inferior) de A, tal que $x \leq M$ (respectivament $x \geq M$) per a tot $x \in A$.

Definition 3.3. Sigui K un cos ordenat i $A \subset K$ acotat superiorment (respectivament inferiorment). Diem que α cota superior (respectivament inferior) de A és el suprem (respectivament ínfim) de A si qualsevol cota superior β (respectivament inferior) verifica $\beta \geq \alpha$ (respectivament $\beta \leq \alpha$). El suprem de A i l'ínfim de A, quan existeixin, els designem per $\sup A$ i $\inf A$ respectivament.

Theorem 3.4 (Axioma del suprem). Existeix un únic cos ordenat amb la propietat que tot conjunt acotat superiorment té suprem.

Lemma 3.5. Si $\alpha = \sup A$, aleshores per a tot $\varepsilon > 0$ l'interval $(\alpha - \varepsilon, \alpha]$ conté punts de A.

Proposition 3.6. Els nombres naturals no estan acotats superiorment a $\mathbb{R}$.

Proposition 3.7. Entre dos nombres reals sempre n'hi ha un de racional i un d'irracional.

Definition 3.8. Sigui A un conjunt. Diem que A és numerable si existeix una aplicació bijectiva de A en $\mathbb{N}$.

Proposition 3.9. Tot subconjunt infinit de $\mathbb{N}$ és numerable.

Corollary 3.10. Tot subconjunt infinit d'un conjunt numerable és numerable.

Corollary 3.11. Sigui A un conjunt amb infinits elements. Perquè A sigui numerable, n'hi ha prou que existeixi una aplicació injectiva de A en $\mathbb{N}$.

Proposition 3.12. Si A i B són numerables, aleshores $A \times B$ també ho és.

Theorem 3.13. $\mathbb{Q}$ no és numerable.

Theorem 3.14. $\mathbb{R}$ no és numerable.

1.3.2 | Successions

Definition 3.15. Diem que la successió (a_n) està acotada superiorment (respectivament acotada inferiorment) si existeix un nombre real K de manera que $a_n \leq K$ (respectivament $a_n \geq K$) per a tot $n \in \mathbb{N}$.

Definition 3.16. Diem que $\lim a_n = l$ si $\forall \varepsilon > 0$ $\exists n_0$ tal que $|a_n - l| < \varepsilon$ $\forall n > n_0$.

Lemma 3.17. Siguin (a_n) i (b_n) successions convergents amb límits a i b respectivament. Aleshores els següents fets són certs:

1. Les successions $(a_n + b_n)$ i $(a_n b_n)$ són convergents i $\lim(a_n + b_n) = a + b$ i $\lim(a_n b_n) = ab$.
2. Si $a \neq 0$, aleshores $a_n \neq 0$ per a n prou gran la successió $(\frac{1}{a_n})$ és convergent i $\lim \frac{1}{a_n} = \frac{1}{a}$.

Theorem 3.18. Tota successió monòtona i acotada és convergent.

Lemma 3.19 (Teorema del sandvitx). Siguin (a_n), (b_n) i (c_n) tres successions verificant $a_n \leq b_n \leq c_n$ $\forall n \in \mathbb{N}$. Suposem, a més, que $\lim a_n = \lim c_n = l$. Aleshores (b_n) és convergent i $\lim b_n = l$.

Lemma 3.20. $e = \lim S_n = \lim T_n$ on $S_n = \sum_{i=0}^{n} \frac{1}{i!}$ i $T_n = \left(1 + \frac{1}{n}\right)^n$.

Theorem 3.21. El nombre e és irracional.

Lemma 3.22. Si $\lim a_n = l$, llavors qualsevol successió parcial de (a_n) té també límit l.

Proposition 3.23. a és un punt d'acumulació de (a_n) si i només si existeix (a_{k_n}) parcial de (a_n) amb $\lim a_{k_n} = a$.

Proposition 3.24. Tota successió té una parcial monòtona.

Theorem 3.25 (Teorema de Bolzano-Weierstraß). Tota successió de punts d'un interval tancat té una parcial convergent a un punt de l'interval.

Lemma 3.26. Sigui (a_n) acotada. Llavors (a_n) és convergent si i només si $\liminf a_n = \limsup a_n$. En aquest cas es té que $\lim a_n = \limsup a_n = \liminf a_n$

Definition 3.27. Diem que la successió (a_n) és una successió de Cauchy si $\forall \varepsilon > 0$, $\exists n_0$ tal que $|a_n - a_m| < \varepsilon$ per a qualsevol $n, m > n_0$.

Theorem 3.28. Una successió és convergent si i només si és de Cauchy.

Theorem 3.29 (Criteri de Stolz). Sigui (a_n) una successió estrictament monòtona i (b_n) una successió qualsevol. Suposem a més que $\lim \frac{b_n - b_{n-1}}{a_n - a_{n-1}} = l \in \mathbb{R} \cup \{\pm\infty\}$. Aleshores les següents afirmacions són certes:

1. Si $\lim a_n = \pm\infty$, llavors $\lim \frac{b_n}{a_n} = l$.

2. Si $\lim b_n = \lim a_n = 0$, llavors $\lim \frac{b_n}{a_n} = l$.

1.3.3 | Continuïtat

Definition 3.30. Sigui $f : [a, b] \to \mathbb{R}$ una funció i $x_0 \in (a, b)$. Diem que l és el límit de la funció f en el punt x_0 i escrivim $\lim_{x \to x_0} f(x) = l$, si $\forall \varepsilon > 0$, $\exists \delta > 0$ de manera que $|f(x) - l| < \varepsilon$ sempre que $|x - x_0| < \delta$.

Lemma 3.31. Sigui $f : [a, b] \to \mathbb{R}$ i $x_0 \in (a, b)$. Aleshores $\lim_{x \to x_0} f(x) = l$ si i només si per a tota successió a_n de punts de $(a, b) \setminus \{x_0\}$ amb $\lim a_n = x_0$ es compleix que $\lim f(a_n) = l$.

Definition 3.32. Sigui I un interval, $f : I \to \mathbb{R}$ una funció i $x_0 \in I$. Diem que f és contínua en x_0 si existeix el límit de f en x_0 i és igual a $f(x_0)$.

Lemma 3.33. f és contínua en $x_0 \in I$ si i només si per a tota successió $x_n \in I$ amb $\lim x_n = x_0$ es té que $\lim f(x_n) = f(x_0)$.

Proposition 3.34. Siguin $f, g : I \to \mathbb{R}$ contínues en $x_0 \in I$. Llavors:

1. $f + g$ i fg són contínues en x_0.

2. Si $f(x_0) > 0$ (respectivament $f(x_0) < 0$), aleshores $f(x) > 0$ (respectivament $f(x) < 0$) en un entorn de x_0. A més, en ambdós casos, $\frac{1}{f}$ és contínua en x_0.

Proposition 3.35. Siguin $f : I \to \mathbb{R}$ i $g : J \to \mathbb{R}$. Sigui $x_0 \in I$ amb $f(x_0) \in J$ i suposem que f és contínua en x_0 i g també ho és en $f(x_0)$. Aleshores $g \circ f$ és contínua en x_0.

Theorem 3.36 (Teorema de Weierstraß). Sigui $f : [a, b] \to \mathbb{R}$ contínua. Aleshores f està acotada en $[a, b]$. A més, existeixen $y, z \in [a, b]$ tals que $f(y) \leq f(x) \leq f(z)$ $\forall x \in [a, b]$.

Theorem 3.37 (Teorema de Bolzano). Sigui f contínua en $[a, b]$. Si $f(a)f(b) < 0$, llavors existeix $c \in (a, b)$ amb $f(c) = 0$.

Theorem 3.38. Sigui $f : (c, d) \to \mathbb{R}$ contínua. Si f és injectiva i contínua, aleshores f és monòtona. A més, f^{-1} és també contínua en $f((c, d))$.

1.3.4 | Derivació

Definition 3.39. Si $f : (a, b) \to \mathbb{R}$. Diem que f és derivable en $x_0 \in (a, b)$ si existeix el límit

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}$$

Proposition 3.40. Siguin f, g definides en un entorn de a i derivables en a. Aleshores $f + g$ i fg són derivables en a i

$$(f + a)'(a) = f'(a) + g'(a),$$
$$(fg)'(a) = f'(a)g(a) + f(a)g'(a).$$

Si a més $f(a) \neq 0$, llavors $\frac{1}{f}$ està definida en un entorn de a, és derivable en a i

$$\left(\frac{1}{f}\right)'(a) = -\frac{f'(a)}{f^2(a)}$$

Proposition 3.41 (Regla de la cadena). Siguin $g : (a, b) \to \mathbb{R}$ i $f : (c, d) \to \mathbb{R}$. Suposem que g és derivable en $x \in (a, b)$ i f és derivable en $g(x) \in (c, d)$. Aleshores $f \circ g$ és derivable en x i $(f \circ g)'(x) = f'(g(x))g'(x)$.

Proposition 3.42. Sigui $f : (a, b) \to \mathbb{R}$ injectiva i contínua en (a, b) i derivable en $c \in (a, b)$ amb $f'(c) \neq 0$. Aleshores f^{-1} és derivable en $f(c)$ i

$$(f^{-1})'(f(c)) = \frac{1}{f'(c)}$$

Proposition 3.43. Sigui $f : I \to \mathbb{R}$ i sigui $c \in I$ extrem local de f. Si f és derivable en c, $f'(c) = 0$.

Theorem 3.44 (Teorema de Rolle). Sigui $f : [a, b] \to \mathbb{R}$ contínua i derivable en (a, b). Suposem $f(a) = f(b)$. Aleshores existeix un punt $c \in (a, b)$ amb $f'(c) = 0$.

Theorem 3.45 (Teorema del valor mitjà). Sigui $f : [a, b] \to \mathbb{R}$ contínua en $[a, b]$ i derivable en (a, b). Aleshores existeix un punt $c \in (a, b)$ verificant

$$f'(c) = \frac{f(b) - f(a)}{b - a}$$

Corollary 3.46. Sigui f derivable en (a, b) verificant $f'(x) > 0$ (respectivament $f'(x) < 0$) $\forall x \in (a, b)$. Aleshores f és creixent (respectivament decreixent) en (a, b).

Theorem 3.47 (Teorema de Cauchy). Siguin $f, g : [a, b] \to \mathbb{R}$ contínues en $[a, b]$ i derivables en (a, b). Aleshores existeix un punt $c \in (a, b)$ verificant

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

Theorem 3.48 (Regla de l'Hôpital). Suposem que f, g són dues funcions definides en un entron de a i que o bé

$\lim_{x\to a} f = \lim_{x\to a} g = 0$, o bé $\lim_{x\to a} g = \infty$. Suposem també que existeix el límit $\lim_{x\to a} \frac{f'(x)}{g'(x)}$. Aleshores també existeix el $\lim_{x\to a} \frac{f(x)}{g(x)}$ i

$$\lim_{x\to a} \frac{f(x)}{g(x)} = \lim_{x\to a} \frac{f'(x)}{g'(x)}$$

Theorem 3.49 (Teorema de Darboux). Sigui $f : (a,b) \to \mathbb{R}$ derivable i suposem que existeixen $x, y \in (a,b)$, $x < y$ amb $f'(x)f'(y) < 0$. Aleshores existeix $z \in (x,y)$ tal que $f'(z) = 0$.

1.3.5 | Convexitat i segona derivada

Definition 3.50. Diem que $f : I \to \mathbb{R}$ és convexa si donats dos punts qualssevol $a, b \in I$, $a < b$ el segment que uneix els punts $(a, f(a))$ i $(b, f(b))$ queda per damunt de la gràfica en (a,b). Diem que $f : I \to \mathbb{R}$ és còncava si donats dos punts qualssevol $a, b \in I$, $a < b$ el segment que uneix els punts $(a, f(a))$ i $(b, f(b))$ queda per sota de la gràfica en (a,b).

Lemma 3.51. f és convexa en I si i només si per a qualssevol $a, x, b \in I$ amb $a < x < b$ es té:

$$\frac{f(x) - f(a)}{x - a} < \frac{f(b) - f(a)}{b - a},$$

o, equivalentment,

$$\frac{f(b) - f(a)}{b - a} < \frac{f(b) - f(x)}{b - x}.$$

Si f és còncava, les desigualtats s'inverteixen.

Theorem 3.52. Sigui f derivable en I. Llavors f és convexa (respectivament còncava) si i només si f' és creixent (respectivament decreixent).

Theorem 3.53. Sigui f derivable en I. Llavors f és convexa (respectivament còncava) si i només si qualsevol tangent a la gràfica queda per sota (respectivament sobre) de la gràfica excepte en el punt de contacte.

Theorem 3.54. Sigui f dues vegades derivable en I. Llavors les següents afirmacions són certes:

1. Si f és convexa (respectivament còncava) en I aleshores $f''(x) \geq 0$ (respectivament $f''(x) \leq 0$) $\forall x \in I$.
2. Si $f''(x) > 0$ (respectivament $f''(x) < 0$) $\forall x \in I$ aleshores f és convexa (respectivament còncava) en I.

Proposition 3.55. Sigui f dues vegades derivable en I. Aleshores les següents afirmacions són certes:

1. Si a és un punt d'inflexió aleshores $f''(a) = 0$.
2. Suposem a més que f'' és contínua en $a \in I$. Aleshores si $f''(a) > 0$ (respectivament $f''(a) < 0$) aleshores f és convexa (respectivament còncava) en a.

1.3.6 | Aproximació polinòmica

Definition 3.56. Diem que f i g tenen un contacte d'ordre $\geq n$ en a si

$$\lim_{x\to a} \frac{f(x) - g(x)}{(x-a)^n} = 0$$

Definition 3.57. Diem que f és de classe $\mathcal{C}^n$ en a si f és n vegades derivable en un entorn de a i $f^{(n)}$ és contínua en aquest entorn. Diem que f és de classe $\mathcal{C}^\infty$ en a si f és de classe $\mathcal{C}^n$ en a per a tot $n \in \mathbb{N}$.

Theorem 3.58. Sigui f n vegades derivable en a. Aleshores el polinomi

$$P_{n,f,a} = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \\ + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n$$

té un contacte amb f d'ordre $\geq n$ en a.

Theorem 3.59. Sigui f derivable n vegades en a. Si $f'(a) = f''(a) = \cdots = f^{(n-1)}(a) = 0$ i $f^{(n)}(a) \neq 0$ llavors:

1. Si n és senar a no és extrem relatiu.
2. Si n és parell i $f^{(n)} > 0$, aleshores a és un mínim relatiu.
3. Si n és parell i $f^{(n)} < 0$, aleshores a és un màxim relatiu.

Theorem 3.60. Sigui f derivable $n+1$ vegades en I un entorn de a. Sigui $P = P_{n,f,a}$ i $R_n = f - P$. Sigui $x \in I$. Llavors,

1. Fórmula de Cauchy:

   $$R_n(x) = \frac{f^{(n+1)}(\xi)}{n!}(x-\xi)^n(x-a),$$

   per algun ξ entre a i x.
2. Fórmula de Lagrange:

   $$R_n(x) = \frac{f^{(n+1)}(\eta)}{(n+1)!}(x-a)^{n+1},$$

   per algun η entre a i x.
3. Si, a més $f^{(n+1)}$, és integrable en $[a,x]$:

   $$R_n(x) = \int_a^x \frac{f^{(n+1)}(t)}{n!}(x-t)^n dt$$

Definition 3.61. Diem que f és analítica en a si és de classe $\mathcal{C}^\infty$ en un entorn de a i $\lim_{n\to\infty} R_n(x) = 0$ per a tot x en aquest entorn.

1.3.7 | Integral de Riemann

Definition 3.62. Definim la suma inferior i superior de f, respectivament, associada a la partició P com:

$$L(f,P)=\sum_{i=1}^{n} m_i(t_i-t_{i-1}) \qquad U(f,P)=\sum_{i=1}^{n} M_i(t_i-t_{i-1})$$

on $m_i=\inf\{f(x_i): x_i\in[t_{i-1},t_i]\}$ i $M_i=\sup\{f(x_i): x_i\in[t_{i-1},t_i]\}$.

Definition 3.63. Sigui f acotada en $I=[a,b]$. Diem que f és integrable en I si $\underline{\int_a^b} f = \overline{\int_a^b} f$ on $\underline{\int_a^b} = \inf\{L(f,P) : P \text{ partició de } [a,b]\}$ i $\overline{\int_a^b} = \sup\{U(f,P) : P \text{ partició de } [a,b]\}$

Lemma 3.64. Sigui f acotada en $I=[a,b]$. Aleshores f és integrable en I si i només si $\forall\varepsilon>0$ existeix una partició P de I amb $U(f,P)-L(f,P)<\varepsilon$.

Theorem 3.65. Sigui f monòtona i fitada a $I=[a,b]$. Llavors f és integrable en I.

Theorem 3.66. Sigui $I=[a,b]$ i f contínua en I. Aleshores f és uniformement contínua en I.

Theorem 3.67. Sigui f contínua en $I=[a,b]$. Aleshores f és integrable en I.

Proposition 3.68. Siguin f i g integrables en $I=[a,b]$ i $c\in\mathbb{R}$. Llavors $f+g$ i cf són integrables en I i

$$\int_a^b (f+g)=\int_a^b f+\int_a^b g \qquad \int_a^b cf = c\int_a^b f$$

Theorem 3.69. Sigui f integrable en $[a,b]$ amb $f([a,b])\subset[c,d]$ i g contínua en $[c,d]$. Llavors $g\circ f$ és integrable en $[a,b]$.

Corollary 3.70. Siguin f i g integrables en $[a,b]$. Llavors fg és integrable en $[a,b]$.

Proposition 3.71. Siguin f,g integrables en $[a,b]$ amb $f(x)\le g(x)\ \forall x\in[a,b]$. Aleshores $\int_a^b f\le\int_a^b g$.

Corollary 3.72. Si f és integrable en $[a,b]$ amb $m\le f(x)\le M\ \forall x\in[a,b]$. Aleshores $m(b-a)\le\int_a^b f\le M(b-a)$. Si a més f és contínua, aleshores existeix $c\in[a,b]$ amb $\int_a^b f=f(c)(b-a)$.

Proposition 3.73. Si f és integrable en $[a,b]$, aleshores $|f|$ també ho és i

$$\left|\int_a^b f\right|\le\int_a^b |f|$$

Proposition 3.74. Sigui f una funció integrable en $[a,b]$ i g una funció definida en $[a,b]$ diferent de f en un nombre finit de punts. Aleshores g és integrable i

$$\int_a^b g=\int_a^b f$$

Proposition 3.75. Si $f:[a,c]\to\mathbb{R}$ i $b\in(a,c)$ aleshores f integrable en $[a,c]$ si i només si f és integrable en $[a,b]$ i en $[b,c]$. A més,

$$\int_a^c f=\int_a^b f+\int_b^c f$$

Theorem 3.76 (Teorema fonamental del càlcul). Si f és integrable en [a,b], aleshores

$$F(x)=\int_a^x f$$

és contínua en $[a,b]$. Si, a més, f és contínua en $c\in[a,b]$, aleshores F és derivable en c i $F'(c)=f(c)$.

Theorem 3.77. Sigui f integrable en $[a,b]$ i G una primitiva de f. Llavors $\int_a^b f=G(b)-G(a)$.

Corollary 3.78 (Integració per parts). Siguin f,g integrables en $[a,b]$ amb primitives F i G respectivament. Aleshores es compleix:

$$\int_a^b Fg=F(b)G(b)-F(a)G(a)-\int_a^b fG$$

Corollary 3.79 (Canvi de variable). Sigui $\varphi:[c,d]\to[a,b]$ de classe $\mathcal{C}^1$ i tal que $\varphi(c)=a$ i $\varphi(d)=b$. Sigui f contínua en $[a,b]$. Aleshores es compleix:

$$\int_a^b f=\int_c^d (f\circ\varphi)\varphi'$$

Definition 3.80. Una suma de Riemann de f, $S(f,P)$, associada a una partició P, és qualsevol nombre obtingut de la següent manera:

$$S(f,P)=\sum_{i=1}^{n} f(x_i)(t_i-t_{i-1})$$

on $x_i\in[t_{i-1},t_i]$.

Theorem 3.81. Sigui f contínua en $[a,b]$. Aleshores $\forall\varepsilon>0\ \exists\delta>0$ tal que si $P=\{t_0,\dots,t_n\}$ és una partició de $[a,b]$ amb $t_i-t_{i-1}<\delta$, llavors

$$\left|\int_a^b f-S(f,P)\right|<\varepsilon$$

per a tota suma de Riemann associada a P.

Corollary 3.82. Sigui f contínua en $[a,b]$ i sigui P_n una successió de particions de $[a,b]$ tal que $t_i-t_{i-1}<1/n$ sempre que t_i i t_{i-1} siguin punts consecutius de P_n. Aleshores per a tota elecció $S(f,P_n)$ de sumes de Riemann associades a les particions P_n es té que

$$\int_a^b=\lim_{n\to\infty} S(f,P_n)$$

Definition 3.83. Sigui $f : [a,b] \to \mathbb{R}$ i $P = \{t_0, \ldots, t_n\}$ una partició de $[a,b]$. Definim

$$l(f,P) := \sum_{i=1}^{n} \sqrt{(t_i - t_{i-1})^2 + (f(t_i) - f(t_{i-1}))^2}$$

Si el conjunt $\mathcal{L} = \{l(f,P) : P \text{ partició de } [a,b]\}$ està acotat superiorment, diem que la gràfica és rectificable i definim la seva longitud $l(f,[a,b]) := \sup \mathcal{L}$.

Proposition 3.84. Sigui f de classe $\mathcal{C}^1$ a $[a,b]$. Aleshores f és rectificable a $[a,b]$ i

$$l(f,[a,b]) = \int_a^b \sqrt{1 + (f')^2}$$

Proposition 3.85. Sigui $\varphi : [a,b] \to \mathbb{R}^2$ amb $\varphi(t) = (x(t), y(t))$. Suposem que les funcions $x(t)$, $y(t)$ són de classe $\mathcal{C}^1$ a $[a,b]$. Aleshores la corba φ és rectificable a $[a,b]$ i

$$l(f,[a,b]) = \int_a^b \sqrt{(x')^2 + (y')^2}$$

Proposition 3.86. Sigui $f : [a,b] \to \mathbb{R}$ acotada i integrable. Aleshores el volum de revolució de f respecte l'eix horitzontal és

$$V_x = \pi \int_a^b f^2$$

Proposition 3.87. Sigui $f : [a,b] \to \mathbb{R}$ contínua. Aleshores el volum de revolució de f respecte l'eix vertical és

$$V_y = 2\pi \int_a^b x f(x) dx$$

Proposition 3.88. Sigui $f : [a,b] \to \mathbb{R}_+$ de classe $\mathcal{C}^1$. Aleshores la superfície de revolució del gràfic de f al voltant de l'eix horitzontal és

$$S_x = 2\pi \int_a^b f(x)\sqrt{1 + (f'(x))^2} dx$$

Proposition 3.89. Si $a > 0$ i $f : [a,b] \to \mathbb{R}$. La superfície de revolució del gràfic de f al voltant de l'eix vertical és

$$S_y = 2\pi \int_a^b x\sqrt{1 + (f'(x))^2} dx$$

Proposition 3.90. El centre de masses (x_0, y_0) d'un sòlid molt prim de secció i densitat uniforme és:

$$x_0 = \frac{\int_a^b x\sqrt{1 + (f'(x))^2} dx}{\int_a^b \sqrt{1 + (f'(x))^2} dx}, y_0 = \frac{\int_a^b f(x)\sqrt{1 + (f'(x))^2} dx}{\int_a^b \sqrt{1 + (f'(x))^2} dx}$$

El moment d'inèrcia d'un sòlid és:

$$I_x = \rho\sigma \int_a^b x^2 f(x) dx \qquad I_y = \rho\sigma \int_a^b x^2 \sqrt{1 + (f'(x))^2} dx$$

on ρ és la densitat del sòlid i σ el gruix, suposat constant, del sòlid.

Chapter 2

Second year

2.1 Algebraic structures

2.1.1 | Groups

Groups and subgroups

Definition 1.1 (Group). A group is a non-empty set G together with a binary operation

$$\begin{aligned} \cdot : G \times G &\longrightarrow G \\ (g_1, g_2) &\longmapsto g_1 \cdot g_2 \end{aligned}$$

satisfying the following properties:

1. Associativity:
$$(g_1 \cdot g_2) \cdot g_3 = g_1 \cdot (g_2 \cdot g_3) \quad \forall g_1, g_2, g_3 \in G.$$
2. Identity element:
$$\exists e \in G : e \cdot g = g \cdot e = g \quad \forall g \in G^{1}.$$
3. Inverse element:
$$\forall g \in G, \exists h \in G : g \cdot h = h \cdot g = e.$$
We denote h by g^{-1}.

If, moreover, we have $g_1 \cdot g_2 = g_2 \cdot g_1 \ \forall g_1, g_2 \in G$, we say the group $(G, \cdot)$ is *commutative* or *abelian*[2].

Lemma 1.2. Let $(G, \cdot)$ be a group. Then,

1. The identity element is unique.
2. Given an element $g \in G$, $\exists! h \in G$ such that $g \cdot h = h \cdot g = e$.
3. Given $g, h \in G$ such that $g \cdot h = e$, we have $h = g^{-1}$.

Definition 1.3 (Subgroup). Let $(G, \cdot)$ be a group and H be a subset of G. $(H, \cdot)$ is called a *subgroup* of $(G, \cdot)$[3] if satisfies:

1. If $h_1, h_2 \in H$, then $h_1 \cdot h_2 \in H$.
2. $e \in H$.
3. If $h \in H$, then $h^{-1} \in H$.

Proposition 1.4. Let $(G, \cdot)$ be a group and $H \neq \emptyset$ be a subset of G. Then

$$(H, \cdot) \text{ is a subgroup } \iff h_1 \cdot h_2^{-1} \in H \quad \forall h_1, h_2 \in H.$$

Proposition 1.5. If $(H, +)$ is a subgroup of $(\mathbb{Z}, +)$, then $\exists n \in \mathbb{Z}$ such that $H = n\mathbb{Z} = \{nk : k \in \mathbb{Z}\}$.

Proposition 1.6. Let $(G_i, *_i)$, for $i = 1, \dots, n$, be groups. Then the product

$$(G_1, *_1) \times \cdots \times (G_n, *_n)$$

induces a group with the operation $\cdot$ defined as

$$(g_1, \dots, g_n) \cdot (g_1', \dots, g_n') = (g_1 *_1 g_1', \dots, g_n *_n g_n'),$$

where $g_i, g_i' \in G_i$.

Definition 1.7. The *order of a group* $(G, \cdot)$ is the number of elements in its set, that is, $|G|$.

Lemma 1.8. Let $(G, \cdot)$ be a group and $\{(H_i, \cdot) : i \in I\}$ be a set of subgroups of $(G, \cdot)$. Then if

$$H = \bigcap_{i \in I} H_i,$$

we have that $(H, \cdot)$ is also a subgroup of $(G, \cdot)$.

Definition 1.9. Let $(G, \cdot)$ be a group and $X \subseteq G$ be a subset of G. The *subgroup of* $(G, \cdot)$ *generated by* X, $(\langle X \rangle, \cdot)$, is the smallest subgroup of $(G, \cdot)$ containing X, that is,

$$\langle X \rangle = \bigcap_{X \subseteq H \leq G} H.$$

Definition 1.10. Let $(G, *)$ be a group, $g \in G$ and $n \in \mathbb{Z}$. We define g^n as:

$$g^n = \begin{cases} g *^{(n)} \cdots * g & \text{if} \quad n > 0 \\ 1 & \text{if} \quad n = 0 \\ (g^{-1}) *^{(|n|)} \cdots * (g^{-1}) & \text{if} \quad n < 0 \end{cases}$$

Lemma 1.11. Let $(G, \cdot)$ be a group and $g \in G$. Then for all $n, m \in \mathbb{Z}$ we have

1. $g^n \cdot g^m = g^{n+m} = g^m \cdot g^n$.
2. $(g^n)^m = g^{nm} = (g^m)^n$.

Proposition 1.12. Let $(G, *)$ be a group and $X \subseteq G$ be a subset of G. Then:

$$\langle X \rangle = \{e\} \cup \{g_1^{\alpha_1} * \cdots * g_n^{\alpha_n} : n \in \mathbb{N}, \alpha_i \in \mathbb{Z}, g_i \in X\}.$$

Corollary 1.13. Let $(G, \cdot)$ be a group and $g \in G$. Then,

$$\langle g \rangle = \{g^i : i \in \mathbb{Z}\}.$$

Definition 1.14. Let $(G, \cdot)$ be a group and $g \in G$. A subgroup $(\langle g \rangle, \cdot)$ of $(G, \cdot)$ generated by a single element g is called a *cyclic group*.

Definition 1.15. Let $(G, \cdot)$ be a group and $g \in G$. The *order of* g is $\text{ord}(g) := |\langle g \rangle|$.

[1] From now on, we will denote e or e_G the identity element of the group $(G, \cdot)$.
[2] Sometimes to simplify the notation and if the context is clear, we will refer to G directly as the group as well as the set.
[3] Sometimes we will denote that $(H, \cdot)$ is a subgroup of $(G, \cdot)$ as $H \leq G$.

Proposition 1.16. Let $(G,\cdot)$ be a group and $g \in G$. Then,

$$\operatorname{ord}(g) = \min\{i \in \mathbb{N} : g^i = e\}.$$

If no such i exists, we say $\operatorname{ord}(g) = \infty$.

Corollary 1.17. Let $n \in \mathbb{N}$, $n > 1$, and $\bar{a} \in \mathbb{Z}/n\mathbb{Z}$. Then:

$$\operatorname{ord}(\bar{a}) = \frac{n}{\gcd(a,n)}.$$

Lemma 1.18. Let $(G,\cdot)$ be a group, $g \in G$ and $\operatorname{ord}(g) = n$. Then:

1. $g^m = e \iff n \mid m$.
2. $g^m = g^{m'} \iff m = m' \mod n$.
3. If $0 \le i \le n$, then $g^{-i} = (g^i)^{-1} = g^{n-i}$.

Corollary 1.19. Let $(G_i, *_i)$, for $i = 1,\dots,n$, be groups. For $i = 1,\dots,n$, let $g_i \in G_i$ and consider the element $g = (g_1,\dots,g_n) \in (G_1,*_1) \times \cdots \times (G_n,*_n)$. Then:

$$\operatorname{ord}(g) = \operatorname{lcm}(\operatorname{ord}(g_1),\dots,\operatorname{ord}(g_n)).$$

Group morphisms

Definition 1.20 (Group morphism). Let $(G,*)$, $(H,\cdot)$ be two groups. A *group morphism from* $(G,*)$ *to* $(H,\cdot)$ is a function $\phi : G \to H$ such that

$$\phi(g_1 * g_2) = \phi(g_1)\cdot\phi(g_2) \quad \forall g_1, g_2 \in G.$$

Lemma 1.21. Let $\phi : G_1 \to G_2$ be a morphism between $(G_1,*)$ and $(G_2,\cdot)$. Then,

1. $\phi(e_1) = e_2$.
2. $\phi(g^{-1}) = \phi(g)^{-1} \quad \forall g \in G_1$.
3. $\phi(g^n) = \phi(g)^n \quad \forall g \in G_1$ and $\forall n \in \mathbb{Z}$.

Definition 1.22. We say a subgroup $(H,\cdot)$ of a group $(G,\cdot)$ is *normal*, $H \lhd G$, if and only if $\forall h \in H$ and $\forall g \in G$, we have $g \cdot h \cdot g^{-1} \in H$.

Definition 1.23. Let $(G_1,*)$, $(G_2,\cdot)$ be two groups and $\phi : G_1 \to G_2$ be a group morphism. The *kernel of* ϕ is

$$\ker\phi = \{g \in G_1 : \phi(g) = e_2\}.$$

The *image of* ϕ is

$$\operatorname{im}\phi = \{h \in G_2 : \phi(g) = h \text{ for some } g \in G_1\}.$$

Proposition 1.24. Let $(G_1,*)$, $(G_2,\cdot)$ be two groups and $\phi : G_1 \to G_2$ be a group morphism. Then:

1. $(\ker\phi, *)$ is a normal subgroup of $(G_1,*)$ and $(\operatorname{im}\phi,\cdot)$ is a subgroup of $(G_2,\cdot)$.
2. Let $g, g' \in G_1$. The following statements are equivalent:
   i) $\phi(g) = \phi(g')$.
   ii) $g * g'^{-1} \in \ker\phi$.
   iii) $g'^{-1} * g \in \ker\phi$.
3. ϕ is injective if and only if $\ker\phi = \{e_1\}$.
4. ϕ is surjective if and only if $\operatorname{im}\phi = G_2$.

Definition 1.25. Let $(G,*)$, $(H,\cdot)$ be two groups. An *isomorphism between* $(G,*)$ *and* $(H,\cdot)$ is a bijective morphism between these groups.

Definition 1.26. Two groups $(G,*)$, $(H,\cdot)$ are *isomorphic*, $G \cong H$, if there exists an isomorphism $\phi : G \to H$.

Proposition 1.27. Let $(G_1,\cdot_1)$, $(G_2,\cdot_2)$, $(G_3,\cdot_3)$ be three groups and $\phi : G_1 \to G_2$, $\psi : G_2 \to G_3$ be two group morphisms. Then the composition $\psi \circ \phi$ is also a group morphism.

Proposition 1.28. Let $(G_1,*)$, $(G_2,\cdot)$ be groups and let $\phi : G_1 \to G_2$ be an isomorphism. Then $\phi^{-1} : G_2 \to G_1$ is also an isomorphism.

Theorem 1.29 (Classification of cyclic groups). Let $(G,\cdot)$ be a group and $g \in G$ be an element such that $\langle g\rangle = G$.

- If $|G| = \infty$, then $G \cong \mathbb{Z}$. We can define the isomorphism as follows:
$$\begin{aligned}\phi : \mathbb{Z} &\longrightarrow G\\ k &\longmapsto g^k\end{aligned}$$
- If $|G| = n$, then $G \cong \mathbb{Z}/n\mathbb{Z}$. We can define the isomorphism as follows:
$$\begin{aligned}\phi : \mathbb{Z}/n\mathbb{Z} &\longrightarrow G\\ \bar{k} &\longmapsto g^k\end{aligned}$$

Corollary 1.30. Let $(G,\cdot)$ be a group and $g \in G$ be such that $\langle g\rangle = G$. Then all subgroups of G are cyclic. Moreover:

- If $|G| = \infty$, subgroups of $(G,\cdot)$ are of the form $\langle g^n\rangle$, $n \in \mathbb{N} \cup \{0\}$.
- If $|G| = n$, then there is a unique subgroup $(H,\cdot)$ of $(G,\cdot)$ for every divisor $d > 0$ of n. In fact, if $n = dq$, then $H = \langle g^q\rangle$ and $|H| = d$.

Definition 1.31. Let X be a set. We define the *symmetric group* $(S(X),\circ)$ as

$$S(X) = \{f : X \to X : f \text{ is bijective}\}^4.$$

[4] Observe that if $X = \{1,\dots,n\}$, then $S(X) = S_n$.

Definition 1.32. Let $(G,\cdot)$ be a group. We define the functions:

$$\begin{aligned} \ell_g : G &\longrightarrow G & r_g : G &\longrightarrow S(G) \\ x &\longmapsto g\cdot x & x &\longmapsto x\cdot g \end{aligned}$$

Lemma 1.33. Let $(G,\cdot)$ be a group. The functions ℓ_g, r_g are bijective and its inverses are $\ell_{g^{-1}}$, $r_{g^{-1}}$, respectively.

Proposition 1.34. Let $(G,\cdot)$ be a group. We define the functions:

$$\begin{aligned} \phi : G &\longrightarrow S(G) & \psi : G &\longrightarrow G \\ g &\longmapsto \ell_g & g &\longmapsto r_{g^{-1}} \end{aligned}$$

Then, ϕ and ψ are injective group morphisms.

Theorem 1.35 (Cayley's theorem). Let $(G,\cdot)$ be a group. Then there is an injective morphism

$$\phi : G \longrightarrow S(G)$$

Corollary 1.36. If $(G,\cdot)$ is a group with $|G| = n$, then $(G,\cdot)$ is isomorphic to a subgroup of $(S_n,\circ)$.

Cosets

Definition 1.37. Let $(G,\cdot)$ be a finite group, $(H,\cdot)$ be a subgroup of $(G,\cdot)$ and $g_1, g_2 \in G$.

- We say $g_1 \sim g_2 \iff g_1\cdot g_2^{-1} \in H$.
- We say $g_1 \approx g_2 \iff g_2^{-1}\cdot g_1 \in H$.

Lemma 1.38. Let $(G,\cdot)$ be a finite group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. Then:

1. $\sim$ and $\approx$ are equivalence relations.
2. If $g \in G$, then

$$\begin{aligned} [g]_\sim &= H\cdot g = \{h\cdot g : h \in H\}, \\ [g]_\approx &= g\cdot H = \{g\cdot h' : h' \in H\}. \end{aligned}$$

Usually we say that $H\cdot g$ are the *right cosets in* G and $g\cdot H$, the *left cosets in* G.

Definition 1.39. Let $(G,\cdot)$ be a finite group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. We define the *set of left cosets* and the *set of right cosets* as follows:

$$G\Big/_\sim = \{H\cdot g : g \in G\}, \quad G\Big/_\approx = \{g\cdot H : g \in G\}.$$

Proposition 1.40. Let $(G,\cdot)$ be a group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. The following statements are equivalent:

1. $H \lhd G$.
2. $\forall g \in G,\ g\cdot H = H\cdot g$.

Theorem 1.41 (Lagrange's theorem). Let $(G,\cdot)$ be a finite group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. Then $|H| \mid |G|$.

Definition 1.42. Let $(G,\cdot)$ be a finite group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. We define the *index of* $(H,\cdot)$ *in* $(G,\cdot)$ as

$$[G:H] := \frac{|G|}{|H|}.$$

Corollary 1.43. Let $(G,\cdot)$ be a finite group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. Then:

$$[G:H] = \left|G\Big/_\sim\right| = \left|G\Big/_\approx\right|.$$

Corollary 1.44. Let $(G,\cdot)$ be a finite group.

1. If $g \in G$, then $\operatorname{ord}(g) \mid |G|$.
2. If $|G|$ is prime, then $(G,\cdot)$ is cyclic.
3. If $(H,\cdot)$ and $(K,\cdot)$ are subgroups of $(G,\cdot)$ and $\gcd(|H|,|K|) = 1$, then $H \cap K = \{e\}$.

Definition 1.45 (Quotient group). Let $(G,\cdot)$ be a finite group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$ such that $H \lhd G$. We define the *quotient group* $(G/H,*)$ as

$$G\Big/_H = G\Big/_\sim = G\Big/_\approx$$

and

$$\begin{aligned} * : G\Big/_H \times G\Big/_H &\longrightarrow G\Big/_H \\ (g_1\cdot H, g_2\cdot H) &\longmapsto (g_1\cdot g_2)\cdot H \end{aligned}$$

Lemma 1.46. Let $(G,\cdot)$ be a finite group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$ such that $H \lhd G$. The projection

$$\begin{aligned} \pi : G &\longrightarrow G\Big/_H \\ g &\longmapsto [g] = g\cdot H \end{aligned}$$

is a group morphism.

Isomorphism theorems

Theorem 1.47 (First isomorphism theorem). Let $(G_1,*)$, $(G_2,\cdot)$ be groups, $\phi : G_1 \to G_2$ be a group morphism and $(H,*)$ be a subgroup of $(G_1,*)$ such that $H \lhd G_1$. If $(H,*)$ is a subgroup of $(\ker\phi,*)$, then there exists a unique group morphism $\psi : G_1/H \to G_2$ such that the diagram of figure 2.1 is commutative, that is, $\phi = \psi\circ\pi$.

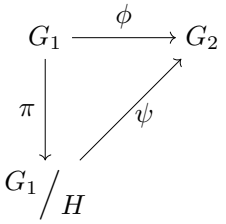

Figure 2.1

In particular, if $H = \ker\phi$, then ψ is injective and therefore there is an isomorphism $\psi : G_1/\ker\phi \to \operatorname{im}\phi$.

Theorem 1.48. Let

$$\begin{aligned}\phi : \mathbb{Z} &\longrightarrow {}^{\mathbb{Z}}\big/_{n\mathbb{Z}} \times {}^{\mathbb{Z}}\big/_{m\mathbb{Z}} \\ 1 &\longmapsto (\bar{1}, \bar{1})\end{aligned}$$

be a group morphism. Then, ϕ induces a morphism $\psi : \mathbb{Z}/nm\mathbb{Z} \to \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$. Moreover, ψ is injective if and only if $\gcd(n, m) = 1$ and in this case ψ is an isomorphism.

Corollary 1.49. Let $n, m \in \mathbb{Z}$ be two coprime integers and $a, b \in \mathbb{Z}$. The system of congruences

$$\begin{cases} x \equiv a \mod n \\ x \equiv b \mod m \end{cases}$$

has solutions and these are of the form $x \equiv c \mod nm$, where $c \equiv a \mod n$ and $c \equiv b \mod m$.

Definition 1.50. Let $(G, \cdot)$ be a group and $(H, \cdot)$, $(K, \cdot)$ be subgroups of $(G, \cdot)$. We define the *products of group subsets K, H* as the sets

$$\begin{aligned} H \cdot K &= \{h \cdot k : h \in H, k \in K\}, \\ K \cdot H &= \{k \cdot h : k \in K, h \in H\}. \end{aligned}$$

Proposition 1.51. Let $(G, \cdot)$ be a group and $(H, \cdot)$, $(K, \cdot)$ be subgroups of $(G, \cdot)$ such that $H \lhd G$. Then, $(H \cdot K, \cdot)$ is a subgroup of $(G, \cdot)$ and $H \cdot K = K \cdot H$.

Proposition 1.52. Let $(G, \cdot)$ be a group and $(H, \cdot)$, $(K, \cdot)$ be subgroups of $(G, \cdot)$ such that $H \cap K = \{e\}$. If $H, K \lhd G$, then the map

$$\begin{aligned}\phi : H \times K &\longrightarrow H \cdot K \\ (h, k) &\longmapsto h \cdot k\end{aligned}$$

is an isomorphism. In particular, $\forall h \in H$ and $\forall k \in K$, $h \cdot k = k \cdot h$.

Theorem 1.53 (Second isomorphism theorem). Let $(G, \cdot)$ be a group and $(H, \cdot)$, $(K, \cdot)$ be subgroups of $(G, \cdot)$ such that $H \lhd G$. Then $H \cap K \lhd K$ and

$${}^{K}\big/_{H \cap K} \cong {}^{H \cdot K}\big/_{H}.$$

Corollary 1.54. Let $(G, \cdot)$ be a group and $(H, \cdot)$, $(K, \cdot)$ be subgroups of $(G, \cdot)$. Then,

$$|H||K| = |H \cap K||H \cdot K|.$$

Lemma 1.55. Let $(G, \cdot)$ be a group and $(H, \cdot)$, $(K, \cdot)$ be subgroups of $(G, \cdot)$ such that $H \lhd G$ and $H \subseteq K$. Then $H \lhd K$, $(K/H, *)$ is a subgroup of $(G/H, *)$ and moreover

$${}^{K}\big/_{H} \lhd {}^{G}\big/_{H} \iff K \lhd G.$$

Theorem 1.56 (Correspondence theorem). Let $(G, \cdot)$ be a group and $(H, \cdot)$ be a subgroup of $(G, \cdot)$ such that $H \lhd G$. Then there is a bijection ϕ from the set $\mathcal{G}$ of all subgroups $(K, \cdot)$ of $(G, \cdot)$ such that $H \subseteq K$ onto the set $\mathcal{H}$ of all subgroups $(K/H, *)$ of $(G/H, *)$. More precisely, the bijection is:

$$\begin{aligned}\phi : \mathcal{G} &\longrightarrow \mathcal{H} \\ K &\longmapsto K/H\end{aligned}$$

Theorem 1.57 (Third isomorphism theorem). Let $(G, \cdot)$ be a group and $(H, \cdot)$, $(K, \cdot)$ be subgroups of $(G, \cdot)$ such that $H, K \lhd G$ and $H \subseteq K$. Then $K/H \lhd G/H$ and

$$\left({}^{G}\big/_{H}\right)\Big/\left({}^{K}\big/_{H}\right) \cong {}^{G}\big/_{K}.$$

Group actions

Definition 1.58. Let X be a set and $(G, \cdot)$ be a group. A *(left) group action of $(G, \cdot)$ on X* is a function

$$\begin{aligned} * : G \times X &\longrightarrow X \\ (g, x) &\longmapsto g * x \end{aligned}$$

satisfying the following properties:

1. $e * x = x$, $\forall x \in X$.
2. $(g_1 \cdot g_2) * x = g_1 * (g_2 * x)$, $\forall x \in X$ and $\forall g_1, g_2 \in G$.

A set X together with an action $*$ of $(G, \cdot)$ is usually called a *(left) G-set*.

Lemma 1.59. Let $(G, \cdot)$ be a group and X be a G-set. For all $g \in G$ the function

$$\begin{aligned} \ell_g : X &\longrightarrow X \\ x &\longmapsto g * x \end{aligned}$$

is bijective an its inverse is $\ell_{g^{-1}}$.

Definition 1.60. Let $(G, \cdot)$ be a group and X be a G-set. For all $x, y \in X$, we say $x \backsim y \iff \exists g \in G : y = g * x$.

Lemma 1.61. The relation $\backsim$ is an equivalence relation.

Definition 1.62. Let $(G, \cdot)$ be a group and X be a G-set. If $x \in X$, we define the *orbit of x* as:

$$\mathcal{O}_x = [x]_{\backsim} = \{g * x : g \in G\}.$$

Definition 1.63. Let $(G, \cdot)$ be a group and X be a G-set. For $x \in X$, we define the *stabilizer of $(G, \cdot)$ with respect to x* as the set:

$$G_x = \{g \in G : g * x = x\}.$$

Proposition 1.64. Let $(G, \cdot)$ be a group and X be a G-set. For all $x \in X$, $(G_x, \cdot)$ is a subgroup of $(G, \cdot)$.

Theorem 1.65 (Orbit-stabilizer theorem). Let $(G,\cdot)$ be a group, X be a G-set and $x \in X$. The surjective map

$$\begin{aligned}\phi : G &\longrightarrow \mathcal{O}_x \\ g &\longmapsto g * x\end{aligned}$$

induces a bijective map $\psi : G/\approx \to \mathcal{O}_x$, where $\approx$ is the equivalence relation $g_1 \approx g_2 \iff g_2^{-1}\cdot g_1 \in G_x$ $\forall g_1, g_2 \in G$[5]. In particular, if G is finite,

$$|\mathcal{O}_x| = |[G : G_x]|.$$

Corollary 1.66 (Orbits formula). Let $(G,\cdot)$ be a finite group and X be a finite G-set. If $x_1,\dots,x_m$ are the elements of X and $|\mathcal{O}_{x_i}| = 1$ for $i = 1,\dots,r$, then:

$$|X| = r + \sum_{i=r+1}^{m} |\mathcal{O}_{x_i}| = r + \sum_{i=r+1}^{m} |[G : G_{x_i}]|. \tag{2.1}$$

Applications of orbits formula

Theorem 1.67 (Cauchy's theorem). Let $(G,\cdot)$ be a finite group of order n and p be a prime number. If $p \mid n$, then $(G,\cdot)$ has an element of order p.

Corollary 1.68. Let p be an odd prime number. Then groups of order $2p$ are isomorphic to $(\mathbb{Z}/2p\mathbb{Z},+)$ or $(\mathrm{D}_{2p},\circ)$[6].

Proposition 1.69. Let $(G,\cdot)$ be a group. The map

$$\begin{aligned}G \times G &\longrightarrow G \\ (g,x) &\longmapsto g\cdot x\cdot g^{-1}\end{aligned}$$

is an action of $(G,\cdot)$ over itself. It is called the *conjugation action*.

Definition 1.70 (Center of a group). Let $(G,\cdot)$ be a group. We define the *center of* $(G,\cdot)$ as

$$Z(G) = \{z \in G : z\cdot g = g\cdot z \ \forall g \in G\}$$[7].

Proposition 1.71. Let p be a prime number and $(G,\cdot)$ be a finite group of order p^n for some $n \geq 1$. Then, $|Z(G)| > 1$.

Lemma 1.72. Let $(G,\cdot)$ be a group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. Consider the application

$$\begin{aligned}H \times {G}\big/_{\approx} &\longrightarrow {G}\big/_{\approx} \\ (h, g\cdot H) &\longmapsto (h\cdot g)\cdot H\end{aligned}$$

This application defines an action of the subgroup $(H,\cdot)$ over the set $G/\approx$.

Definition 1.73. Let $(G,\cdot)$ be a group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. The *normalizer of* $(H,\cdot)$ *in* $(G,\cdot)$ is

$$N_G(H) = \{g \in G : g\cdot h\cdot g^{-1} \in H \ \forall h \in H\}.$$

Lemma 1.74. Let $(G,\cdot)$ be a group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. $(N_G(H),\cdot)$ is a subgroup of $(G,\cdot)$ containing H and, moreover, $H \lhd N_G(H)$.

Corollary 1.75. Let $(G,\cdot)$ be a finite group and $(H,\cdot)$ be a subgroup of $(G,\cdot)$. Then by orbits formula applied to action defined on lemma 1.72, we have:

$$[G : H] = [N_G(H) : H] + \sum_{|\mathcal{O}_x|>1} |\mathcal{O}_x|.$$

Proposition 1.76. Let $(G,\cdot)$ be a group of order $n \in \mathbb{N}$, p be a prime number such that $p \mid n$ and $(H,\cdot)$ be a subgroup of $(G,\cdot)$ of order p^i, $i \geq 1$. Suppose $p \mid [G : H]$. Then $p \mid [N_G(H) : H]$.

Sylow's theorems

Corollary 1.77. Let $(G,\cdot)$ be a group of order $n \in \mathbb{N}$, p be a prime number and $(H,\cdot)$ be a subgroup of $(G,\cdot)$ such that $|H| = p^i$, $i \geq 0$. Suppose $p \mid [G : H]$. Then, there is a subgroup $(H',\cdot)$ of $(G,\cdot)$ such that $(H,\cdot)$ is a subgroup of $(H',\cdot)$ and $|H'| = p^{i+1}$. Moreover, $H \lhd H'$ and $H'/H \cong \mathbb{Z}/p\mathbb{Z}$.

Theorem 1.78 (First Sylow theorem). Let $(G,\cdot)$ be a finite group and p be a prime number. Suppose $|G| = p^r m$, where $r \geq 0$ and $\gcd(p,m) = 1$. Then, there is a subgroup $(K,\cdot)$ of $(G,\cdot)$ of order p^r. Moreover there is a chain of subgroups $(H_i,\cdot)$ satisfying:

$$\{e\} = H_0 \lhd H_1 \lhd \cdots \lhd H_r = K,$$

such that $H_{i+1}/H_i \cong \mathbb{Z}/p\mathbb{Z}$ for $0 \leq i < r$.

Definition 1.79. Let p be a prime number. A group $(G,\cdot)$ is a *p-group* if $|G| = p^r$, for some $r \in \mathbb{N}$.

Definition 1.80. Let p be a prime number and $(G,\cdot)$ be a group. A *Sylow p-subgroup* is a p-subgroup of $(G,\cdot)$ of maximum order.

Definition 1.81. Let $(G,\cdot)$ be a finite group. We say $(G,\cdot)$ is *soluble* if there is a chain of subgroups $(H_i,\cdot)$ of $(G,\cdot)$ satisfying:

$$\{e\} = H_0 \lhd H_1 \lhd \cdots \lhd H_r = K,$$

and such that the subgroups $(H_{i+1}/H_i, *)$ are cyclic, $0 \leq i < r$.

[5] Note that the notation $\approx$ for the equivalence relation correspond with the one defined in definition 1.37.
[6] See the end of section 2.1.1.
[7] Note that, by orbits formula (2.1), if we consider the conjugation action we have:

$$|G| = |Z(G)| + \sum_{|\mathcal{O}_x|>1} |\mathcal{O}_x|$$

Theorem 1.82 (Second Sylow theorem). Let $(G,\cdot)$ be a finite group and p be a prime number. Suppose $|G| = p^r m$, where $r \geq 0$ and $\gcd(p,m) = 1$. Let $(K,\cdot)$ be a Sylow p-subgroup of $(G,\cdot)$. Then, if $(H,\cdot)$ is a subgroup of $(G,\cdot)$ of order p^i, $\exists g \in G$ such that $g \cdot H \cdot g^{-1} \subseteq K$. In particular two different Sylow p-subgroups $(K_1,\cdot)$ and $(K_2,\cdot)$ are conjugate, that is, there exists an element $g \in G$ such that $g \cdot K_1 \cdot g^{-1} = K_2$.

Theorem 1.83 (Third Sylow theorem). Let $(G,\cdot)$ be a finite group and p be a prime number. Suppose $|G| = p^r m$, where $r \geq 0$ and $\gcd(p,m) = 1$. Let $(K,\cdot)$ be a Sylow p-subgroup of $(G,\cdot)$ and n_p be the number of different Sylow p-subgroups of $(G,\cdot)$. Then, $n_p = [G : N_G(K)]$. Therefore, $n_p \mid m$ and $n_p \equiv 1 \mod p$.

Examples of groups

Let $n, p \in \mathbb{N}$ such that p is a prime number.

- $(\mathbb{Z},+)$, $(\mathbb{Z}/n\mathbb{Z},+)$, $(\mathbb{Q},+)$, $(\mathbb{R},+)$, $(\mathbb{C},+)$
- $((\mathbb{Z}/p\mathbb{Z})^*,\cdot)$, $(\mathbb{Q}^*,\cdot)$, $(\mathbb{R}^*,\cdot)$, $(\mathbb{C}^*,\cdot)$
- $(S_n,\circ)$
- $(A_n,\circ)$, where $A_n = \{\sigma \in S_n : \varepsilon(\sigma) = 1\}$. Note that $|A_n| = \frac{S_n}{2} = \frac{n!}{2}$.
- $(\mathrm{GL}_n(\mathbb{A}),\cdot)$, where $\mathrm{GL}_n(\mathbb{A}) = \{M \in \mathcal{M}_n(\mathbb{A}) : M \text{ is invertible}\}$ and $\mathbb{A} = \mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$.
- $(\mathrm{SL}_n(\mathbb{A}),\cdot)$, where $\mathrm{SL}_n(\mathbb{A}) = \{M \in \mathrm{GL}_n(\mathbb{A}) : \det M = 1\}$ and $\mathbb{A} = \mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ or $\mathbb{C}$.
- $(\mathrm{D}_{2n},\circ)$, where D_{2n} is the set of rotations and reflections that leave invariant the regular polygon of n vertices centered at origin. It can be seen that $\mathrm{D}_{2n} = \langle r, s : \mathrm{ord}(r) = 4, \mathrm{ord}(s) = 2, r \circ s = s \circ r^{-1}\rangle$. This group is called the *dihedral group*.
- $(\mathrm{Q}_8,\cdot)$, where $\mathrm{Q}_8 = \langle a, b : \mathrm{ord}(a) = \mathrm{ord}(b) = 4, b \cdot a = a^{-1} \cdot b\rangle$. This group is called the *quaternion group*.

2.1.2 | Rings

Rings, subrings and ring morphisms

Definition 1.84 (Ring). A *ring* is a set R equipped with two binary operations

$$+ : R \times R \longrightarrow R \qquad \cdot : R \times R \longrightarrow R$$
$$(r_1, r_2) \longmapsto r_1 + r_2 \qquad (r_1, r_2) \longmapsto r_1 \cdot r_2$$

satisfying the following properties:

1. $(R,+)$ is an abelian group.
2. $(R,\cdot)$ satisfies[8]:
   i) Associativity:
   $$(r_1 \cdot r_2) \cdot r_3 = r_1 \cdot (r_2 \cdot r_3) \quad \forall r_1, r_2, r_3 \in R.$$
   ii) Identity element[9]:
   $$\exists 1 \in R : 1 \cdot r = r \cdot 1 = r \quad \forall r \in R.$$
   iii) Commutativity:
   $$r_1 \cdot r_2 = r_2 \cdot r_1 \quad \forall r_1, r_2 \in R.$$
3. Multiplication is distributive with respect to addition:
   $$(r_1 + r_2) \cdot r_3 = r_1 \cdot r_3 + r_2 \cdot r_3 \quad \forall r_1, r_2, r_3 \in R.$$

In this context we say $(R,+,\cdot)$ is a ring.

Definition 1.85 (Field). Let $(R,+,\cdot)$ be a ring. If every nonzero element of R has a multiplicative inverse (that is, $(R,\cdot)$ is an abelian group), we say R is a *field*.

Lemma 1.86. Let $(R,+,\cdot)$ be a ring. Then,

1. The multiplicative identity element is unique.
2. $\forall r \in R$, $0 \cdot r = 0$.
3. $\forall r \in R$, $(-1) \cdot r = -r$, where -1 is the additive inverse of 1.
4. $\forall r, s \in R$, $(-r) \cdot s = -(r \cdot s)$ and $(-r) \cdot (-s) = r \cdot s$.

Definition 1.87 (Subring). Let $(R,+,\cdot)$ be a ring and $S \subseteq R$ be a subset of R. $(S,+,\cdot)$ is called a *subring* of $(R,+,\cdot)$ if satisfies:

1. $(S,+)$ is a subgroup of $(R,+)$.
2. $\forall s_1, s_2 \in S$, $s_1 \cdot s_2 \in S$.
3. $1 \in S$.

Definition 1.88 (Ring morphism). Let $(R,+,\cdot)$, $(S,\oplus,\odot)$ be two rings. A *ring morphism from* $(R,+,\cdot)$ *to* $(S,\oplus,\odot)$ is a function $\phi : R \to S$ such that:

1. ϕ is a group morphism between groups $(R,+)$ and $(S,\oplus)$.
2. $\phi(r_1 * r_2) = \phi(r_1) \odot \phi(r_2) \quad \forall r_1, r_2 \in R$.
3. $\phi(1_R) = 1_S$.

Lemma 1.89. Let $(R,+,\cdot)$, $(S,\oplus,\odot)$ be two rings and $\phi : R \to S$ be a ring morphism. Then knowing that $\ker\phi = \{r \in R : f(r) = 0\}$, then:

1. $(\ker\phi,+)$ is a subgroup of $(R,+)$.
2. $\forall k \in \ker\phi$ and $\forall r \in R$, $k \cdot r \in \ker\phi$.

Proposition 1.90. Let $(R,+,\cdot)$, $(S,\oplus,\odot)$ be two rings and $\phi : R \to S$ be a ring morphism. Then:

1. $f(0) = 0$.

[8] Some definitions state that the commutative property is not necessary to define a ring. However, in these notes we will take the definition given.

[9] It is common to denote the additive identity element as 0 and the multiplicative identity element as 1.

2. $\forall r \in R,\ f(-r) = -f(r)$.
3. If $r \in R$ has a multiplicative inverse, then $f(r)$ so it has and, moreover, $f(r^{-1}) = f(r)^{-1}$.

Proposition 1.91. Let $(R_1, +_1, \cdot_1)$, $(R_1, +_1, \cdot_1)$ and $(R_1, +_1, \cdot_1)$ be rings and $\phi : R_1 \to R_2$, $\psi : R_2 \to R_3$ be two ring morphisms. Then the composition $\psi \circ \phi$ is also a ring morphism.

Proposition 1.92. Let $(R, +, \cdot)$, $(S, \oplus, \odot)$ be rings and let $\phi : R \to S$ be a bijective ring morphism. Then $\phi^{-1} : S \to R$ is also a bijective ring morphism.

Ideals

Definition 1.93. Let $(R, +, \cdot)$ be a ring. A subgroup $(I, +)$ of $(R, +)$ is an *ideal* if $\forall x \in I$ and $\forall r \in R$, $x \cdot r \in I$.

Lemma 1.94 (Principal ideal). Let $(R, +, \cdot)$ be a ring and $a \in R$. The set

$$(a) := aR = \{a \cdot r : r \in R\}$$

is an ideal of $(R, +, \cdot)$ and it is called the *principal ideal generated by* a.

Proposition 1.95. Let $(R, +, \cdot)$ be a ring. R is a field if and only if $(R, +, \cdot)$ has only two ideals: $\{0\}$ and R.

Definition 1.96. Let $(R, +, \cdot)$ be a ring. An element $r \in R$ is a *unit* if has a multiplicative invertible. The set of units in $(R, +, \cdot)$ is denoted by R^* and $(R^*, \cdot)$ is a group called *multiplicative group of* $(R, +, \cdot)$.

Lemma 1.97. Let $(R, +, \cdot)$, $(S, \oplus, \odot)$ be rings and $u \in R^*$. Then,

1. If $r \in R$, then $r \cdot R = r \cdot u \cdot R$.
2. $f : (R^*, \cdot) \to (S^*, \odot)$ is a group morphism.

Proposition 1.98. Let $(k, +, \cdot)$ be a field. Then ideals of $k[x]$ are all principal. Moreover if $I \neq \{0\}$ is an ideal of $k[x]$, exists a monic polynomial $p(x) \in k[x]$ such that $I = p(x) \cdot k[x]$.

Proposition 1.99. Let $(R, +, \cdot)$ be a ring and I, J be a ideals. Then the sets

$$I \cap J := \{x : x \in I,\ x \in J\},$$
$$I + J := \{x + y : x \in I,\ y \in J\},$$
$$I \cdot J := \{\sum_{i=1}^{n} x_i y_i : n \geq 0,\ x_i \in I,\ y_i \in J\},$$

are all ideals. In particular $I \cap J$ is the largest ideal contained in I and J and $I + J$ is the smallest ideal containing I and J.

Definition 1.100. A ring is *noetherian* if all its ideals are finitely generated.

Lemma 1.101. Let $(R, +, \cdot)$, $(S, \oplus, \odot)$ be two rings and $\phi : R \to S$ be a ring morphism. Then:

1. $\ker \phi$ is an ideal of $(R, +, \cdot)$.
2. $\operatorname{im} \phi$ is a subring of $(S, \oplus, \odot)$.

Theorem 1.102. Let $(R, +, \cdot)$, $(S, \oplus, \odot)$ be two rings and $\phi : R \to S$ be a ring morphism. Then there is a bijection ψ from the set $\mathcal{R}$ of all ideals I of $(R, +, \cdot)$ such that $\ker \phi \subseteq I$ onto the set $\mathcal{S}$ of all ideals of $\operatorname{im} \phi$. More precisely, the bijection is:

$$\begin{aligned} \psi : \mathcal{R} &\longrightarrow \mathcal{S} \\ I &\longmapsto \phi(I) \end{aligned}$$

Ideal quotient

Definition 1.103. Let $(R, +, \cdot)$ be a ring and I be an ideal of $(R, +, \cdot)$. For all $r_1, r_2 \in R$, we say $r_1 \sim r_2 \iff r_1 - r_2 \in I$. Since $(I, +)$ is a subgroup of $(R, +)$, $\sim$ is an equivalence relation and we denote $R \,/\, I := R \,/\, \sim$ the set of equivalence classes.

Proposition 1.104. Let $(R, +, \cdot)$ be a ring and I be an ideal of $(R, +, \cdot)$. Then $R \,/\, I$ is a ring with operation defined as:

- $\forall r_1, r_2 \in R,\ \overline{r_1} + \overline{r_2} = \overline{r_1 + r_2}$ being $\overline{0}$ the identity element with respect to this operation.
- $\forall r_1, r_2 \in R,\ \overline{r_1} \cdot \overline{r_2} = \overline{r_1 \cdot r_2}$ being $\overline{1}$ the identity element with respect to this operation.

Moreover the projection:

$$\begin{aligned} \pi : R &\longrightarrow R \Big/ I \\ r &\longmapsto \overline{r} \end{aligned}$$

is a surjective ring morphism with $\ker \pi = I$.

Corollary 1.105. Let $(R, +, \cdot)$ be a ring and I be an ideal of $(R, +, \cdot)$. Ideals of $R \,/\, I$ are of the form $J \,/\, I$ where J is an ideal of $(R, +, \cdot)$ containing I.

Isomorphism theorems

Theorem 1.106 (First isomorphism theorem). Let $(R, +, \cdot)$, $(S, \oplus, \odot)$ be two rings, $\phi : R \to S$ be a ring morphism and I be an ideal such that I is a subgroup of $(\ker \phi, +)$. Then there exists a unique ring morphism $\psi : R/I \to S$ such that the diagram of figure 2.2 is commutative, that is, $\phi = \psi \circ \pi$.

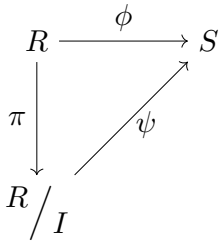

Figure 2.2

In particular, if $I = \ker \phi$, then ψ is injective and therefore there is an isomorphism $\psi : R/\ker \phi \to \operatorname{im} \phi$.

2.2 Discrete mathematics

2.2.1 | Generating functions and recurrence relations

Generating functions

Definition 2.1. Let (a_n) be a sequence of real numbers. We define its *ordinary generating function* as the following formal power series:

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots = \sum_{n=0}^{\infty} a_n x^n.$$

Proposition 2.2. Let $\sum_{n=0}^{\infty} a_n x^n, \sum_{n=0}^{\infty} b_n x^n$ be two formal power series. Then:

- $\sum_{n=0}^{\infty} a_n x^n + \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} (a_n + b_n)x^n.$
- $\lambda \sum_{n=0}^{\infty} a_n x^n = \sum_{n=0}^{\infty} \lambda a_n x^n.$
- $\left(\sum_{n=0}^{\infty} a_n x^n\right)\left(\sum_{n=0}^{\infty} b_n x^n\right) = \sum_{n=0}^{\infty} (a_0b_n + a_1b_{n-1} + \cdots + a_nb_0)x^n.$
- $\left(\sum_{n=0}^{\infty} a_n x^n\right)' = \sum_{n=1}^{\infty} n a_n x^{n-1}.$

Proposition 2.3 (Closed forms). We can write the following ordinary generating functions with their corresponding closed forms:

- $\sum_{n=0}^{N} x^n = \frac{1 - x^{N+1}}{1 - x}.$
- $\sum_{n=0}^{\infty} x^n = \frac{1}{1 - x}.$
- $\sum_{n=0}^{\infty} \binom{n + k - 1}{n} x^n = \left(\frac{1}{1 - x}\right)^k.$

Proposition 2.4. Suppose A and B are two finite disjoint sets. We set some restrictions for the non-ordered selection of elements of $A \cup B$. For every $n \geq 0$, let:

- a_n be the number of non-ordered selection of n elements of A satisfying the restrictions,
- b_n be the number of non-ordered selection of n elements of B satisfying the restrictions,
- c_n be the number of non-ordered selection of n elements of $A \cup B$ satisfying the restrictions.

And let $f(x)$, $g(x)$, $h(x)$ be the ordinary generating functions of (a_n), (b_n), (c_n), respectively. Then we have:

$$h(x) = f(x)g(x).$$

Definition 2.5. Let (a_n) be a sequence of real numbers. We define its *exponential generating function* as the following formal power series:

$$a_0 + a_1x + a_2\frac{x^2}{2!} + a_3\frac{x^3}{3!} + \cdots = \sum_{n=0}^{\infty} a_n \frac{x^n}{n!}.$$

Definition 2.6. Let (a_n) be a sequence of real numbers such that $a_i = 1\ \forall i$. Then its exponential generating function associated is the so called *exponential series*:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Proposition 2.7. The exponential series has the following properties:

1. $e^{x+y} = e^x e^y\ \forall x, y \in \mathbb{R}.$
2. $(e^x)^n = e^{nx}\ \forall x, n \in \mathbb{R}.$

Proposition 2.8. Suppose A and B are two finite disjoint sets. We set some restrictions for the ordered selection of elements of $A \cup B$. For every $n \geq 0$, let:

- a_n be the number of ordered selection of n elements of A satisfying the restrictions,
- b_n be the number of ordered selection of n elements of B satisfying the restrictions,
- c_n be the number of ordered selection of n elements of $A \cup B$ satisfying the restrictions.

And let $f(x)$, $g(x)$, $h(x)$ be the exponential generating functions of (a_n), (b_n), (c_n), respectively. Then we have:

$$h(x) = f(x)g(x).$$

Recurrence relations

Definition 2.9. Let (a_n) be a sequence of real numbers. A *recurrence relation of order k for* (a_n) is an expression that express a_n in terms of k consecutive terms of the sequence, $a_{n-1}, \ldots, a_{n-k}$, for $k \leq n$. We say a sequence is *recurrent* if it satisfies a recurrence relation or, equivalently, if it's a solution of the recurrence relation.

Definition 2.10. The *initial values* of a recurrence relation of order k are the values of the first k terms for which the recurrence relation is still not valid, that is, the values $a_0, a_1, \ldots, a_{k-1}$.

Lemma 2.11. The solution of a recurrence relation of order k with k initial conditions is unique.

Definition 2.12. A *linear recurrence relation of order k* is a recurrence relation that can be written as the form

$$a_n + c_1 a_{n-1} + \cdots c_k a_{n-k} = g(n)$$

where $c_1, \ldots c_k \in \mathbb{R}, c_k \neq 0$ and $g : \mathbb{N} \to \mathbb{N}$ is an arbitrary function.

Definition 2.13. We say a linear recurrence relation is *homogeneous* if $g(n) = 0$, that is, if it's of the form:

$$a_n + c_1 a_{n-1} + \cdots c_k a_{n-k} = 0, \quad \text{with } c_k \neq 0.$$

Proposition 2.14. The general solution to a recurrence relation

$$a_n + c_1 a_{n-1} + \cdots c_k a_{n-k} = g(n)$$

can be expressed as

$$(a_n^{\text{part}}) + (a_n^{\text{hom}}),$$

where (a_n^{part}) is a particular solution of the recurrence relation and (a_n^{hom}) is the general solution of its associated homogeneous recurrence relation.

Proposition 2.15. Given $c_1, \ldots, c_k \in \mathbb{R}$, the set of sequences that are solution of the homogeneous linear recurrence relation $a_n + c_1 a_{n-1} + \cdots + c_k a_{n-k} = 0$ form a real vector space.

Definition 2.16. Let $a_n + c_1 a_{n-1} + \cdots + c_k a_{n-k} = 0$ be a homogeneous linear recurrence relation of order k. The *characteristic polynomial* of the recurrence is:

$$x^k + c_1 x^{k-1} + \cdots + c_k = 0.$$

Proposition 2.17. Consider an homogeneous linear recurrence relation with characteristic polynomial

$$(x - r_1)(x - r_2) \cdots (x - r_k) = 0$$

where $r_1, \ldots, r_k \in \mathbb{C}$ are different complex numbers. Then the general term of the sequences that satisfy the recurrence relation is

$$a_n = \lambda_1 r_1^n + \cdots + \lambda_k r_k^n$$

for arbitrary numbers $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$.

2.2.2 | Graph theory

Definition 2.18. A *graph* G is an structure based on a set $V(G)$ of vertices and a set $E(G)$ of edges, which are non-ordered pairs of vertices.

Definition 2.19. Let G be a graph. The *order of* G is $n = |V(G)|$ and the *size of* G is $m = |E(G)|$.

Definition 2.20. Let G be a graph. Two vertices $a, b \in V(G)$ are said to be *adjacent* to one another if exists an edge $e \in E(G)$ that connects them. In this case we say the edge e is *incident* on vertices a and b.

Definition 2.21. An edge that connects a vertex with itself is called a *loop*.

Definition 2.22. Two or more edges incidents with the same vertices are called *multiple edges*.

Definition 2.23. A graph G is *finite* if $V(G)$ and $E(G)$ are finite.

Definition 2.24. A graph is *simple* if it has neither multiples edges nor loops.

Definition 2.25. A *complete graph* is a graph in which each pair of different vertices is joined by an edge. We denote by K_n the complete graph of order n.

Definition 2.26. Let G be a finite graph. The *degree of a vertex* is the number of edges that are incident to it. If $v \in V(G)$ we denote the degree of v by $\deg v$ or $\deg_G v$[10].

Lemma 2.27 (Handshaking lemma). For every graph G we have:

$$\sum_{v \in V(G)} \deg v = 2|E(G)|.$$

Corollary 2.28. In any graph, the number of odd-degree vertices is even.

Definition 2.29. Let G be a graph with $V(G) = \{v_1, \ldots, v_n\}$. The *degree sequence of* G is the decreasing sequence

$$(\deg v_{i_1}, \ldots, \deg v_{i_n}).$$

Definition 2.30. We say a graph G is *k-regular* if $\deg v = k \ \forall v \in V(G)$.

Definition 2.31. Let G be a graph. A graph F is an *induced subgraph of* G if $V(F) \subseteq V(G)$ and $E(F) \subseteq E(G)$.

Definition 2.32. A *walk of lenght k* in a graph G is a sequence of vertices $(u_1, \ldots, u_k)$ where $u_i u_{i+1} \in E(G)$ for $i = 1, \ldots, k-1$.

Definition 2.33. A walk in a graph is *closed* if it starts and ends in the same vertex.

Definition 2.34. A walk in a graph is a *trail* if all the edges of the walk are distinct.

Definition 2.35. A walk in a graph is a *path* if all the vertices (and therefore the edges) of the walk are distinct.

Definition 2.36. A closed walk in a graph is a *closed trail* if all the edges of the closed walk are distinct.

Definition 2.37. A closed path is called a *cycle*.

Proposition 2.38. Let G be a graph. Given $u, v \in V(G)$, there exists a walk between u and v if and only if there exists a path between u and v.

[10] Observe that with this definition every loop counts as two edges.

Definition 2.39. Let G be a graph. Given $u, v \in V(G)$, we say that u and v are connected if there is a path in G between u and v.

Proposition 2.40. The relation $u \sim v$ if and only if u and v are connected is an equivalence relation. The equivalent classes are the *connected components of* G.

Definition 2.41. A graph G is *connected* if $\forall u, v \in V(G)$, u and v are connected.

Definition 2.42. A graph G is *bipartite* if $V(G) = X \sqcup Y$ and $\forall e \in E(G)$ we have $e = xy$ with $x \in X$ and $y \in Y$.

Definition 2.43. Let G be a graph such that $E(G) \neq \emptyset$. Take an edge $e \in E(G)$. We denote by $G - e$ the induced graph of G such that

$$V(G - e) = V(G) \qquad E(G - e) = E(G) \setminus \{e\}.$$

Definition 2.44. Given a connected graph G, we say that $e \in E(G)$ is a *bridge of* G if $G - e$ is non-connected.

Proposition 2.45. Let G be a connected graph. $e \in E(G)$ is a bridge if and only if e doesn't belong to any cycle of G.

Definition 2.46. Let G be a connected graph. An *Eulerian trail* in G is a trail that contain all the edges of G. An *Eulerian circuit* in G is a closed eulerian trail. G is called *Eulerian* if it admits an eulerian circuit.

Theorem 2.47 (Euler theorem). Let G be a connected graph. G is Eulerian $\iff \deg v = 2k \ \forall v \in V(G)$, $k \in \mathbb{N}$.

Definition 2.48. Let G be a graph of order n with $V(G) = \{v_1, \dots, v_n\}$. We define the *adjacency matrix of* G, $A(G) \in \mathcal{M}_n(\mathbb{R})$, as a_{ij} to be the number of edges incident with v_i and v_j.

Proposition 2.49. Let G be a graph of order n with $V(G) = \{v_1, \dots, v_n\}$ and let $A(G) = (a_{ij})$ be the adjacency matrix of G. Then:

1. $A(G)$ is symmetric.
2. $\displaystyle\sum_{j=1}^{n} a_{jk} = \sum_{j=1}^{n} a_{kj} = \deg v_k, \quad k = 1, \dots, n.$
3. For $k \in \mathbb{N}$, consider $A(G)^k = (b_{ij}^k)$. Then b_{ij}^k is equal to the number of walks of length k between vertices v_i and v_j.

Definition 2.50. A *tree* is an acyclic connected graph, that is, a connected graph that has no cycles.

Definition 2.51. Let T be a tree. A *leave* of T is a vertex of degree 1.

Definition 2.52. Let G be a graph. A *generator tree* is a induced subgraph T of G such that $|V(G)| = |V(T)|$ and T is a tree.

Proposition 2.53. Let G be a graph such that $|V(G)| = n \geq 2$. The following are equivalent:

1. G is a tree.
2. G is connected and every edge of G is a bridge.
3. G is connected and $|E(G)| = n - 1$.
4. G is acyclic and $|E(G)| = n - 1$.
5. For $v_i, v_j \in V(G)$, $i \neq j$, there exists a unique path between v_i, v_j.
6. G is acyclic but adding a new edge creates exactly one cycle.

Definition 2.54. Let G be a connected graph. G is called *traversable* if admits an Eulerian trail.

Theorem 2.55. Let G be a connected graph. G is traversable if and only if G has exactly to odd-degree vertices.

Definition 2.56. Two graphs G, H are said to be *isomorphic* if exists a bijective map $f : V(G) \to V(H)$ such that $vv' \in E(G) \iff f(v)f(v') \in E(H)$.

Proposition 2.57. Two finite isomorphic graphs have the same order, size and degree sequence.

Theorem 2.58. Two graphs G, H are isomorphic if and only if exists a permutation matrix P such that

$$PA(G)P^t = A(H)$$

where $A(G)$, $A(H)$ are adjacency matrices of G, H, respectively.

2.2.3 | Linear programming

Definition 2.59. Given vectors $c, u, v \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and a matrix $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, we define the *linear programming to maximize*[11] as

$$\text{LP} = \begin{cases} \max: & z = c^t x & \textit{(objective function)} \\ \text{subject to}: & Ax \lesseqgtr b & \textit{(restrictions)} \\ & u \leq x \leq v \end{cases}$$

Definition 2.60. Given vectors $c, u, v \in \mathbb{R}^n$, $b \in \mathbb{R}^m$ and a matrix $A \in \mathcal{M}_{m\times n}(\mathbb{R})$, we define the *canonical form of a linear programming to maximize* as

$$\text{LP} = \begin{cases} \max: & z = c^t x & \textit{(objective function)} \\ \text{subject to}: & Ax \leq b & \textit{(restrictions)} \\ & u \leq x \leq v \end{cases}$$

Analogously we define the *canonical form of a linear programming to minimize* as

$$\text{LP} = \begin{cases} \min: & z = c^t x \\ \text{subject to}: & Ax \geq b \\ & u \leq x \leq v \end{cases}$$

[11] Analogously we can define a *linear programming to minimize* changing the objective function to a minimize function.

Definition 2.61. Given a linear program, the *feasible region* of the program is the set

$$\mathfrak{F} = \{x \in \mathbb{R}^n : Ax \lesseqgtr b, u \leq x \leq v\}.$$

That is, the set of the points that satisfy the conditions of the problem. Given an $x \in \mathbb{R}^n$, x is a *feasible solution* of the linear program if and only if $x \in \mathfrak{F}$.

Definition 2.62. A *polyhedron* P is a set of $\mathbb{R}^n$ that can be expressed as an intersection of a finite collection of half-spaces, that is

$$P = \{x \in \mathbb{R}^n : Ax \geq b, A \in \mathcal{M}_{m\times n}(\mathbb{R}), b \in \mathbb{R}^m\}.$$

A *polytope* is a non-empty and bounded polyhedron. The feasible region of any linear program is a polyhedron.

Definition 2.63. Let $P \subset \mathbb{R}^n$ be a polyhedron. A point $x \in \mathbb{R}^n$ is an extreme point of P if there is neither a pair of points $y, z \in P$, nor a scalar $\lambda \in [0, 1]$ such that $x = \lambda y + (1 - \lambda)z$.

Definition 2.64. Let LP be a linear program. We define the *standard form of LP* as

$$\text{LP} = \begin{cases} \min: & z = c^t x \\ \text{subject to}: & Ax = b \\ & x \geq 0 \end{cases}$$

Definition 2.65. Let $\text{LP} = \min_{x\in\mathbb{R}^n}\{c^t x : Ax = b, x \geq 0\}$. Feasible solution in which free variables or non-basic variable equal zero with respect to basis of basic variables are called *basic feasible solutions*.

Proposition 2.66. If a linear program admits feasible solutions, exists a basic feasible solution. If a linear program admits an optimal solution, exists an optimal basic feasible solution.

Theorem 2.67. Let P be a non-empty polyhedron of a linear program in standard form with maximum rank and let $x \in P$. Then x is an extreme point of P if and only if x is a basic feasible solution.

Definition 2.68 (Simplex method: Phase I). Given a linear program in standard form

$$\text{LP} = \begin{cases} \min: & z = c^t x \\ \text{subject to}: & Ax = b \\ & x \geq 0 \end{cases}$$

its associated problem in phase I (LP_1) is

$$\text{LP}_1 = \begin{cases} \min: & w = \sum_{i=1}^{m} y_i \\ \text{subject to}: & Ax + I_m y = b \\ & x, y \geq 0 \end{cases}$$

A condition necessary for LP having basic feasible solutions is that the optimal solution of LP_1 must be $w = 0$. In fact, if $w \neq 0$, then the original linear program has no feasible solutions[12].

Proposition 2.69 (Simplex method: Phase II). Suppose in a simplex table with positive pivots and therefore independent-terms vector $d \geq 0$, there is a coefficient $c_j < 0$.

$$\left(\begin{array}{c|c} * & d^t \\ \hline c & z - z_0 \end{array}\right).$$

To find a basic feasible solution with lower cost, we make the following change of variable:

1. The variable in column j becomes a basic variable.
2. The variable in row i such that

$$\frac{d_i}{a_{ij}} = \min\left\{\frac{d_k}{a_{kj}} : a_{kj} > 0\right\}$$

becomes a non-basic variable. If this variable does not exists, that is, $a_{kj} \leq 0\ \forall k$ then the linear program is not bounded.

Definition 2.70 (Dual program). Let $\text{LP} = \min_{x\in\mathbb{R}^n}\{c^t x : Ax \geq b, x \geq 0\}$. We define the *dual program of LP* as

$$\text{LP}^* = \begin{cases} \max: & z = b^t y \\ \text{subject to}: & A^t y \leq c \\ & y \geq 0 \end{cases}$$

The linear program LP is called *primal*.

Theorem 2.71 (Weak duality theorem). Let x be a feasible solution of the primal linear program and y a feasible solution of the dual linear program. Then we have:

- $c^t x \leq d^t y$ if the primal linear program is in canonical form to maximize.
- $c^t x \geq d^t y$ if the primal linear program is in canonical form to minimize.

Corollary 2.72. Let x, y be feasible solutions of the primal and dual linear programs respectively such that $c^t x = d^t y$. Then x and y are optimal solutions.

Theorem 2.73 (Strong duality theorem). Any linear program has an optimal solution if and only if its dual linear program does, and in this case, the values coincide.

Theorem 2.74 (Complementary property). Suppose that the optimal table of the primal linear program is of the form

$$\left(\begin{array}{c|c} * & d^t \\ \hline c & z - z_0 \end{array}\right),$$

where $c = (c_1, \ldots, c_{n+m})$ and $d = (d_1, \ldots, d_m)$ with $c_i \geq 0$, $i = 1, \ldots, n + m$. If $(y_1, \ldots y_m, t_1^*, \ldots, t_n^*)$ is the optimal solution of the dual linear program, expressed in standard form, then

$$c_1 = t_1^*, \ldots, c_n = t_n^*, c_{n+1} = y_1, \ldots, c_{n+m} = y_m.$$

[12] This phase is useful to find, if there is, an initial basic feasible solution.

2.3 Functions of several variables

2.3.1 | Topology of $\mathbb{R}^n$

Definition 3.1. Let M be a set. A *distance* in M is an function $d : M \times M \to \mathbb{R}$ such that $\forall x, y, x \in M$ the following properties are satisfied:

1. $d(x, y) \geq 0$.
2. $d(x, y) = 0 \iff x = y$.
3. $d(x, y) = d(y, x)$.
4. $d(x, y) \leq d(x, z) + d(z, y)$ *(triangular inequality).*

We define a *metric space* as a pair (M, d) that satisfy the previous properties.

Definition 3.2. Let E be a real vector space. A *norm* on E is a function $\|\cdot\| : E \to \mathbb{R}$ such that $\forall u, v \in E$ and $\forall \lambda \in \mathbb{R}$ the following properties are satisfied:

1. $\|u\| \geq 0$.
2. $\|u\| = 0 \iff u = 0$.
3. $\|\lambda u\| = |\lambda|\|u\|$.
4. $\|u + v\| \leq \|u\| + \|v\|$ *(triangular inequality).*

We define a *normed vector space* as a pair $(E, \|\cdot\|)$ that satisfy the previous properties.

Proposition 3.3. Let $(E, \|\cdot\|)$ be a normed vector space. Then (E, d) is a metric space with associated distance $d(u, v) := \|u - v\|$, $\forall u, v \in E$.

Definition 3.4. Let E be a real vector space. A *dot product* on E is a function $\langle\cdot,\cdot\rangle : E \times E \to \mathbb{R}$ such that $\forall u, v, w \in E$ and $\forall \alpha, \beta \in \mathbb{R}$ the following properties are satisfied:

1. $\langle \alpha u + \beta w, v\rangle = \alpha\langle u, v\rangle + \beta\langle u, v\rangle$,
   $\langle u, \alpha v + \beta w\rangle = \alpha\langle u, v\rangle + \beta\langle u, w\rangle$.
2. $\langle u, v\rangle = \langle v, u\rangle$.
3. $\langle u, u\rangle \geq 0$.
4. $\langle u, u\rangle = 0 \iff u = 0$.

We define an *euclidean space* as a pair $(E, \langle\cdot,\cdot\rangle)$ that satisfy the previous properties[13].

Proposition 3.5. Let $(E, \langle\cdot,\cdot\rangle)$ be an euclidean space. Then $(E, \|\cdot\|)$ is a normed space with associated norm $\|u\| := \sqrt{\langle u, u\rangle}$.

Proposition 3.6. Let $\langle\cdot,\cdot\rangle_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ be a map defined by

$$\langle u, v\rangle_2 = \sum_{i=1}^{n} u_i v_i$$

$\forall u, v \in \mathbb{R}^n$, being $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$. Then the pair $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_2)$ is an euclidean space.

Corollary 3.7. Consider the norm $\|\cdot\|_2$ and distance d_2 in $\mathbb{R}^n$ defined as follows:

$$\|u\|_2 = \sqrt{\langle u, u\rangle_2} = \sqrt{\sum_{i=1}^{n} u_i^2},$$

$$d_2(u, v) = \|u - v\| = \sqrt{\sum_{i=1}^{n} (u_i - v_i)^2}.$$

Then, $(\mathbb{R}^n, \|\cdot\|_2)$ is a normed space and $(\mathbb{R}^n, d_2)$ is a metric space.

Proposition 3.8. Let $(E, \langle\cdot,\cdot\rangle)$ be an euclidean space with the norm defined as $\|u\| := \sqrt{\langle u, u\rangle}$. Then for all $u, v \in E$ the following properties are satisfied:

1. $\langle u, v\rangle \leq \|u\|\|v\|$ *(Cauchy-Schwarz inequality).*
2. $\|u - v\| \geq \|u\| - \|v\|$.
3. $\|u+v\|^2+\|u-v\|^2 = 2(\|u\|^2+\|v\|^2)$ *(Parallelogram law).*
4. $\|u + v\|^2 - \|u - v\|^2 = 4\langle u, v\rangle$.
5. On $(\mathbb{R}^n, \langle\cdot,\cdot\rangle_2)$, if $u = (u_1, \ldots, u_n)$, then:

$$|u_i| \leq \|u\| \leq \sum_{i=1}^{n} |u_i|.$$

Definition 3.9. Let $L : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. We define the *norm of* L as

$$\|L\| = \sup\{\|L(x)\| : \|x\| = 1\}.$$

Lemma 3.10. Let $\Phi : \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}$ be a map defined as $\Phi(L) = \|L\|$. Then Φ is a norm on the vector space $\mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Proposition 3.11. Let $L \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$. Then:

$$\|L\| = \inf\{C : \|L(x)\| \leq C\|x\|\}.$$

Corollary 3.12. Let $L, M \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ be linear maps with associated matrices $L = (a_{ij})$, $M = (b_{ij})$ respectively. The following properties are satisfied:

1. $\|L(x)\| \leq \|L\|\|x\|$.
2. $\|L\| \leq \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}$.
3. $|a_{ij} - b_{ij}| < \varepsilon, \forall i, j \iff \|L - M\| < \varepsilon'$.

[13] Sometimes the notation $u \cdot v$ is used, instead of $\langle u, v\rangle$, to denote the dot product between u and v.

Definition 3.13. Let (M,d) be a metric space. A *sphere with center p and radius $r \in \mathbb{R}^+$* is the set $S(p,r) = \{x \in M : d(x,p) = r\}$.

Definition 3.14. Let (M,d) be a metric space. An *open ball with center p and radius $r \in \mathbb{R}^+$* is the set $B(p,r) = \{x \in M : d(x,p) < r\}$.

Definition 3.15. Let (M,d) be a metric space. A *closed ball with center p and radius $r \in \mathbb{R}^+$* is the set $\overline{B}(p,r) = \{x \in M : d(x,p) \leq r\}$.

Definition 3.16. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. A is a *bounded set* if exists a ball containing it.

Definition 3.17. Let (M,d) be a metric space. A *neighborhood of p* is a bounded set $E(p) \subset M$ such that $\exists r \in \mathbb{R}^+$ satisfying $B(p,r) \subset E(p)$.

Definition 3.18. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. p is an *interior point of A* if $\exists r \in \mathbb{R}^+$ such that $B(p,r) \subset A$. The *interior of A* is the set $\mathring{A}$ containing all interior points of A.

Definition 3.19. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. p is an *exterior point of A* if $\exists r \in \mathbb{R}^+$ such that $B(p,r) \cap A = \emptyset$. The *exterior of A* is the set $\mathring{A}$ containing all exterior points of A.

Definition 3.20. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. p is an *adherent point* of A if $\forall r \in \mathbb{R}^+$, $B(p,r) \cap A \neq \emptyset$. The *adherence* of A is the set $\overline{A}$ containing all adherent points of A.

Definition 3.21. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. p is a *limit point* of A if $\forall r \in \mathbb{R}^+$, $B(p,r) \setminus \{p\} \cap A \neq \emptyset$. The *limit set* of A is the set A' containing all limit points of A.

Definition 3.22. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. p is an *isolated point* of A if it is an adherent but not limit point, that is, if $p \in A$ and $\exists r \in \mathbb{R}^+$ such that $B(p,r) \setminus \{p\} \cap A = \emptyset$.

Definition 3.23. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. p is a *boundary point* of A if $\forall r \in \mathbb{R}^+$, $B(p,r) \cap A \neq \emptyset$ and $B(p,r) \cap A^c \neq \emptyset$. The *boundary* of A is the set ∂A containing all boundary points of A.

Proposition 3.24. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. If p is a limit point of A, then $B(p,r)$ has infinity many point of A, $\forall r \in \mathbb{R}^+$.

Theorem 3.25 (Bolzano-Weierstraß theorem). Let $B \subset \mathbb{R}^n$ be a set. If B has infinity many points and it is bounded, then it has at least a limit point.

Definition 3.26. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. A is *open* if $\forall p \in A$, $\exists r \in \mathbb{R}^+$ such that $B(p,r) \subset A$.

Definition 3.27. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. A is *closed* if its complementary A^c is open.

Proposition 3.28. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. A is closed $\iff A = \overline{A} \iff \partial A \subset A \iff A' \subset A$.

Proposition 3.29. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M. A is open $\iff A = \mathring{A}$.

Proposition 3.30. Let (M,d) be a metric space and $A \subseteq M$ be a subset of M.

- $\mathring{A}$ is the biggest open set contained in A, that is, if $B \subset A$ is open, $B \subset \text{Int } A$.
- $\overline{A}$ is the smallest set contained in A, that is, if $B \supset A$ is closed, $\overline{A} \supset B$.

Proposition 3.31.

- Union of open sets is open.
- Intersection of a finite number of open sets is open.
- Union of a finite number of closed sets is closed.
- Intersection of closed sets is closed.

Definition 3.32. We say a set A is *connected* if there are no open sets $U, V \neq \emptyset$ such that:

$$A \subseteq U \cup V, \quad A \cap U \cap V = \emptyset, \quad A \cap U \neq \emptyset, \quad A \cap V \neq \emptyset.$$

Definition 3.33. Let (M,d) be a metric space. A *sequence* (x_n) in M is a map

$$\begin{aligned} \mathbb{N} &\longrightarrow M \\ n &\longmapsto x_n \end{aligned}$$

Definition 3.34. Let (M,d) be a metric space. We say $(x_n) \subset M$ is *convergent* to $p \in M$ if

$$\forall \varepsilon \in \mathbb{R}^+, \ \exists n_0 \in \mathbb{N} : d(x_n, p) < \varepsilon \text{ if } n > n_0.$$

Definition 3.35. Let (M,d) be a metric space. We say a sequence (x_n) is a *Cauchy sequence* if $\forall \varepsilon > 0$ $\exists n_0$ such that $d(x_n, x_m) < \varepsilon$, for all $m, n \geq n_0$.

Definition 3.36. A metric space (M,d) is *complete* if every Cauchy sequence in M converges in M.

Definition 3.37. A subset $K \subset \mathbb{R}^n$ is *compact* if it is closed and bounded.

Theorem 3.38. Let $K \subset \mathbb{R}^n$ be an arbitrary set and $(x_m) \in K$ be a sequence. Then K is compact if and only if there exists a partial sequence (x_{m_k}) and $x \in K$ such that $\lim_{k \to \infty} x_{m_k} = x$.

2.3.2 | Continuity

Definition 3.39 (Graph of a function). Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$. We define the *graph of* f as the following subset of $\mathbb{R}^{n+1}$:

$$\text{graph}(f) = \{(x, f(x)) \in \mathbb{R}^{n+1} : x \in U\}.$$

Definition 3.40. Given a function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, we define the *level set* $C_k(f)$ as $C_k(f) = \{x \in \mathbb{R}^n : f(x) = k\}$.

Definition 3.41. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $p \in U'$. We say $\lim_{x\to p} \boldsymbol{f}(x) = L$ if $\forall \varepsilon > 0$, $\exists \delta > 0$ such that $\|\boldsymbol{f}(x) - L\| < \varepsilon$ if $\|x - p\| < \delta$.

Proposition 3.42. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{f} = (f_1, \dots, f_m)$, and $p \in U'$.

1. The limit of $\boldsymbol{f}$ at point p, if exists, is unique.
2. $\lim_{x\to p} \boldsymbol{f}(x) = L \iff \lim_{x\to p} f_j(x) = L_j \quad \forall j = 1, \dots, m.$

Lemma 3.43. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $\ell \in U'$. $\exists \lim_{x\to\ell} \boldsymbol{f}(x) = L \iff \forall (x_n) \in \mathbb{R}^n : \lim_{n\to\infty} x_n = \ell$ and $x_n \neq \ell$ for all n we have $\lim_{n\to\infty} \boldsymbol{f}(x_n) = L$.

Definition 3.44. We say that $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is *continuous* at $p \in U'$ if exists $\lim_{x\to p} \boldsymbol{f}(x) = \boldsymbol{f}(p)$. We say that $\boldsymbol{f}$ is continuous on U, if it is at each point $p \in U$.

Definition 3.45. We say that $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is *uniformly continuous* on U if $\forall \varepsilon > 0, \exists \delta > 0 : \|\boldsymbol{f}(x) - \boldsymbol{f}(y)\| < \varepsilon$, $\forall x, y \in U : \|x - y\| < \delta$.

Corollary 3.46. A uniformly continuous function is continuous.

Theorem 3.47 (Heine's theorem). Let $\boldsymbol{f} : K \subset \mathbb{R}^n \to \mathbb{R}^m$ be continuous function and K a compact set. Then $\boldsymbol{f}$ is uniformly continuous on K.

Theorem 3.48. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be an uniformly continuous function and $(x_n) \in U$ be a Cauchy sequence. Then $(\boldsymbol{f}(x_n)) \in \mathbb{R}^m$ is a Cauchy sequence.

Theorem 3.49. Let $\boldsymbol{f} : K \subset \mathbb{R}^n \to \mathbb{R}^m$ be a continuous function and K be a compact set. Then $\boldsymbol{f}(K)$ is a compact set.

Theorem 3.50 (Weierstraß' theorem). Let $f : K \subset \mathbb{R}^n \to \mathbb{R}$ be a continuous function and K a compact set. Then f attains a maximum and a minimum on K.

Theorem 3.51 (Intermediate value theorem). Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a continuous function and U be a connected set. Then $\forall x, y \in U$ and $\forall c \in [f(x), f(y)]$, $\exists z \in U : f(z) = c$.

Definition 3.52. A function $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is called *Lipschitz continuous* if $\exists k > 0$ such that

$$\|\boldsymbol{f}(x) - \boldsymbol{f}(y)\| \leq k\|x - y\|$$

$\forall x, y \in U$. If $0 \leq k < 1$ we say that $\boldsymbol{f}$ is a *contraction*.

Proposition 3.53. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz continuous function at $p \in U$. Then $\boldsymbol{f}$ is continuous at p.

Definition 3.54. Let (M, d) be a metric space and $f : M \to \mathbb{R}$ a function. We define the *modulus of continuity of* f as the function $\omega_f : (0, \infty) \to [0, \infty]$ defined as

$$\omega_f(\delta) := \sup\{|f(x) - f(y)| : d(x, y) < \delta, x, y \in M\}.$$

2.3.3 | Differential calculus

Differential of a function

Definition 3.55. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $a \in U$. The function $\boldsymbol{f}$ is *differentiable at* a if there exists a linear map $D\boldsymbol{f}(a) : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\lim_{x\to a} \frac{\|\boldsymbol{f}(x) - \boldsymbol{f}(a) - D\boldsymbol{f}(a)(x - a)\|}{\|x - a\|} = \\ = \lim_{h\to 0} \frac{\|\boldsymbol{f}(a + h) - \boldsymbol{f}(a) - D\boldsymbol{f}(a)h\|}{\|h\|} = 0.$$

$D\boldsymbol{f}(a)$ is called the *differential of* $\boldsymbol{f}$ *at point* a. Furthermore, we say $\boldsymbol{f}$ is differentiable on $B \subseteq U$ if it is differentiable at each point of B.

Proposition 3.56. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $a \in U$. $\boldsymbol{f} = (f_1, \dots, f_m)$ is differentiable at a if and only if every component function $f_j : U \subseteq \mathbb{R}^n \to \mathbb{R}$ is differentiable at a.

Definition 3.57. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}$, $a \in U$ and $\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v}\| = 1$. The *directional derivative of* $\boldsymbol{f}$ *at* a in the direction of $\mathbf{v}$ is

$$D_{\mathbf{v}}\boldsymbol{f}(a) = \lim_{t\to 0} \frac{\boldsymbol{f}(a + t\mathbf{v}) - \boldsymbol{f}(a)}{t}.$$

Definition 3.58. Let $U \subseteq \mathbb{R}^n$ be an open set, $\boldsymbol{f} : U \to \mathbb{R}$ and $a \in U$. If the following limit exists, we define the *partial derivative with respect to* x_j *of* $\boldsymbol{f}$ *at* a as

$$\frac{\partial \boldsymbol{f}}{\partial x_j}(a) = \lim_{h\to 0} \frac{\boldsymbol{f}(a + h\mathbf{e}_j) - \boldsymbol{f}(a)}{h}.$$ [14]

Definition 3.59. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $a \in U$. If all partial derivatives of $\boldsymbol{f}$ at a exist, we call *Jacobian matrix of* $\boldsymbol{f}$ *at* a the matrix associated with $D\boldsymbol{f}(a)$ (with respect to the canonical basis of $\mathbb{R}^n$ and $\mathbb{R}^m$):

$$D\boldsymbol{f}(a) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{pmatrix}.$$

[14] Here $\mathbf{e}_j$ is the j-th vector of the canonical basis of $\mathbb{R}^n$, that is, $\mathbf{e}_j = (0, \dots, 0, \overset{(j)}{1}, 0, \dots, 0)$.

If $n = m$, we define the *Jacobian determinant* as $J\boldsymbol{f}(a) = \det D\boldsymbol{f}(a)$.

Definition 3.60. Let $U \subseteq \mathbb{R}^n$ be an open set, $f : U \to \mathbb{R}$ and $a \in U$ such that f is differentiable at $a \in U$. The *gradient of f at a* is

$$\nabla f(a) := Df(a) = \left(\frac{\partial f}{\partial x_1}(a), \ldots, \frac{\partial f}{\partial x_n}(a)\right).$$

Proposition 3.61. Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ be a differentiable function at $a \in U$. Then there exists the tangent hyperplane to the graph of f at a and has the equation

$$x_{n+1} = f(a) + \nabla f(a) \cdot (x - a)^{15}.$$

Theorem 3.62. Let $U \subseteq \mathbb{R}^n$ be an open set, $f : U \to \mathbb{R}$, $a \in U$ and $\mathbf{v} \in \mathbb{R}^n : \|\mathbf{v}\| = 1$. If f is differentiable at a, then $D_{\mathbf{v}} f(a)$ exists and

$$D_{\mathbf{v}} f(a) = \nabla f(a) \cdot \mathbf{v}.$$

Proposition 3.63. Let $U \subseteq \mathbb{R}^n$ be an open set, $f : U \to \mathbb{R}$ be a differentiable function on U and C_k be the level set of value $k \in \mathbb{R}$. Then $\nabla f(a) \perp C_k$ at $a \in C_k$.

Proposition 3.64. Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ a differentiable function at $a \in U$ and $\mathbf{v} \in \mathbb{R}^n$. Then:

- $\max\{D_{\mathbf{v}} f(a) : \|\mathbf{v}\| = 1\} = \|\nabla f(a)\|$ and it is attained when $\mathbf{v} = \dfrac{\nabla f(a)}{\|\nabla f(a)\|}$.
- $\min\{D_{\mathbf{v}} f(a) : \|\mathbf{v}\| = 1\} = -\|\nabla f(a)\|$ and it is attained when $\mathbf{v} = -\dfrac{\nabla f(a)}{\|\nabla f(a)\|}$.

Theorem 3.65. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a differentiable function at $a \in U$. Then $\boldsymbol{f}$ is locally Lipschitz continuous at a.

Theorem 3.66. Let $\boldsymbol{f}, \boldsymbol{g} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be two differentiable functions at a point $a \in U$ and let $c \in \mathbb{R}$. Then:

1. $\boldsymbol{f} + \boldsymbol{g}$ is differentiable at a and
$$D(\boldsymbol{f} + \boldsymbol{g})(a) = D\boldsymbol{f}(a) + D\boldsymbol{g}(a).$$
2. $c\boldsymbol{f}$ is differentiable at a and
$$D(c\boldsymbol{f})(a) = cD\boldsymbol{f}(a).$$
3. If $m = 1$, then $(fg)(x) = f(x)g(x)$ is differentiable at a and
$$D(fg)(a) = g(a)Df(a) + f(a)Dg(a).$$
4. If $m = 1$ and $g(a) \neq 0$, then $\left(\dfrac{f}{g}\right)(x) = \dfrac{f(x)}{g(x)}$ is differentiable at a and
$$D\left(\frac{f}{g}\right)(a) = \frac{g(a)Df(a) - f(a)Dg(a)}{[g(a)]^2}.$$

Theorem 3.67 (Chain rule). Let $U \subseteq \mathbb{R}^n$ and $V \subseteq \mathbb{R}^m$ be open sets. Let $\boldsymbol{f} : U \to \mathbb{R}^m$ and $\boldsymbol{g} : V \to \mathbb{R}^p$. Suppose that $\boldsymbol{f}(U) \subset V$, $\boldsymbol{f}$ is differentiable at $a \in U$ and $\boldsymbol{g}$ is differentiable at $\boldsymbol{f}(a)$. Then $\boldsymbol{g} \circ \boldsymbol{f}$ is differentiable at a and

$$D(\boldsymbol{g} \circ \boldsymbol{f})(a) = D\boldsymbol{g}(\boldsymbol{f}(a)) \circ D\boldsymbol{f}(a).$$

Definition 3.68. Let $U \subseteq \mathbb{R}^n$ be an open set and $\boldsymbol{f} : U \to \mathbb{R}^m$. We say that $\boldsymbol{f}$ is a *function of class $\mathcal{C}^k(U)$*, $k \in \mathbb{N}$, if all partial derivatives of order k exists and are continuous on U. We say that $\boldsymbol{f}$ is *function of class $\mathcal{C}^\infty(U)$* if it is of class $\mathcal{C}^k(U)$, $\forall k \in \mathbb{N}$.

Theorem 3.69 (Differentiability criterion). Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{f}(x) = (f_1(x), \ldots, f_m(x))$. If all partial derivatives $\dfrac{\partial f_i(x)}{\partial x_j}$ exists in a neighborhood of $a \in U$ and are continuous at a, then $\boldsymbol{f}$ is differentiable at $a \in U$.

Proposition 3.70. Let $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $A \subseteq U$. If all partial derivatives of $\boldsymbol{f}$ exist on A and are bounded functions on A, then $\boldsymbol{f}$ is uniformly continuous on A.

Theorem 3.71 (Mean value theorem). Let $f : B \to \mathbb{R}$ be a function of class $\mathcal{C}^1$ in an open connected set B. Let $x, y \in B$. Then,

$$f(x) - f(y) = \nabla f(z) \cdot (x - y),$$

for some $z \in [x, y]$.

Theorem 3.72 (Mean value theorem for vector-valued functions). Let $\boldsymbol{f} : B \to \mathbb{R}^m$ be a function of class $\mathcal{C}^1$ in an open connected set B. Let $x, y \in B$. Then,

$$\|\boldsymbol{f}(x) - \boldsymbol{f}(y)\| \leq \|Df(z)\|\|x - y\|,$$

for some $z \in [x, y]$.

Higher order derivatives

Definition 3.73. Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$. We denote the *partial derivative of f of order k with respect to the variables $x_{i_1}, \ldots, x_{i_k}$* as

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}.$$

[15] In general (not only the case of the graph of a function) the tangent hyperplane to function f at a point a is given by the equation

$$\nabla f(a) \cdot (x - a) = 0.$$

Definition 3.74. Let $U \subseteq \mathbb{R}^n$ be an open set. If $f : U \to \mathbb{R}$ has second order partial derivatives at $a \in U$, we define the *hessian matrix of f at a point a* as

$$Hf(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n x_1}(a) \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 x_n}(a) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(a) \end{pmatrix}.$$

Theorem 3.75 (Schwarz's theorem). Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$. If f has mixed partial derivatives of order k and are continuous functions on $A \subseteq U$, then for any permutation $\sigma \in S_k$ we have

$$\frac{\partial^k f}{\partial x_{i_k} \cdots \partial x_{i_1}}(a) = \frac{\partial^k f}{\partial x_{\sigma(i_k)} \cdots \partial x_{\sigma(i_1)}}(a), \quad \forall a \in A.$$

Inverse and implicit function theorems

Lemma 3.76. Let $U \subseteq \mathbb{R}^n$ be an open set and $\boldsymbol{f} : U \to \mathbb{R}^m$ with $\boldsymbol{f} \in \mathcal{C}^1(U)$. Given an $a \in U$ and $\varepsilon > 0$, $\exists B(a,r) \subset U$ such that

$$\|\boldsymbol{f}(x) - \boldsymbol{f}(y)\| \leq (\|D\boldsymbol{f}(a)\| + \varepsilon)\|x - y\|, \quad \forall x, y \in B(a,r).$$

Lemma 3.77. Let $U \subseteq \mathbb{R}^n$ be an open set and $\boldsymbol{f} : U \to \mathbb{R}^n$ with $\boldsymbol{f} \in \mathcal{C}^1(U)$. Suppose that for some $a \in U$, $J\boldsymbol{f}(a) \neq 0$. Then $\exists B(a,r) \subset U$ and $c > 0$ such that

$$\|\boldsymbol{f}(y) - \boldsymbol{f}(x)\| \geq c\|y - x\|, \quad \forall x, y \in B(a,r).$$

In particular, $\boldsymbol{f}$ is injective on $B(a,r)$.

Theorem 3.78 (Inverse function theorem). Let $U \subseteq \mathbb{R}^n$ be an open set, $\boldsymbol{f} : U \to \mathbb{R}^n$ with $\boldsymbol{f} \in \mathcal{C}^1(U)$ and $a \in U$ such that $J\boldsymbol{f}(a) \neq 0$. Then $\exists B = B(a,r) \subset U$ such that:

1. $\boldsymbol{f}$ is injective on B.
2. $\boldsymbol{f}(B) = V$ is an open set of $\mathbb{R}^n$.
3. $\boldsymbol{f}^{-1} : V \to B$ is of class $\mathcal{C}^1$ on V.

Moreover, it is satisfied that $D\boldsymbol{f}^{-1}(\boldsymbol{f}(a)) = D\boldsymbol{f}(a)^{-1}$

Definition 3.79. A function $\boldsymbol{f} : U \subseteq \mathbb{R}^n \to \mathbb{R}^n$ is a *diffeomorphism of class $\mathcal{C}^k$* if it is bijective and both $\boldsymbol{f}$ and $\boldsymbol{f}^{-1}$ are of class $\mathcal{C}^k$.

Theorem 3.80 (Implicit function theorem). Let $U \subseteq \mathbb{R}^{n+m}$ be an open set, $\boldsymbol{f} : U \to \mathbb{R}^m$ with $\boldsymbol{f} \in \mathcal{C}^1(U)$ and $(a,b) = (a_1, \ldots, a_n, b_1, \ldots, b_m) \in U$ such that $\boldsymbol{f}(a,b) = 0$. If $D\boldsymbol{f}(x) = (Df_1(x) \mid Df_2(x))$ with $Df_1(x) \in \mathcal{M}_{m\times n}(\mathbb{R})$, $Df_2(x) \in \mathcal{M}_m(\mathbb{R})$ and $\det Df_2(x) \neq 0$ (i.e. rang $D\boldsymbol{f}(a,b) = m$), then exists an open set $W \subseteq \mathbb{R}^n$ such that $a \in W$ and a function $\boldsymbol{g} : W \to \mathbb{R}^m$ with $\boldsymbol{g} \in \mathcal{C}^1(W)$, such that

$$\boldsymbol{g}(a) = b \quad \text{and} \quad \boldsymbol{f}(x, \boldsymbol{g}(x)) = 0 \quad \forall x \in W.$$

Moreover, is is satisfied that

$$D\boldsymbol{g}(a) = -Df_2(a, \boldsymbol{g}(a))^{-1} \circ Df_1(a, \boldsymbol{g}(a)).$$

Taylor's polynomial and maxima and minima

Theorem 3.81 (Taylor's theorem). Let $U \subseteq \mathbb{R}^n$ be an open set, $f : U \to \mathbb{R}$, $a \in U$ and $f \in \mathcal{C}^{k+1}(U)$. Then:

$$f(x) = f(a) + \\ + \sum_{m=1}^{k} \frac{1}{m!} \left(\sum_{i_m, \ldots, i_1 = 1}^{n} \frac{\partial^m f}{\partial x_{i_m} \cdots \partial x_{i_1}}(a) \prod_{j=1}^{m} (x_{i_j} - a_{i_j}) \right) + \\ + R_k(f,a),$$

where

$$R_k(f,a) = \\ = \frac{1}{(k+1)!} \sum_{i_{k+1}, \ldots, i_1 = 1}^{n} \frac{\partial^{k+1} f}{\partial x_{i_{k+1}} \cdots \partial x_{i_1}}(\xi) \prod_{j=1}^{k+1} (x_{i_j} - a_{i_j}) = \\ = o(\|x - a\|^k)$$

for some $\xi \in [a, x]$. In particular, for $k = 2$ we have:

$$f(x) = f(a) + Df(a)(x - a) + \frac{1}{2} Hf(a)(x - a, x - a) + \\ + R_2(f,a),$$

where $R_2(f,a) = o(\|x - a\|^2)$.

Definition 3.82. Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$. We say that f has a *local maximum at $a \in U$* if $\exists B(a,r) \subset U : f(x) \leq f(a)$, $\forall x \in B(a,r)$. Analogously, we say that f has a *local minimum at $a \in U$* if $\exists B(a,r) \subset U : f(x) \geq f(a)$, $\forall x \in B(a,r)$. A *local extremum* is either a local maximum or a local minimum. Moreover, if $f(x) \leq f(a)$ $\forall x \in U$, we say that f has a *global maximum at $a \in U$*. Similarly if $f(x) \geq f(a)$ $\forall x \in U$, we say that f has a *global minimum at $a \in U$*.

Proposition 3.83. Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ be a differentiable function at $a \in U$. If f has a local extremum at a, then $\nabla f(a) = 0$.

Definition 3.84. Let $U \subseteq \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$. We say that $a \in U$ is a *critical point of f* if $\nabla f(a) = 0$. We say that $a \in U$ is a *saddle point* if a is a critical point but not a local extremum.

Theorem 3.85. Let $\mathcal{Q}$ be a quadratic form. Then for all $x \neq 0$ we have:

$$\mathcal{Q} \text{ is defined positive} \iff \exists \lambda \in \mathbb{R}^+ : \mathcal{Q}(x) \geq \lambda \|x\|^2.$$
$$\mathcal{Q} \text{ is defined negative} \iff \exists \lambda \in \mathbb{R}^- : \mathcal{Q}(x) \leq \lambda \|x\|^2.$$

Proposition 3.86 (Sylvester's criterion). Let $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{R})$ be a symmetric matrix. A is defined positive if and only if all its principal minors are positive, that is:

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \ldots, \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} > 0.$$

A is defined negative if and only if its principal minor of order k have sign $(-1)^k$, that is:

$$a_{11} < 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \dots, (-1)^n \begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} > 0.$$

Theorem 3.87. Let $U \subseteq \mathbb{R}^2$ be an open set, $f : U \to \mathbb{R}$ a function of class $\mathcal{C}^2(U)$ and $a \in U : \nabla f(a) = 0$. Let $Hf(a)$ be the hessian matrix of f at a and $\mathcal{H}f(a)$ be its associated quadratic form. Then:

1. If $\mathcal{H}f(a)$ is defined positive $\implies$ f has a local minimum at a.
2. If $\mathcal{H}f(a)$ is defined negative $\implies$ f has a local maximum at a.
3. If $\mathcal{H}f(a)$ is undefined $\implies$ f has a saddle point at a.

Theorem 3.88 (Lagrange multipliers theorem). Let $f, g_i : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be functions of class $\mathcal{C}^1(U)$ for $i = 1, \dots, k$ and $1 \leq k < n$. Let $S = \{x \in U : g_i(x) = 0,\ \forall i\}$ and $a \in S$ such that $f_{|S}(a)$ is a local extremum. If the vectors $\nabla g_1(a), \dots, \nabla g_k(a)$ are linearly independents, then $\exists \lambda_1, \dots, \lambda_k \in \mathbb{R}$ such that:

$$\nabla f(a) = \sum_{i=1}^{k} \lambda_i \nabla g_i(a).$$

2.3.4 | Integral calculus

Integration over compact rectangles

Definition 3.89. A *rectangle* R of $\mathbb{R}^n$ is a product $R = I_1 \times \cdots \times I_n$ where $I_j \in \mathbb{R}$ are bounded and non-degenerate[16] intervals.

Definition 3.90. The *n-dimensional volume* (*surface* if $n = 2$ or *length* if $n = 1$) of a bounded rectangle $R = I_1 \times \cdots \times I_n$, $I_i = [a_i, b_i]$ is:

$$\text{vol}(R) = \prod_{i=1}^{n} (b_i - a_i).$$

Definition 3.91. Given a rectangle $R = I_1 \times \cdots \times I_n$, a *partition of R* is the product $\mathcal{P} = \mathcal{P}_1 \times \cdots \times \mathcal{P}_n$ where $\mathcal{P}_j$ is a partition of the interval I_j. A partition $\mathcal{P}$ is *regular* if for all j, $\mathcal{P}_j$ is regular, that is, all subintervals in $\mathcal{P}_j$ have the same size. We denote by $\mathbf{P}(R)$ the set of all partitions of R.

Definition 3.92. Given two partitions $\mathcal{P} = I_1 \times \cdots \times I_n$ and $\mathcal{P}' = I_1' \times \cdots \times I_n'$ of a rectangle R, we say that $\mathcal{P}'$ is *finer than* $\mathcal{P}$ if each $\mathcal{P}_j'$ is finer than $\mathcal{P}_j$.

Definition 3.93. Let $R \subset \mathbb{R}^n$ be a compact rectangle, $f : R \to \mathbb{R}$ be a bounded function and $\mathcal{P} \in \mathbf{P}(R)$. For each subrectangle R_j, $j = 1, \dots, m$, determined by $\mathcal{P}$ let

$$m_j = \inf\{f(x) : x \in R_j\} \quad \text{i} \quad M_j = \sup\{f(x) : x \in R_j\}.$$

We define the *lower sum* and the *upper sum of f with respect to $\mathcal{P}$* as:

$$L(f, \mathcal{P}) = \sum_{j=1}^{m} m_j \text{vol}(R), \qquad U(f, \mathcal{P}) = \sum_{j=1}^{m} M_j \text{vol}(R)^{17}.$$

Definition 3.94. Let $R \subset \mathbb{R}^n$ be a compact rectangle and $f : R \to \mathbb{R}$ be a bounded function. We define the *lower integral* and *upper integral of f on R* as

$$\underline{\int_R} f = \sup\{L(f, \mathcal{P}) : \mathcal{P} \in \mathbf{P}\},$$
$$\overline{\int_R} f = \inf\{U(f, \mathcal{P}) : \mathcal{P} \in \mathbf{P}\}.$$

We say that f is *Riemann-integrable on R* if $\underline{\int_R} f = \overline{\int_R} f$.

Proposition 3.95. Let $R \subset \mathbb{R}^n$ be a compact rectangle and $f : R \to \mathbb{R}$ be a bounded function. f is Riemann-integrable if and only if $\forall \varepsilon\ \exists \mathcal{P} \in \mathbf{P}(R)$ such that $U(f, \mathcal{P}) - L(f, \mathcal{P}) < \varepsilon$.

Definition 3.96. Let $R \subset \mathbb{R}^n$ be a compact rectangle; $f : R \to \mathbb{R}$, a bounded function; $\mathcal{P} \in \mathbf{P}(R)$, and ξ_j, an arbitrary point of the subrectangle R_j, $j = 1, \dots, m$. Then we define the *Riemann sum of f associated to $\mathcal{P}$* as:

$$S(f, \mathcal{P}) = \sum_{j=1}^{m} f(\xi_j) \text{vol}(R_j).$$

Theorem 3.97. Let $R \subset \mathbb{R}^n$ be a compact rectangle and $f : R \to \mathbb{R}$ be a bounded function. f is Riemann-integrable over R if and only if $\forall \varepsilon > 0\ \exists \mathcal{P}_\varepsilon \in \mathbf{P}(R)$ such that

$$\left| S(f, \mathcal{P}) - \int_R f \right| = \left| \sum_{j=1}^{m} f(\xi_j) \text{vol}(R_j) - \int_R f \right| < \varepsilon,$$

for any $\mathcal{P} \in \mathbf{P}(R)$ finer than $\mathcal{P}_\varepsilon$ and for any $\xi_j \in R_j$.

[16] That is, non-empty intervals with more than one point.

[17] We will omit the results related to these definitions because of they are a natural extension of results of single-variable functions course and can be deduced easily. That's why we only expose the most important ones here.

Fubini's theorem

Theorem 3.98 (Fubini's theorem). Let $R_1 \subset \mathbb{R}^n$ and $R_2 \subset \mathbb{R}^m$ be closed rectangles and $f : R_1 \times R_2 \to \mathbb{R}$ be an integrable[18] function. Suppose for every $x_0 \in R_1$, $f(x_0, y)$ is integrable over R_2. Then the function $g(x) = \int_{R_2} f(x,y)dy$ is integrable over R_1 and

$$\int_{R_1\times R_2} f(x,y) = \int_{R_1} dx \int_{R_2} f(x,y)dy.$$

Similarly if for every $y_0 \in R_2$, $f(x, y_0)$ is integrable over R_1, then the function $h(y) = \int_{R_1} f(x,y)dx$ is integrable over R_2 and

$$\int_{R_1\times R_2} f(x,y) = \int_{R_2} dy \int_{R_1} f(x,y)dx.$$

Corollary 3.99. Let $R_1 \subset \mathbb{R}^n$ and $R_2 \subset \mathbb{R}^m$ be closed rectangles and let $f : R_1 \times R_2 \to \mathbb{R}$ be a continuous function on $R_1 \times R_2$. Then,

$$\int_{R_1\times R_2} f = \int_{R_1} dx \int_{R_2} f(x,y)dy = \int_{R_2} dy \int_{R_1} f(x,y)dx.$$

Corollary 3.100. Let $R = [a_1, b_1] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ be a rectangle. If $f : R \to \mathbb{R}$ is a continuous function, then

$$\int_R f = \int_{a_n}^{b_n} dx_n \int_{a_{n-1}}^{b_{n-1}} dx_{n-1} \cdots \int_{a_1}^{b_1} f(x_1, \ldots, x_n)dx_1.$$

Definition 3.101. Let $D \subset \mathbb{R}^{n-1}$ be a compact set and $\varphi_1, \varphi_2 : D \to \mathbb{R}$ be continuous functions such that $\varphi_1(x) \le \varphi_2(x)\ \forall x \in D$. The set

$$S = \{(x,y) \subset \mathbb{R}^n : x \in D, \varphi_1(x) \le y \le \varphi_2(x)\}$$

is called an *elementary region in* $\mathbb{R}^n$. In particular, if $n = 2$, we say S is *x-simple*. An elementary region in $V \subset \mathbb{R}^3$ is called *xy-simple* if it is of the form

$$V = \{(x,y,z) \in \mathbb{R}^3 : (x,y) \in U, \phi_1(x,y) \le z \le \phi_2(x,y)\},$$

where U is an elementary region in $\mathbb{R}^2$ and ϕ_1, ϕ_2 are continuous functions on U[19].

Theorem 3.102 (Fubini's theorem for elementary regions). Let $S = \{(x,y) \subset \mathbb{R}^n : x \in D, \varphi_1(x) \le y \le \varphi_2(x)\}$ be an elementary region in $\mathbb{R}^n$ and $f : S \to \mathbb{R}$. If f is integrable over S and for all $x_0 \in D$ the function $f(x_0, y)$ is integrable over $[-M, M]$, $M \in \mathbb{R}$, then

$$\int_S f = \int_D dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x,y)dy.$$

Definition 3.103. Let $D \subset \mathbb{R}^{n-1}$ be a compact set and $S = \{(x,y) \subset \mathbb{R}^n : x \in D, \varphi_1(x) \le y \le \varphi_2(x)\}$ an elementary region. We define the *n-dimensional volume of S* as

$$\mathrm{vol}(S) := \int_S dx = \int_D dx \int_{\varphi_1(x)}^{\varphi_2(x)} dy$$[20].

Corollary 3.104 (Cavalieri's principle). Let $\Omega \subset R \times [a,b]$ be a set in $\mathbb{R}^n$ where $R \subset \mathbb{R}^{n-1}$ is a rectangle. For every $t \in [a,b]$ let

$$\Omega_t = \{(x,y) \in \Omega : y = t\} \subset \mathbb{R}^n$$

be the section of Ω corresponding to the hyperplane $y = t$. If $\nu(\Omega_t)$ is the $(n-1)$-dimensional volume (area if $n = 3$ or length if $n = 2$) of Ω_t, then

$$\mathrm{vol}(\Omega) = \int_a^b \nu(\Omega_t)dt.$$

Definition 3.105 (Center of mass). The *center of mass of an object with mass density* $\rho(x,y,z)$ occupying a region $\Omega \subset \mathbb{R}^3$ is the point $(\overline{x}, \overline{y}, \overline{z}) \in \mathbb{R}^3$ whose coordinates are:

$$\overline{x} = \frac{1}{m}\iiint_\Omega x\rho(x,y,z)dxdydz,$$
$$\overline{y} = \frac{1}{m}\iiint_\Omega y\rho(x,y,z)dxdydz,$$
$$\overline{z} = \frac{1}{m}\iiint_\Omega z\rho(x,y,z)dxdydz,$$

where $m = \iiint_\Omega \rho(x,y,z)dxdydz$ is the total mass of the object.

Definition 3.106 (Moment of inertia). Given a body with mass density $\rho(x,y,z)$ occupying a region $\Omega \subset \mathbb{R}^3$ and a line $L \subset \mathbb{R}^3$, the *moment of inertia of the body about the line* L is

$$I_L = \iiint_\Omega d(x,y,z)^2\rho(x,y,z)dxdydz,$$

where $d(x,y,z)$ denotes the distance from (x,y,z) to the line L. In particular, when L is the z-axis, then

$$I_z = \iiint_\Omega (x^2+y^2)\rho(x,y,z)dxdydz,$$

and similarly for I_x and I_y. The moment of inertia of the body about the xy-plane is definedy by

$$I_{xy} = \iiint_\Omega z^2\rho(x,y,z)dxdydz,$$

and similarly for I_{yz} and I_{zx}.

[18] As we only have defined Riemann-integration, it goes without saying that an *integrable function* means a *Riemann-integrable function.*

[19] Analogously we define *y-simple* regions in $\mathbb{R}^2$ and *yz-simple* or *xz-simple* regions in $\mathbb{R}^3$.

[20] In particular, we define the area of a region $S \subset \mathbb{R}^2$ as $\mathrm{area}(S) = \iint_S dxdy$ and the volume of a region $\Omega \subset \mathbb{R}^3$ as $\mathrm{vol}(\Omega) = \iiint_\Omega dxdydz$.

Change of variable

Theorem 3.107 (Change of variable theorem). Let $U \subseteq \mathbb{R}^n$ be an open set and let $\varphi : U \to \mathbb{R}^n$ be a diffeomorphism. If $f : \varphi(U) \to \mathbb{R}$ is integrable on $\varphi(U)$, then

$$\int_{\varphi(U)} f = \int_U (f \circ \varphi)|J\varphi|.$$

Corollary 3.108 (Integral in polar coordinates). Let $\varphi : [0, \infty) \times [0, 2\pi) \to \mathbb{R}$ be such that

$$\varphi(r, \theta) \longmapsto (r\cos\theta, r\sin\theta).$$

Then we have $|J\varphi| = r$ and therefore:

$$\int_{\varphi(U)} f(x, y)dxdy = \int_U f(r\cos\theta, r\sin\theta)rdrd\theta.$$

Corollary 3.109 (Integral in cylindrical coordinates). Let $\varphi : [0, \infty) \times [0, 2\pi) \times \mathbb{R} \to \mathbb{R}$ be such that

$$\varphi(r, \theta, z) \longmapsto (r\cos\theta, r\sin\theta, z).$$

Then we have $|J\varphi| = r$ and therefore:

$$\int_{\varphi(U)} f(x, y, z)dxdydz = \int_U f(r\cos\theta, r\sin\theta, z)rdrd\theta dz.$$

Corollary 3.110 (Integral in spherical coordinates). Let $\varphi : [0, \infty) \times [0, 2\pi) \times [0, \pi] \to \mathbb{R}$ be such that

$$\varphi(\rho, \theta, \phi) \longmapsto (\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi).$$

Then we have $|J\varphi| = \rho^2\sin\phi$ and therefore:

$$\int_{\varphi(U)} f(x, y, z)dxdydz =$$
$$= \int_U f(\rho\sin\phi\cos\theta, \rho\sin\phi\sin\theta, \rho\cos\phi)\rho^2\sin\phi d\rho d\theta d\phi.$$

2.3.5 | Vector calculus

Arc-length and line integrals

Definition 3.111. Let $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ be a parametrization of a curve and $\mathcal{P} = \{t_0, \dots, t_n\}$ be a partition of $[a, b]$. Then, the *length of the polygonal* created from the vertices $\boldsymbol{\gamma}(t_i)$ is

$$L(\boldsymbol{\gamma}, \mathcal{P}) = \sum_{i=1}^{n} \|\boldsymbol{\gamma}(t_i) - \boldsymbol{\gamma}(t_{i-1})\|.$$

Definition 3.112. Let $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ be a parametrization of a curve c. The *arc length of* c is

$$L(c) = \sup\{L(\boldsymbol{\gamma}, \mathcal{P}) : \mathcal{P} \in \mathbf{P}([a, b])\} \in [0, \infty].$$

Definition 3.113. We say that a curve c is *rectifiable* if it has a finite arc length, that is, if $L(c) < \infty$.

Proposition 3.114. Let $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ be a parametrization of class $\mathcal{C}^1$ of a curve c. Then c is rectifiable and

$$L(c) = \int_a^b \|\boldsymbol{\gamma}'(t)\|dt^{21}.$$

Definition 3.115. Let $\boldsymbol{F} : U \subset \mathbb{R}^m \to \mathbb{R}^n$ be a vector field[22]. If all its component functions F_i are integrable, we define

$$\int_U \boldsymbol{F} = \left(\int_U F_1, \dots, \int_U F_n\right) \in \mathbb{R}^n.$$

Definition 3.116. Let c be a curve in $\mathbb{R}^2$ parametritzed by $\boldsymbol{\gamma} = (x(t), y(t))$. The unit tangent vector to the curve at time t is

$$\mathbf{T} = \frac{\boldsymbol{\gamma}'(t)}{\|\boldsymbol{\gamma}'(t)\|}.$$

The normal vector to the curve is $N(t) = (y'(t), -x'(t))$ and the unit normal vector to the curve is

$$\mathbf{n} = \frac{N(t)}{\|N(t)\|}^{23}.$$

Definition 3.117. Let c be a curve parametritzed by $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ and $\varphi : [c, d] \to [a, b]$ be a diffeomorphism. The composition $\boldsymbol{\gamma} \circ \varphi : [c, d] \to \mathbb{R}^n$ is called a *reparametrization of* c.

Definition 3.118. Let c be a curve of class $\mathcal{C}^1$ parametritzed by $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ an L be its arc length. We define the *arc length parameter* as

$$s(t) = \int_a^t \|\boldsymbol{\gamma}'(t)\|dt.$$

We reparametrize c by $\rho(s) = \boldsymbol{\gamma}(t(s))$, $0 \le s \le L$. Then $\rho'(s)$ is a unit tangent vector to c and $\rho''(s)$ is perpendicular to c at the point $\rho(s)$.

Definition 3.119. Let c be a curve of class $\mathcal{C}^2$ and s be its arc length parameter. We define the *curvature* of c at the point $\rho(s)$ as

$$\kappa(\rho(s)) = \|\rho''(s)\|.$$

Definition 3.120. Let $c = \{\boldsymbol{\gamma}(t) : t \in [a, b]\} \subset \mathbb{R}^n$ be a curve of class $\mathcal{C}^1$ and $f : \mathbb{R}^n \to \mathbb{R}$ be continuous function. We define the *line integral of* f *along* c as

$$\int_c f ds = \int_a^b f(\boldsymbol{\gamma}(t))\|\boldsymbol{\gamma}'(t)\|dt^{24}.$$

[21] It can be seen that the arc length of a curve does not depend on its parametrization.
[22] A *vector field* is nothing more than a vector-valued function.
[23] Observe that $-N(t)$ is also a normal vector to the curve but, by agreement, we take the one pointing to the right of the curve or, if the curve is closed, the one pointing outwards from the curve.
[24] It can be seen that this integral is independent of the parametrization of c.

Definition 3.121. Let $c = \{\boldsymbol{\gamma}(t) : t \in [a,b]\} \subset \mathbb{R}^n$ be a curve of class $\mathcal{C}^1$ and $\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n$ be a continuous vector field. We define the *line integral of* $\boldsymbol{f}$ *along* c as

$$\int_c \boldsymbol{f} \cdot d\mathbf{s} = \int_c \boldsymbol{f} \cdot \mathbf{T} ds = \int_a^b \boldsymbol{f}(\boldsymbol{\gamma}(t)) \cdot \boldsymbol{\gamma}'(t) dt,$$

where $\mathbf{T}$ is the unit tangent vector c[25]. If c is closed, then this integral is called the *circulation of* $\boldsymbol{f}$ *around* c.

Definition 3.122. A *Jordan arc* is the image of an injective continuous map $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$. A *Jordan closed curve* is the image of an injective continuous map $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^n$ such that $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$.

Conservative vector fields

Definition 3.123. Let $U \subseteq \mathbb{R}^n$ be a domain and $f : U \to \mathbb{R}$ be a function of class $\mathcal{C}^1$. We say that $\boldsymbol{f} : U \to \mathbb{R}^n$ is a *conservative* or a *gradient vector field* if

$$\boldsymbol{f}(x) = \nabla f(x), \quad \forall x \in U.$$

The function f is called the *potential* of $\boldsymbol{f}$.

Theorem 3.124. Let $\boldsymbol{f} = \nabla f$ be a conservative vector field on $U \subseteq \mathbb{R}^n$ and c be a closed curve that admits a parametrization $\boldsymbol{\gamma}(t) : [a,b] \to \mathbb{R}^n$ of class $\mathcal{C}^1(U)$. Then

$$\int_c \boldsymbol{f} \cdot d\mathbf{s} = f(\boldsymbol{\gamma}(b)) - f(\boldsymbol{\gamma}(a)).$$

Corollary 3.125. Let $\boldsymbol{f}$ be a conservative vector field on U and c be a closed curve that admits a parametrization of class $\mathcal{C}^1(U)$. Then $\displaystyle\int_c \boldsymbol{f} \cdot d\mathbf{s} = 0$.

Divergence, curl and Laplacian

Definition 3.126. Let $\boldsymbol{f} = (F_1, \ldots, F_n)$ be a vector field of class $\mathcal{C}^1(U)$, $U \subseteq \mathbb{R}^n$. The *divergence of* $\boldsymbol{f}$ is

$$\operatorname{div} \boldsymbol{f} = \nabla \cdot \boldsymbol{f} = \sum_{i=1}^n \frac{\partial F_j}{\partial x_j}.$$

Definition 3.127. Let $\boldsymbol{f} = (F_1, F_2, F_3)$ be a vector field of class $\mathcal{C}^1(U)$, $U \subseteq \mathbb{R}^3$. The *curl de* $\boldsymbol{f}$ is

$$\operatorname{rot} \boldsymbol{f} = \nabla \times \boldsymbol{f} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_1 & F_2 & F_3 \end{vmatrix} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}, \frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}, \frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right).$$

Definition 3.128. Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathcal{C}^2(U)$, $U \subseteq \mathbb{R}^3$. The *Laplacian of* f is

$$\nabla^2 f = \Delta f = \sum_{i=1}^n \frac{\partial^2 f}{\partial x_j^2}.$$

Proposition 3.129. Let U be an open set of $\mathbb{R}^3$ and $f : U \to \mathbb{R}$, $\boldsymbol{f} : U \to \mathbb{R}^3$ be functions of class $\mathcal{C}^2(U)$. Then for all $x \in U$ we have:

$$\operatorname{rot}(\nabla f) = 0, \qquad \operatorname{div}(\operatorname{rot} \boldsymbol{f}) = 0 \quad \text{i} \quad \operatorname{div}(\nabla f) = \nabla^2 f.$$

Surface area and surface integrals

Proposition 3.130. Let S be the graph of a function $z = \Phi(x,y)$ of class $\mathcal{C}^1(U)$, $U \subseteq \mathbb{R}^2$. Then

$$\operatorname{area}(S) = \iint_U \sqrt{1 + \left(\frac{\partial \Phi}{\partial x}\right)^2 + \left(\frac{\partial \Phi}{\partial y}\right)^2}\, dxdy.$$

Definition 3.131. A *parametritzed surface* $S \subset \mathbb{R}^3$ is the image of a map $\Phi : U \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ of class $\mathcal{C}^1(U)$ defined by $\Phi(u,v) = (x(u,v), y(u,v), z(u,v))$.

Proposition 3.132. Let $S = \Phi(U)$ be a surface in $\mathbb{R}^3$ parametritzed by $\Phi \in \mathcal{C}^1(U)$. Then the unit normal vector to S at the point $\Phi(u,v)$ is

$$\mathbf{n}(u,v) = \frac{\frac{\partial \Phi}{\partial u} \wedge \frac{\partial \Phi}{\partial v}}{\left\| \frac{\partial \Phi}{\partial u} \wedge \frac{\partial \Phi}{\partial v} \right\|}.$$

Proposition 3.133. Let $S = \Phi(U)$ be a surface in $\mathbb{R}^3$ parametritzed by $\Phi \in \mathcal{C}^1(U)$. Then,

$$\operatorname{area}(S) = \iint_U \left\| \frac{\partial \Phi}{\partial u} \wedge \frac{\partial \Phi}{\partial v} \right\| dudv.$$

Definition 3.134. Let $S = \Phi(U)$ be a surface in $\mathbb{R}^3$ parametritzed by $\Phi \in \mathcal{C}^1(U)$ and $f : \mathbb{R}^3 \to \mathbb{R}$ be a continuous function whose domain contain S. We define the *surface integral* f *over* S as

$$\iint_S f dS = \iint_U f(\Phi(u,v)) \left\| \frac{\partial \Phi}{\partial u} \wedge \frac{\partial \Phi}{\partial v} \right\| dudv$$[26].

Definition 3.135. Let $S = \Phi(U)$ be a surface in $\mathbb{R}^3$ parametritzed by $\Phi \in \mathcal{C}^1(U)$ and $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3$ be a continuous vector field whose domain contain S. We define the *surface integral* $\boldsymbol{f}$ *over* S or the *flux of* $\boldsymbol{f}$ *across* S as

$$\iint_S \boldsymbol{f} \cdot d\mathbf{S} = \iint_S \boldsymbol{f} \cdot \mathbf{n} dS = \iint_U \boldsymbol{f}(\Phi(u,v)) \cdot \left(\frac{\partial \Phi}{\partial u} \wedge \frac{\partial \Phi}{\partial v} \right) dudv,$$

where $\mathbf{n}$ is the unit normal vector to S[27].

[25] It can be seen that the latter integral is independent of the parametrization of c except for a factor of -1 that depends on the orientation of the parametrization.

[26] It can be seen that this integral is independent of the parametrization of S.

[27] It can be seen that the latter integral is independent of the parametrization of S except for a factor of -1 that depends on the orientation of the normal vector $\mathbf{n}$.

Theorems of vector calculus on $\mathbb{R}^2$

Definition 3.136. Let $U \subseteq \mathbb{R}^3$ be an open set. A *differential 1-form on U* is an expression of the form

$$\omega = F_1 dx + F_2 dy + F_3 dz,$$

where F_1, F_2, F_3 are scalar functions defined on U[28].

Theorem 3.137 (Green's theorem). Let $\boldsymbol{f} = (f_1, f_2)$ be a vector field of class $\mathcal{C}^1(U)$, $U \subseteq \mathbb{R}^2$, and $c = \partial U$ be the curve formed from the boundary of U[29]. Then

$$\int_{\partial U} \boldsymbol{f} \cdot d\mathbf{s} = \iint_U \operatorname{rot} \boldsymbol{f} dxdy$$[30].

Corollary 3.138. Let U be a region in $\mathbb{R}^2$ and ∂U be its boundary. Then,

$$\operatorname{area}(U) = \int_{\partial U} xdy = -\int_{\partial U} ydx = \frac{1}{2}\int_{\partial U} (xdy - ydx).$$

Theorem 3.139 (Divergence theorem on $\mathbb{R}^2$). Let $\boldsymbol{f} = (f_1, f_2)$ be a vector field of class $\mathcal{C}^1(U)$, $U \subseteq \mathbb{R}^2$ with boundary ∂U. Then,

$$\int_{\partial U} \boldsymbol{f} \cdot \mathbf{n} ds = \iint_U \operatorname{div} \boldsymbol{f} dxdy$$[31].

Theorems of vector calculus on $\mathbb{R}^3$

Theorem 3.140 (Stokes' theorem). Let S be a parametritzed surface of class $\mathcal{C}^1$ and ∂S be its boundary. Let $\boldsymbol{f} = (f_1, f_2, f_3)$ be a vector field of class $\mathcal{C}^1$ in a domain containing $S \cup \partial S$. Then

$$\int_{\partial S} \boldsymbol{f} \cdot d\mathbf{s} = \iint_S \operatorname{rot} \boldsymbol{f} \cdot \mathbf{n} dS.$$

Corollary 3.141. Let $a \in \mathbb{R}^3$ and $\mathbf{n}$ be a unit vector. Suppose $D_r = D(a, r)$ is a disk of radius r centered at a and perpendicular to $\mathbf{n}$. Let $\boldsymbol{f}$ be a vector field of class $\mathcal{C}^1(D_r)$. Then

$$\operatorname{rot} \boldsymbol{f}(a) \cdot \mathbf{n} = \lim_{r \to 0} \frac{1}{\operatorname{area}(D_r)} \int_{\partial D_r} \boldsymbol{f} \cdot d\mathbf{s}.$$

Therefore, the $\mathbf{n}$-th component of $\operatorname{rot} \boldsymbol{f}(a)$ is the circulation of $\boldsymbol{f}$ in a small circular surface perpendicular to $\mathbf{n}$, per unit of area.

Theorem 3.142 (Gauß' or divergence theorem on $\mathbb{R}^3$). Let $\boldsymbol{f} = (f_1, f_2, f_3)$ be a vector field of class $\mathcal{C}^1$ on a symmetric region[32] $V \subset \mathbb{R}^3$ with boundary ∂V. Then,

$$\iint_{\partial V} \boldsymbol{f} \cdot \mathbf{n} dS = \iiint_V \operatorname{div} \boldsymbol{f} dxdydz$$

[28] Extending this notion, we can define 2-forms and 3-forms as:

$$\begin{aligned} \omega = F_1 dxdy + F_2 dydz + F_3 dzdy \quad & \text{2-form,} \\ \omega = F dxdydz \quad & \text{3-form.} \end{aligned}$$

[29] It goes without saying that the orientation is chosen positive, that is counterclockwise.

[30] Alternatively, using differential forms, we get

$$\int_{\partial U} (F_1 dx + F_2 dy) = \iint_U \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dxdy.$$

[31] The first integral represents the flux of $\boldsymbol{f}$ across the curve ∂U.

[32] A region on $\mathbb{R}^3$ is *symmetric* if is xy-simple, yz-simple and xz-simple.

Corollary 3.143. Let $B_r = B(a, r)$ be a ball of radius r centered at $a \in \mathbb{R}^3$ and $\boldsymbol{f}$ be a vector field of class $\mathcal{C}^1(B_r)$. Then

$$\operatorname{div} \boldsymbol{f}(a) = \lim_{r \to 0} \frac{1}{\operatorname{vol}(B_r)} \iint_{\partial B_r} \boldsymbol{f} \cdot \mathbf{n} dS.$$

Therefore, $\operatorname{div} \boldsymbol{f}(a)$ is the flux of $\boldsymbol{f}$ outward form a, in the normal direction across the surface of a small ball centered on a, per unit of volume.

2.4 Linear geometry

2.4.1 | The foundations of geometry

In this section we will only study geometry in the plane.

Euclidean geometry

Axiom 4.1 (Euclid's axioms).

1. It is possible to draw, from any point to any point, a straight line.
2. It is possible to extend any segment by either of its two ends.
3. With center at any point it is possible to draw a circle that passes through any other point.
4. All right angles are equal.
5. If a line segment intersects two straight lines forming two interior angles on the same side that sum to less than two right angles, then the two lines, if extended indefinitely, meet on the side on which the angles sum to less than two right angles.

5′. *(Playfair's axiom)* Given a line and a point not on it, at most one line parallel to the given line can be drawn through the point.

Hilbert's axioms

Definition 4.2. In elementary plane geometry[33], there are two types of objects, *points* and *lines*, which can have three types of relationships between them:

- An *incidence relation*. We say, for example, that a point lies on a line or a line passes through a point.
- An *order relation*. We say, for example, that a point lies between two other points.
- A *congruence relation*. We say, for example, that a segment is congruent to another or an angle is congruent to another[34].

Axiom 4.3 (Incidence axioms).

1. For every two points there exists no more than one line containing both.
2. There exist at least two points on a line.
3. There exist at least three points that do not lie on the same line.

Axiom 4.4 (Order axioms).

1. If a point B lies between A and C, then B lies between C and A and there exists a line containing the distinct points A, B, C.
2. If A and B are two points, there exists at least one point C such that B lies between A and C.
3. Given three point on a line, there is no more than one which lies between the other two.
4. *(Pasch's axiom)* Let A, B, C be three points not-lying in the same line and let r be a line not passing through any of the points A, B, C and passing through a point of the segment AB. Then it also passes through either a point of the segment BC or a point of the segment AC.

Definition 4.5. A *ray* or *half-line* is a point A, called vertex, and all the points of a line passing through A lying on the same side with respect to A.

Definition 4.6. A *half-plane* is a straight line r and all the points lying on the same side with respect to r.

Definition 4.7. An *angle* is a non-ordered pair of rays with same vertices that belong to different straight lines.

Axiom 4.8 (Congruence axioms).

1. Congruence of angles and congruence of rays are equivalence relations.
2. Let a and b be two lines not necessarily different, A and B be points on a and A' be a point on b. We fix a side of the line b with respect to A'. Then, there exists a point B' lying on this side of b such that $AB \equiv A'B'$.
3. Let a, a' be two lines not necessarily different. Let AB, BC be segments on a that intersect only in one point and $A'B', B'C'$ be segments on a' that also intersect only in one point. If $AB \equiv A'B'$ and $BC \equiv B'C'$, then $AC \equiv A'C'$.
4. Let $\angle hk$ be an angle, k' be a ray and H be one of the two half-planes that k' defines. Then, there is one and only one angle $\angle h'k'$ such that $\angle hk \equiv \angle h'k'$ and h' belongs to H.
5. *(SAS criterion)* Consider two triangles[35] ABC and $A'B'C'$ (not necessarily different). If $AC \equiv A'C'$, $AB \equiv A'B'$ and $\alpha \equiv \alpha'$, then $\beta \equiv \beta'$.

Axiom 4.9 (Continuity axioms).

1. *(Axiom of Archimedes)* If AB and CD are any segments, then there exists a number n such that n segments CD constructed contiguously from A, along the ray from A to B, will pass beyond the point B.

[33] In this section we only study the geometry in the plane.
[34] We will use the notation $\equiv$ to say that two angles or segments are congruent.
[35] We will use the following notation with respect to the angles of a triangle ABC: $\alpha = \angle CAB$, $\beta = \angle ABC$ and $\gamma = \angle BCA$.

2. *(Axiom of completeness)* An extension of a set of points on a line with order and congruence relations that would preserve the relations existing among the original elements as well as the rest of the axioms is impossible.
3. *(RC)* If a straight line passes through a point inside a circle, it intersects the circle in two points.
4. *(CC)* If a circle passes through points inside and outside another circle, the two circle intersect in two points.

Axiom 4.10 (Axiom of Parallels). Let a be any line and A be a point not on it. Then there is at most one line that passes through A and does not intersect a.

Definition 4.11. Different types of geometry:

- A *Hilbert plane* is a geometry where axioms 4.3, 4.4 and 4.8 are satisfied.
- A *Pythagorean plane* is a Hilbert plane in which axiom of Parallels is satisfied.
- An *Euclidean plane* is a Pythagorean plane in which axioms RC and CC are satisfied.
- The *Cartesian geometry of* $\mathbb{R}^2$ is the unique geometry satisfying all Hilbert's axioms.

Absolute geometry

Definition 4.12. *Absolute geometry* is the part of Euclidean geometry that only uses axioms 4.3, 4.4 and 4.8.

Theorem 4.13. In an isosceles triangles, the angles opposite the congruent sides are congruent.

Theorem 4.14 (SAS criterion). If two sides of a triangle and the angle between them are congruent to the corresponding sides and angle of a second triangle, then the two triangles are congruent.

Theorem 4.15. Adjacent angles of congruent angles are congruent.

Theorem 4.16. Opposite angles[36] are congruent.

Theorem 4.17. If A and B are each on one of the sides of an angle with vertex O, any ray with vertex O that passes through an interior point of the angle intersects the segment AB.

Theorem 4.18. There exist right angles.

Theorem 4.19. Let $\alpha, \alpha', \beta, \beta'$ be angles. If $\alpha \equiv \alpha'$ and $\beta \equiv \beta'$, then $\alpha + \beta \equiv \alpha' + \beta'$.

Theorem 4.20 (SSS criterion). If two triangles have all its sides congruent, they have all its angles congruent.

Theorem 4.21. Right angles are congruent.

Theorem 4.22 (Exterior angle theorem). An exterior angle of a triangle is greater than any of the non-adjacent interior angles.

Theorem 4.23. If ℓ is a line and P is a point not lying on ℓ, there exists a line L passing through P and such that not intersects ℓ.

Theorem 4.24 (ASA criterion). If two triangles have a side and the two angles of this side congruent, the triangles are congruent.

Theorem 4.25 (SAA criterion). If two triangles have a side, an angle of this side and the angle opposite to this side congruent, the triangles are congruent.

Theorem 4.26. In any triangle the greater side is opposite to the greater angle.

Theorem 4.27. If a triangle has two congruent angles, it is isosceles.

Theorem 4.28. Every segment has a midpoint.

Theorem 4.29. Every angle has an angle bisector.

Theorem 4.30. Every segment has a perpendicular bisector.

Theorem 4.31 (Saccheri–Legendre theorem). The sum of the angles of a triangle is at most two right angles.

Cartesian geometry

Definition 4.32. An *ordered field* k is a field together with a total order of its elements, satisfying:

- $x \leq y \implies x + z \leq y + z \ \forall z \in k$.
- $x, y \geq 0 \implies xy \geq 0$.

Definition 4.33. We say a field k is *Pythagorean* if $\forall a \in k$, $1 + a^2 = b^2$ for some $b \in k$.

Theorem 4.34. k^2 is a Pythagorean plane if and only if k is an ordered Pythagorean field.

Definition 4.35. An ordered field k is *Archimedean* if axiom of Archimedes is valid in k.

Definition 4.36. An ordered field k is *Euclidean* if $\forall a \in k$, $a > 0$, there exists a $b \in k$ such that $b^2 = a$.

Theorem 4.37. k^2 is a Euclidean plane if and only if k is an ordered Euclidean field.

[36] Opposite angles are angles that are opposite each other when two lines intersect.

Definition 4.38. The smallest Pythagorean field is called *Hilbert field* (Ω) and it can be defined as the intersection of all Pythagorean fields of $\mathbb{R}$. Alternatively, it can be defined as the field whose elements are the real numbers obtained from rational numbers with the operations of addition, subtraction, multiplication, multiplicative inverse and the operation $a \mapsto \sqrt{1+a^2}$.

Definition 4.39. The smallest Euclidean field is called *constructible field* ($\mathbb{K}$) and it can be defined as the intersection of all Euclidean fields of $\mathbb{R}$. Alternatively, it can be defined as the field whose elements are the real numbers obtained from rational numbers with the operations of addition, subtraction, multiplication, multiplicative inverse and the square root of positive numbers.

Non-Euclidean geometries

Definition 4.40 (Hyperbolic geometry). *Hyperbolic geometry* is the non-Euclidean geometry where axiom of Parallels fails.

Proposition 4.41. Properties of hyperbolic geometry:

- There are infinity lines parallel to a given line ℓ that pass through a point not lying on ℓ.
- There are lines inside an angle that do not intersect the sides of the angle.
- The sum of the angles of any triangle is less than two right angles.

Definition 4.42. Hyperbolic geometry models:

- Beltrami-Klein model:
  - Points: $\mathcal{K} := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$.
  - Lines: Lines of $\mathbb{R}^2$ that intersect with $\mathcal{K}$.
  - Incidence and order relations are the same as in ordinary Euclidean geometry of $\mathbb{R}^2$.
  - Two segments $AB, A'B' \in \mathcal{K}$ are congruent if and only if there is an Euclidean motion[37] f such that $f(A) = A'$ and $f(B) = B'$. Two angles $hk, h'k' \in \mathcal{K}$ are congruent if and only if there is an Euclidean motion f such that $f(h) = h'$ and $f(k) = k'$.

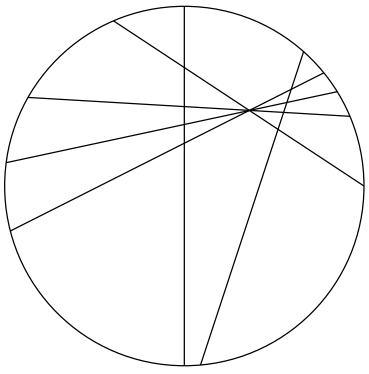

Figure 2.3: Beltrami–Klein model

- Poincaré disk model:
  - Points: $\mathcal{D} := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$.
  - Lines:
    1. Lines of $\mathbb{R}^2$ that pass through the origin.
    2. Circles of $\mathbb{R}^2$ that intersect orthogonally the circle $\mathcal{C} = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$.
  - Incidence and order relations are the same as in ordinary Euclidean geometry of $\mathbb{R}^2$.
  - Is a conformal model: The hyperbolic measure of an angle coincides with the Euclidean measure of it whereas the distance between two points $A, B \in \mathcal{D}$ is measured using the following formula:

$$d_h(A,B) := -\log \frac{d(A,P)d(B,Q)}{d(A,Q)d(B,P)}.$$

    where $P, Q \in \mathcal{C}$ are the boundary points of $\mathcal{D}$ on the line passing through A and B so that A lies between P and B.

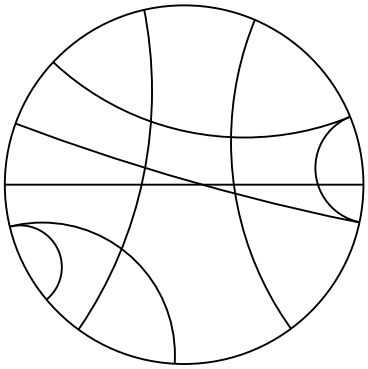

Figure 2.4: Poincare disk model

- Poincaré half-plane model:
  - Points: $\mathcal{H} := \{(x,y) \in \mathbb{R}^2 : y > 0\}$.
  - Lines:
    1. Vertical straight lines of $\mathbb{R}^2$.
    2. Circles of $\mathbb{R}^2$ with center on the x-axis.
  - Incidence and order relations are the same as in ordinary Euclidean geometry of $\mathbb{R}^2$.
  - Is a conformal model. The distance between two points $A, B \in \mathcal{D}$ is measured using the following formula:

$$d_h(A,B) := -\log \frac{d(A,P)d(B,Q)}{d(A,Q)d(B,P)}.$$

    where $P, Q \in \{(x,y) \in \mathbb{R}^2 : y = 0\}$ are the points where the semicircle meet the boundary line $y = 0$.

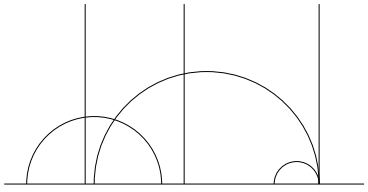

Figure 2.5: Poincare half-plane model

[37] See section 2.4.3.

Definition 4.43 (Non-Paschian geometry). *Non-Paschian geometry* is the non-Euclidean geometry where axiom of Archimedes fails.

Proposition 4.44 (Construction of a non-Paschian geometry). Suppose we have a total order relation $\trianglelefteq$ on $\mathbb{R}$ such that:

1. $x \trianglelefteq y \implies x + z \trianglelefteq y + z \ \forall x, y, z \in \mathbb{R}$.
2. $\exists a, b \in \mathbb{R} : a \trianglerighteq 0, b \trianglerighteq 1$ and $ab \trianglelefteq 0$.

Then, the ordinary affine geometry of $\mathbb{R}^2$ together with $\trianglelefteq$, satisfy all Hilbert's axioms except Pasch's axiom.

Definition 4.45 (Non-SAS geometry). *Non-SAS geometry* is the non-Euclidean geometry where SAS criterion fails.

Proposition 4.46 (Construction of a non-SAS geometry).

- Points: $\mathcal{S} = \{(x, y, z) \in \mathbb{R}^3 : x + z = 0\} = \{(x, y, -x) \in \mathbb{R}^3\}$.
- Lines: Ordinary straight lines of $\mathbb{R}^2$ contained in $\mathcal{S}$.
- Incidence and order relations are the same as in ordinary Euclidean geometry of $\mathbb{R}^2$.
- Congruence of angles is the same as in the ordinary geometry of $\mathbb{R}^3$. Congruence of segments is based in the following distance:

  $$d'((x, y, -x), (x', y' - x'))^2 = (x - x')^2 + (y - y')^2.$$

  That is, two segments are congruent if so are their projections to the plane $z = 0$.

Definition 4.47 (Non-Archimedean geometry). *Non-Archimedean geometry* is the non-Euclidean geometry where SAS criterion fails.

Axiomatic projective space

Definition 4.48. An *axiomatic projective space* is a system of points and lines with an incidence relation that satisfy:

1. Every line contains at least 3 points
2. Any two distinct points lie on a unique line.
3. *(Projective axiom)* If A, B, C, D are four different points and lines AB and CD intersect, then lines AC and BD also intersect.

Definition 4.49. Let X be a projective space. A *projective subvariety of* X is a set $Z \neq \emptyset$ of points of X such that if $x, y \in Z$ are different points, then all the points lying on the line passing through x and y belong to Z. Thus, Z is also a projective space.

Proposition 4.50. Let X be a projective space. The intersection of subvarieties of X is also a subvariety of X.

Proposition 4.51. If A and B are subvarieties of a projective space X, we define its sum $A+B$ as the intersection of all subvarieties containing $A \cup B$. As a consequence, $A + B$ is a subvariety of X.

Definition 4.52. Let X, Y be a projective spaces. A *collineation between X and Y* is a bijection map $f : X \to Y$ such that $A, B, C \in X$ are three collinear points if and only if $f(A), f(B), f(C) \in Y$ are also collinear.

Definition 4.53. If X is a projective space, the *dimension of X* is the maximum n such that there is a chain of inclusions

$$X_0 \subsetneq X_1 \subsetneq X_2 \subsetneq \cdots \subsetneq X_n,$$

where each X_i is a non-empty subvariety of X. If this n doesn't exist, we say X has infinite dimension.

Definition 4.54. A *projective plane* is a projective space of dimension 2 that satisfies the following axioms:

1. Any two distinct points lie on a unique line.
2. Any two distinct lines meet on a unique point.
3. There exist at least four points of which no three are collinear.

Theorem 4.55. X is a projective space of dimension 2 if and only if X satisfies axioms 4.54.

Theorem 4.56 (Duality principle). If a statement $\mathcal{P}$ (which only involves points and lines) is true in any projective plane, then the statement obtained from $\mathcal{P}$ exchanging points by lines (and correctly changing all the connectors to make a consistent statement) is also true in any projective plane.

Affine and projective spaces

Definition 4.57. An *affine plane* is a set of points and lines satisfying the following axioms:

1. Any two distinct points lie on a unique line.
2. If r is a line and $P \notin r$ is a point, there exists a unique line s such that $P \in s$ and r and s does not intersect.
3. Any line has at least two distinct points.
4. There exist at least two distinct lines.

Proposition 4.58 (Passage from the projective plane to the affine plane). Suppose X is a projective plane and $r \in X$ is an arbitrary line of X. Let $\mathbb{A} := X - r$. Then, $\mathbb{A}$ is an affine plane.

Proposition 4.59 (Passage from the affine plane to the projective plane). Suppose $\mathbb{A}$ is an affine plane. Let $\mathcal{R}$ be the set of all lines of $\mathbb{A}$. We define

$$L = \mathcal{R}\Big/_{\sim} \quad \text{where } r \sim s \iff r \parallel s.$$

Construction of a projective plane X:

1. The points of X are the points of $\mathbb{A}$ and L.
2. The lines of X are the lines of $\mathbb{A}$ an another line ℓ.
3. Incidence relation on X: Let $P \in X$ be a point and $r \in X$ a line. Then:
   - If $P \in \mathbb{A}$ and $r \in \mathbb{A}$, then $P \in r$ has the same meaning on X and A.
   - If $P \in \mathbb{A}$ and $r = \ell$, then $P \notin r$.
   - If $P \in X \setminus \mathbb{A} = L$, then $P \in \ell$.
   - If $P \in X \setminus \mathbb{A} \neq L$, then P is an equivalence class of lines of $\mathbb{A}$ and, if $r \in \mathbb{A}$, we say $P \in r$ if $r \in X$.

2.4.2 | Projective geometry

Projective space

Definition 4.60. Let V be a $n+1$-dimensional vector space over a field k. We define the *n-dimensional projective space $\mathcal{P}(V)$ of V* in either of these two equivalent ways:

- $\mathcal{P}(V) := \{\text{1-dimensional vector subspaces of } V\}$.
- $\mathcal{P}(V) := (V \setminus \{0\})/\sim$ where the relation $\sim$ is defined as $v \sim u \iff v = \lambda u$, $\lambda \neq 0$[38].

Definition 4.61. Let V, W be two vector spaces over a field k and $\mathcal{P}(V), \mathcal{P}(W)$ be their associated projective spaces. If $\phi : V \to W$ is an isomorphism, we can consider the map:

$$\begin{aligned} \mathcal{P}(\phi) : \mathcal{P}(V) &\to \mathcal{P}(W) \\ [v] &\mapsto [\phi(v)] \end{aligned}$$

We say $\mathcal{P}(\phi)$ is an *homography between $\mathcal{P}(V)$ and $\mathcal{P}(W)$*.

Definition 4.62. Let V be a vector space over a field k and W be a vector space over a field k'. An *semilinear isomorphism* $\phi : V \to W$ is a bijective map associated with a field isomorphism $r : k \to k'$ such that

$$\begin{aligned} &\phi(u+v) = \phi(u) + \phi(v) \quad \forall u, v \in V. \\ &\phi(\lambda v) = r(\lambda)\phi(v) \quad \forall v \in V,\ \forall \lambda \in k. \end{aligned}$$

Definition 4.63. Let V be a vector space over a field k, W be a vector space over a field k' and $\phi : V \to W$ a semilinear isomorphism. We say $\mathcal{P}(\phi) : \mathcal{P}(V) \to \mathcal{P}(W)$ is an *isomorphism between projective spaces* and we write $\mathcal{P}(V) \cong \mathcal{P}(W)$ to denote that $\mathcal{P}(V), \mathcal{P}(W)$ are isomorphic.

Proposition 4.64. Let V be a $n+1$-dimensional vector space over a field k. Then there is an homography $\mathcal{P}(V) \cong \mathcal{P}(k^{n+1})$[39].

Definition 4.65. Let V be a $n+1$-dimensional vector space over a field k and $E \subseteq V$ be a $m+1$-dimensional vector subspace. Consider the natural inclusion $\mathcal{P}(E) \subseteq \mathcal{P}(V)$. We say $\mathcal{P}(E)$ is a *m-dimensional projective subvariety of $\mathcal{P}(V)$*. In particular, we call *line of $\mathcal{P}(V)$* any 1-dimensional projective subvariety and we call *hyperplane of $\mathcal{P}(V)$* any $n-1$-dimensional projective subvariety.

Homogeneous coordinates and Graßmann formula

Definition 4.66. Let V be a $n+1$-dimensional vector space over a field k, $(v_0, \dots, v_n)$ be a basis of V and $\mathcal{P}(V)$ be a projective space. Given $x \in \mathcal{P}(V)$ such that $x = [v]$ for some $v \in V$, $v = \lambda_0 v_0 + \cdots + \lambda_n v_n$, we define the *homogeneous coordinates of x* as

$$x = \{\lambda_0, \dots, \lambda_n\}.$$

Definition 4.67. Let $\mathcal{P}(V)$ be a n-dimensional projective space. A *projective frame on $\mathcal{P}(V)$* is a tuple of $n+2$ points of $\mathcal{P}(V)$, such that any $n+1$ points of the tuple are not contained in a hyperplane.

Theorem 4.68. Let $\mathcal{P}(V)$ be a n-dimensional projective space. If $U_0, \dots, U_n, U$ is a projective frame of $\mathcal{P}(V)$, there exists a basis $(v_0, \dots, v_n)$ of V such that

$$U_i = [v_i] \text{ for } i = 0, \dots, n \text{ and } U = [v_1 + \cdots + v_n].$$

If $(u_0, \dots, u_n)$ is another basis of V that satisfies the same property, then $\exists \tau \neq 0 : u_i = \tau v_i$, for $i = 0, \dots, n$.

Definition 4.69. Let $\mathcal{P}(V)$ be a n-dimensional projective space and let $H \subset \mathcal{P}(V)$ be a hyperplane. The *equation of the hyperplane* is

$$x_0 a_0 + \ldots + x_n a_n = 0.$$

Definition 4.70. Let $\mathcal{P}(V)$ be a projective space and let $Y_1 = \mathcal{P}(E_1)$ and $Y_2 = \mathcal{P}(E_2)$ be two projective subvarieties of $\mathcal{P}(V)$. Then

- $Y_1 \cap Y_2 = \mathcal{P}(E_1 \cap E_2)$.
- $Y_1 + Y_2 = \mathcal{P}(E_1 + E_2)$.

Theorem 4.71 (Graßmann formula). Let $\mathcal{P}(V)$ be a projective space and $Y_1 = \mathcal{P}(E_1)$, $Y_2 = \mathcal{P}(E_2)$ be two projective subvarieties of $\mathcal{P}(V)$. Then:

$$\dim(Y_1 \cap Y_2) + \dim(Y_1 + Y_2) = \dim Y_1 + \dim Y_2$$[40].

[38] Observe that $\sim$ is an equivalence relation.
[39] From now on we will use the notation $P_n(k) := \mathcal{P}(k^{n+1})$.
[40] The formula is also valid for the case $Y_1 \cap Y_2 = \emptyset$ if we consider, by agreement, $\dim \emptyset := -1$.

Fano and Pappus configurations

Definition 4.72. A *configuration* is a finite set of points and lines satisfying the following axioms:

1. There are four points such that no three of them are collinear.

2. Two distinct points lie on at most one line.

Definition 4.73. Let X be a projective geometry and $\mathcal{C}$ be a configuration. We say $\mathcal{C} \subseteq X$ if there exists injective maps i_p, i_ℓ from the points and lines of $\mathcal{C}$ to the points and lines of X, respectively, such that if A is a point and s is a line satisfying $A \in s$, then $i_p(A) \in i_\ell(s)$.

Definition 4.74. Let X be a projective geometry and $\mathcal{C}$ be a configuration. We say $\mathcal{C}$ is *realizable on* X if there is an inclusion $\mathcal{C} \subseteq X$.

Definition 4.75. Let X be a projective geometry and $\mathcal{C}$ be a configuration. We say $\mathcal{C}$ is a *theorem in* X if satisfies that for any line $r \in \mathcal{C}$, the inclusion $\mathcal{C} - r \subseteq X$ can be extended to an inclusion $\mathcal{C} \subseteq X$.

Definition 4.76. *Fano configuration* is a configuration of 7 points and 7 lines defined in either of the following ways:

- It's the configuration described in figure 2.6.

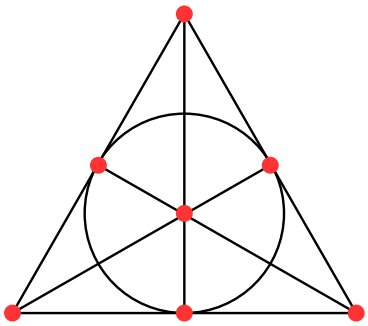

Figure 2.6: Fano configuration

- It's the unique projective plane of order 2[41].

- It's the projective plane $P_2(\mathbb{F}_2)$.

Theorem 4.77. If $n \geq 2$, Fano configuration is a theorem in $P_n(k)$ if and only if $\operatorname{char} k = 2$.

Definition 4.78. *Pappus configuration* is a configuration of 9 points and 9 lines defined in either of the following ways:

- It's the configuration described in figure 2.7.

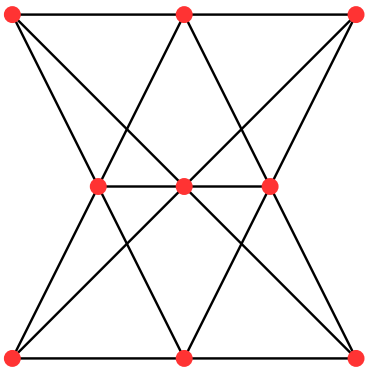

Figure 2.7: Pappus configuration

- It's the configuration whose points are the elements of the group $(\mathbb{Z}/9\mathbb{Z}, +)$ and whose lines are triples $\{i, j, k\}$ such that $i + j + k = 0$ where i, j, k are different modulo 3.

- It's the configuration obtained from the affine plane over $\mathbb{F}_3$ eliminating three parallel lines.

Theorem 4.79. Let k be a division ring. Pappus configuration is a theorem in $P_n(k)$ if and only if k is a field.

Desargues configuration

Definition 4.80. Two triangles ABC and $A'B'C'$ are said to be in *perspective with respect to a point* if lines AA', BB' and CC' intersect at the point P. This point is called *centre of perspectivity.*

Definition 4.81. Two triangles ABC and $A'B'C'$ of sides a, b, c and a', b', c' respectively are said to be in *perspective with respect to a line* if points $a \cap a'$, $b \cap b'$ and $c \cap c'$ lie on the same line r. This line is called *axis of perspectivity.*

Theorem 4.82 (Desargues' theorem). If two triangles are in perspective with respect to a point, so are in perspective with respect to a line[42].

Definition 4.83. *Desargues configuration* is a configuration of 10 points and 10 lines defined in either of the following ways:

- It's the configuration described in figure 2.8.

[41] A finite projective plane of order n is a projective plane in which every line has $n + 1$ points and every point lies on $n + 1$ lines.

[42] Desargues' theorem is valid in any axiomatic projective space of dimension 3 and, generally, in any axiomatic projective space that is a subvariety of an axiomatic projective space of dimension 3. In particular, it is valid in $P_n(k)$ for any division ring k and $n \geq 2$.

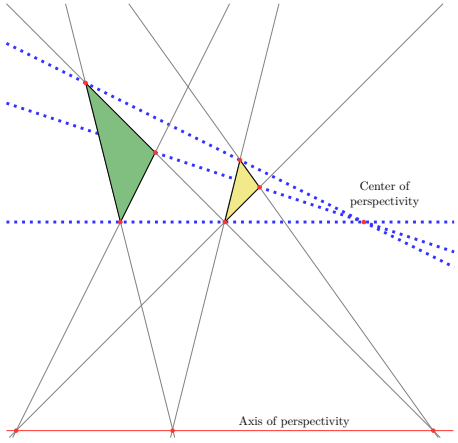

Figure 2.8: Desargues configuration

- It's the configuration whose points are the elements of the set $S = \{1, 2, 3, 4, 5\}$ and whose lines are the subsets of cardinal 3 of S.
- It's the configuration created from two triangles that are simultaneously in perspective with respect to a point and in perspective with respect to a line.

Definition 4.84. Projective planes in which Desargues' theorem is not satisfied are called *non-Desarguesian planes*.

Theorem 4.85 (Coordination theorem). Let X be an axiomatic projective space of finite dimension $n > 1$ where Pappus' theorem is valid. Then there exist a field k and an isomorphism $X \cong P_n(k)$[43].

Fundamental theorem of projective geometry and cross ratio

Theorem 4.86 (Fundamental theorem of projective geometry). Let $f : \mathcal{P}(V) \to \mathcal{P}(W)$ be a collineation between projective spaces of finite dimension greater than 1. Then, there exists a semilinear isomorphism $\phi : V \to W$ such that $f = P(\phi)$.

Definition 4.87 (Cross ratio). Let $A, B, C, D \in \mathcal{P}(V)$ be four collinear points lying on a line $L \in \mathcal{P}(V)$ with A, B, C different. As we have $A = [v_1]$, $B = [v_2]$, $C = [v_3]$ and $D = [v_4]$ for some vectors $v_1, v_2, v_3, v_4 \in V$ then $L = \langle v_1, v_2 \rangle$. Therefore, $v_3 = \lambda_1 v_1 + \lambda_2 v_2$ and $v_4 = \mu_1 v_1 + \mu_2 v_2$, for some $\lambda_1, \lambda_2, \mu_1, \mu_2 \in k$. We define the *cross ratio between* A, B, C, D as

$$(A, B, C, D) := \begin{cases} \dfrac{\lambda_2 \mu_1}{\lambda_1 \mu_2} & \text{si} \quad \lambda_1 \mu_2 \neq 0, \\ \infty & \text{si} \quad \lambda_1 \mu_2 = 0. \end{cases}$$

Definition 4.88. Let $A, B, C, D \in \mathcal{P}(V)$ be four collinear points. If $(A, B, C, D) = -1$ we say the points A, B, C, D form an *harmonic ratio*.

Definition 4.89. Let a, b, c, d be four lines on a plane (with a, b, c different) intersecting at the point P. Let r be a different line such that $P \notin r$ and let $A := a \cap r$, $B := b \cap r$, $C := c \cap r$, $D := d \cap r$. We define the *cross ratio between* a, b, c, d as

$$(a, b, c, d) := (A, B, C, D).$$

Definition 4.90. Let X be a projective space such that $\dim X \geq 2$. Let $L_1, L_2 \in X$ be two lines intersecting at the point $P \in X$ and $f : L_1 \to L_2$ be a function such that $f(A) = L_2 \cap PA$. We say f is a *prespectivity*. The composition of prespectivities is called a *projectivity*.

Theorem 4.91. If $f : L_1 \to L_2$ is a projectivity, then f preserves cross ratio, that is:

$$(f(A), f(B), f(C), f(D)) = (A, B, C, D).$$

Theorem 4.92. Let V be a 2-dimensional vector space and $f : \mathcal{P}(V) \to \mathcal{P}(V)$ be a bijection. There exists linear map $\phi : V \to V$ such that $f = P(\phi)$ if and only if f preserves cross ratio.

Plücker coordinates

Proposition 4.93. Let $r \in \mathcal{P}_3(k)$ be a line and $A, B \in r$ two points with coordinates $A = \{a_0, a_1, a_2, a_3\}$ and $B = \{b_0, b_1, b_2, b_3\}$. Consider the matrix:

$$A = \begin{pmatrix} a_0 & a_1 & a_2 & a_3 \\ b_0 & b_1 & b_2 & b_3 \end{pmatrix}$$

Now consider the six minors of A:

$$p_{01} = \begin{vmatrix} a_0 & a_1 \\ b_0 & b_1 \end{vmatrix}, \quad p_{02} = \begin{vmatrix} a_0 & a_2 \\ b_0 & b_2 \end{vmatrix}, \quad p_{03} = \begin{vmatrix} a_0 & a_3 \\ b_0 & b_3 \end{vmatrix},$$
$$p_{23} = \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix}, \quad p_{31} = \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix}, \quad p_{12} = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix}.$$

The coordinates $\{p_{01}, p_{02}, p_{03}, p_{23}, p_{31}, p_{12}\}$ doesn't depend on the points A, B on the line r. We define the *Plücker coordinates of* r as the coordinates $\{p_{01}, p_{02}, p_{03}, p_{23}, p_{31}, p_{12}\}$

Proposition 4.94. Two lines are equal if and only if they have the same Plücker coordinates.

Proposition 4.95. Let r be a line with Plücker coordinates $\{p_{01}, p_{02}, p_{03}, p_{23}, p_{31}, p_{12}\}$. Then the points $x = \{x_0, x_1, x_2, x_3\} \in r$ satisfy

$$\begin{pmatrix} p_{12} & -p_{02} & p_{01} & 0 \\ -p_{31} & -p_{03} & 0 & p_{01} \\ p_{23} & 0 & -p_{03} & p_{02} \\ 0 & p_{23} & p_{31} & p_{12} \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

[43] If Pappus theorem is not valid but Desargues' theorem is, then $X \cong P_n(k)$ for some division ring k.

2.4.3 | Affine geometry

Affine space

Definition 4.96. Let V be a vector space over a field k. An *affine space over* V is a set $\mathbb{A}$ together with a map:

$$\begin{aligned}\mathbb{A} \times V &\to \mathbb{A}\\ (P, \vec{v}) &\mapsto P + \vec{v}\end{aligned}$$

such that:

1. $P + \vec{0} = P \ \forall P \in X$.
2. $P + (\vec{v} + \vec{w}) = (P + \vec{v}) + \vec{w} \ \forall P \in X$ and $\forall \vec{v}, \vec{w} \in V$.
3. For all $P, Q \in X \ \exists! \vec{v} \in V : Q = P + \vec{v}$. We denote the vector $\vec{v}$ by $\overrightarrow{PQ}$.

Definition 4.97. Let $\mathbb{A}$ be an affine space associated to a vector space V over a field k[44]. We define the *dimension of* $\mathbb{A}$ as $\dim \mathbb{A} = \dim V$.

Proposition 4.98. Let $\mathbb{A}$ be an affine space, $P, Q, R, S \in \mathbb{A}$. Then, the following properties are satisfied:

1. $\overrightarrow{PQ} = \vec{0} \iff P = Q$.
2. $\overrightarrow{PQ} = -\overrightarrow{QP}$.
3. $\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}$.
4. $\overrightarrow{PQ} = \overrightarrow{RS} \implies \overrightarrow{PR} = \overrightarrow{QS}$.

Definition 4.99. Let $\mathbb{A}$ be an affine space, $P_1, \dots, P_n \in \mathbb{A}$ and $\lambda_1, \dots, \lambda_n \in k$ such that $\lambda_1 + \cdots + \lambda_n = 1$. Given an arbitrary point $O \in \mathbb{A}$, we define the *affine combination of* $P_1, \dots, P_n$ as

$$\lambda_1 P_1 + \cdots + \lambda_n P_n := O + (\lambda_1 \overrightarrow{OP_1} + \cdots + \lambda_n \overrightarrow{OP_n}).$$

We say the points $P_1, \dots, P_n$ are *affinely independents* if the vectors $\overrightarrow{P_1P_2}, \dots, \overrightarrow{P_1P_n}$ are linearly independent.

Definition 4.100. Let $\mathbb{A}$ be an affine space and $P_1, \dots, P_r \in \mathbb{A}$. The *barycenter of the points* $P_1, \dots, P_r$ is

$$B := \frac{1}{r}(P_1 + \cdots + P_n).$$

Subvarieties and Graßmann formula

Definition 4.101. Let $\mathbb{A}$ be an affine space. If $P \in \mathbb{A}$ and F is a vector subspace of V, then an *affine subvariety of* $\mathbb{A}$ is the set:

$$P + F := \{P + \vec{v} \in \mathbb{A} : \vec{v} \in F\} = \{Q \in \mathbb{A} : \overrightarrow{PQ} \in F\}.$$

We say F is the *director subspace of the subvariety* $P + F$. If $\dim F = m$, then $\dim(P + F) = m$. If $m = 1$, we say the subvariety is *line*. If $m = \dim \mathbb{A} - 1$, we say la subvariety is a *hyperplane*.

Proposition 4.102. Let $P + F$ be an affine subvariety of an affine space $\mathbb{A}$. Then if $Q \in P + F$, we have $P + F = Q + F$.

Definition 4.103. Two subvarieties $P + F$ and $Q + G$ are said to be *parallel* if $F \subseteq G$ or $G \subseteq F$.

Definition 4.104. Let Y, Z be two subvarieties of an affine space $\mathbb{A}$ such that $Y \cap Z \neq \emptyset$ and let F, G be their director subspaces, respectively. Then if $P \in Y \cap Z$, we have that $Y \cap Z$ is a subvariety of $\mathbb{A}$ and $Y \cap Z = P + F \cap G$.

Definition 4.105. Let $Y = P + F$, $Z = Q + G$ be two subvarieties of an affine space $\mathbb{A}$. We define its *sum* as the subvariety

$$Y + Z := P + (F + G + \langle \overrightarrow{PQ} \rangle)$$ [45].

Theorem 4.106 (Affine Graßmann formulas). Let $L_1 = P_1 + F_1$, $L_2 = P_2 + F_2$ be two subvarieties of an affine space $\mathbb{A}$. Then:

- If $L_1 \cap L_2 \neq \emptyset$,

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(L_1 \cap L_2).$$

- If $L_1 \cap L_2 = \emptyset$,

$$\dim(L_1 + L_2) = \dim L_1 + \dim L_2 - \dim(F_1 \cap F_2) + 1.$$

Coordinates and equations

Definition 4.107. An affine frame in an affine space $\mathbb{A}$ is a pair $\mathcal{R} = \{P; \mathcal{B}\}$ formed by a point $P \in \mathbb{A}$ and a basis $\mathcal{B}$ of V. The point P is called the *origin* of this affine frame.

Definition 4.108. Let $\mathcal{R} = \{P; \mathcal{B}\}$, $\mathcal{B} = (\vec{e_1}, \dots, \vec{e_n})$, be an affine frame in an affine space $\mathbb{A}$ and let $Q \in \mathbb{A}$. We define *affine coordinates of* Q as

$$Q = (\lambda_1, \dots, \lambda_n) \iff \overrightarrow{PQ} = \lambda_1 \vec{e_1} + \cdots + \lambda_n \vec{e_n}.$$

Proposition 4.109. Let $\mathbb{A}$ be an affine space and $P_0, \dots, P_n \in \mathbb{A}$ be points satisfying the following equivalent properties:

1. The points are affinely independent.
2. There is no proper subvariety[46] containing all of them.
3. $P_0 + \cdots + P_n = \mathbb{A}$.
4. The vectors $\overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_n} \in V$ are linearly independent.

Then $\mathcal{R} = \{A_0; \overrightarrow{P_0P_1}, \dots, \overrightarrow{P_0P_n}\}$ is an affine frame in $\mathbb{A}$.

[44] From now on, for simplicity, we will only refer to the affine space by mentioning the set $\mathbb{A}$ without mentioning the associated vector space V over a field k.

[45] As expected, $Y + F$ is the smallest subvariety containing $Y \cup Z$.

[46] A proper subvariety Y of $\mathbb{A}$ is a subvariety such that $Y \neq \emptyset$ and $Y \neq \mathbb{A}$.

Definition 4.110. Let $\{\lambda_0, \dots, \lambda_n\}$ be homogeneous coordinates of a projective space $\mathcal{P}(V)$ and $(\mu_1, \dots, \mu_n)$ affine coordinates of an affine space $\mathbb{A}$. We call *homogenization* the transformation of affine coordinates to homogeneous coordinates as follows:

$$(\mu_1, \dots, \mu_n) \mapsto \{\mu_1, \dots, \mu_n, 1\}.$$

Similarly, we call *deshomogenization* the transformation of homogeneous coordinates to affine coordinates as follows:

$$\{\lambda_0, \dots, \lambda_n\} \mapsto \left(\frac{\lambda_0}{\lambda_n}, \dots, \frac{\lambda_{n-1}}{\lambda_n}\right).$$

Definition 4.111. Let $\mathcal{R} = \{P; \mathcal{B}\}$, $\mathcal{B} = (\overrightarrow{e_1}, \dots, \overrightarrow{e_n})$, be an affine frame in an affine space $\mathbb{A}$ and $L = Q + F$ be a subvariety of $\mathbb{A}$. Let $Q = (q_1, \dots, q_n)$ be a point of $\mathbb{A}$ and $(\overrightarrow{v_1}, \dots, \overrightarrow{v_r})$ be a basis of F. We call *parametric equations of* L the equations

$$(x_1, \dots, x_n) = (q_1, \dots, q_n) + \sum_{i=1}^{r} \lambda_i v_i.$$

If $\lambda_1, \dots, \lambda_r \in k$ we get the coordinates of $(x_1, \dots, x_n)$. If $\overrightarrow{v_j} = \sum_{i=1}^{n} \alpha_{ij} e_i$, $j = 1, \dots, r$ we can rearrange the parametric equations to get:

$$\begin{pmatrix} x_1 - q_1 \\ \vdots \\ x_n - q_n \end{pmatrix} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1r} \\ \vdots & & \vdots \\ \alpha_{n1} & \cdots & \alpha_{nr} \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix}.$$

The *Cartesian equations of* L are those obtained by equating to zero the minors of size $(r+1) \times (r+1)$ of the augmented matrix $(\alpha_{ij} \mid x_i - q_i)$.

Affinities

Definition 4.112. A map $f : \mathbb{A}_1 \to \mathbb{A}_2$ between two affine space over vector spaces V_1, V_2 is an *affinity* if there exists a linear map $\phi : V_1 \to V_2$ such that for all $P \in X$, $\vec{v} \in V_1$

$$f(P + \vec{v}) = f(P) + \phi(\vec{v})^{47}.$$

We call the *differential of* f, denoted by df, the map ϕ.

Proposition 4.113. Let $f : \mathbb{A}_1 \to \mathbb{A}_2$ and $g : \mathbb{A}_2 \to \mathbb{A}_3$ be affinities. Then $g \circ f : \mathbb{A}_1 \to \mathbb{A}_3$ is an affinity and $d(g \circ f) = dg \circ df$.

Proposition 4.114. Let $f : \mathbb{A}_1 \to \mathbb{A}_2$ be an affinity and $P, Q \in \mathbb{A}_1$. Then

$$df(\overrightarrow{PQ}) = \overrightarrow{f(P)f(Q)}.$$

Proposition 4.115. Let $f, g : \mathbb{A}_1 \to \mathbb{A}_2$ be affinities such that $f(P) = g(P)$ for some $P \in \mathbb{A}_1$ and $df = dg$. Then, $f = g$.

[47] If ϕ is a semilinear map, then we say f is a *semiafinity*.

Proposition 4.116. Let $f : \mathbb{A}_1 \to \mathbb{A}_2$ be an affinity and $\lambda_1, \dots, \lambda_r$ such that $\lambda_1 + \cdots + \lambda_r = 1$. Then

$$f(\lambda_1 P_1 + \cdots + \lambda_r P_r) = \lambda_1 f(P_1) + \cdots + \lambda_r f(P_r).$$

Proposition 4.117. Let $f : \mathbb{A}_1 \to \mathbb{A}_2$ be an affinity and $L = P + F$ be a subvariety of $\mathbb{A}$. Then $f(P + F)$ is a subvariety of $\mathbb{A}$ and

$$f(P + F) = f(P) + df(F).$$

Proposition 4.118. Let $f : \mathbb{A}_1 \to \mathbb{A}_2$ be an affinity and $\mathcal{R}_1 = \{P_1; (\overrightarrow{e_1}, \dots, \overrightarrow{e_n})\}$, $\mathcal{R}_2 = \{P_2; (\overrightarrow{v_1}, \dots, \overrightarrow{v_m})\}$ be affine frames of $\mathbb{A}_1$, $\mathbb{A}_2$, respectively. If $x = (x_1, \dots, x_n) \in \mathbb{A}_1$ and $y = (y_1, \dots, y_n) \in \mathbb{A}_2$ then

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \vdots \\ \rho_m \end{pmatrix} + M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

o, equivalently,

$$\begin{pmatrix} y_1 \\ \vdots \\ y_m \\ 1 \end{pmatrix} = \left(\begin{array}{ccc|c} & M & & \rho_1 \\ & & & \vdots \\ & & & \rho_m \\ \hline 0 & \cdots & 0 & 1 \end{array}\right) \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix} = M_{\mathcal{R}_1, \mathcal{R}_2}(f) \begin{pmatrix} x_1 \\ \vdots \\ x_n \\ 1 \end{pmatrix}$$

where M is the matrix associated with df and $(\rho_1, \dots, \rho_m)$ are the coordinates of $\overrightarrow{P_2 f(P_1)}$ in the basis $(\overrightarrow{v_1}, \dots, \overrightarrow{v_m})$. Here, $M_{\mathcal{R}_1, \mathcal{R}_2}(f)$ denote the matrix of f with respect to affine frames $\mathcal{R}_1$, $\mathcal{R}_2$.

Examples of affinities

Definition 4.119. Two affinities $f, g : \mathbb{A} \to \mathbb{A}$ are *similar* if there exist a bijective affinity $h : \mathbb{A} \to \mathbb{A}$ such that $h^{-1} f h = g$.

Proposition 4.120. Two affinities f, g are similar if there exist affine frames $\mathcal{R}$, $\mathcal{R}'$ such that $M_{\mathcal{R}}(f) = M_{\mathcal{R}'}(g)$.

Definition 4.121. A point $P \in \mathbb{A}$ is a *fixed point of* $f : \mathbb{A} \to \mathbb{A}$ if $f(P) = P$.

Definition 4.122. A linear subvariety $L = P + F \subset \mathbb{A}$ is *invariant under an affinity* $f : \mathbb{A} \to \mathbb{A}$ if $f(L) \subset L$.

Proposition 4.123. A linear subvariety $L = P + F \subset \mathbb{A}$ is invariant under an affinity $f : \mathbb{A} \to \mathbb{A}$ if and only if

1. $df(F) \subset F$.
2. $\overrightarrow{Pf(P)} \in F$.

In particular, a line $r = P + \langle \vec{v} \rangle$ is invariant under f if and only if

1. $\vec{v}$ is an eigenvector of df.
2. $\overrightarrow{Pf(P)} \in \langle \vec{v} \rangle$.

Proposition 4.124. If the set of fixed points of an affinity f, $\mathrm{Fix}(f)$, is non-empty, then $\mathrm{Fix}(f)$ is a subvariety.

Definition 4.125. Let f be an affinity. We define the *invariance level of* f, $\rho(f)$, as

$$\rho(f) = \min\{\dim L : f(L) \subset L \subset \mathbb{A}\} \in \{0, \dots, \dim \mathbb{A}\}.$$

Definition 4.126 (Translations). Let $\mathbb{A}$ be an affine space and $\vec{v} \neq 0$. A *translation* with translation vector $\vec{v}$ is an affinity $T_{\vec{v}} : \mathbb{A} \to \mathbb{A}$ defined by $T_{\vec{v}} = P + \vec{v}$.

Proposition 4.127 (Properties of translations). Let $T_{\vec{v}}$ be a translation. Then:

1. $\mathrm{Fix}(T_{\vec{v}}) = \emptyset$.
2. Invariant lines are those with director subspace $\langle \vec{v} \rangle$.
3. If $\mathcal{R} = \{P; (\vec{v}, \overrightarrow{v_2}, \dots, \overrightarrow{v_n})\}$ is an affine frame, then
$$M_{\mathcal{R}}(T_{\vec{v}}) = \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 1 \\ 0 & \ddots & \ddots & \vdots & 0 \\ \vdots & \ddots & 1 & 0 & \vdots \\ 0 & \cdots & 0 & 1 & 0 \\ \hline 0 & \cdots & 0 & 0 & 1 \end{array}\right).$$
4. All translations are similar and $\rho(T_{\vec{v}}) = 1$.

Definition 4.128 (Reflections). Let $\mathbb{A}$ be an affine space and suppose $\mathrm{char}(k) \neq 2$. Let $H = P + E$ be a hyperplane of $\mathbb{A}$ and let $\vec{v} \notin E$. The *reflection of* $\vec{v}$ *with respect to* H is the unique affinity $f : \mathbb{A} \to \mathbb{A}$ such that $f(P) = P$ for all $P \in H$ and $df(\vec{v}) = -\vec{v}$. Usually H is called the *mirror of the reflection* and $\vec{v}$ the *root of the reflection*.

Proposition 4.129 (Properties of reflections). Let f be a reflection with root $\vec{v}$ and mirror $H = P + E$. Then:

1. $\mathrm{Fix}(f) = H$.
2. Invariant lines are those contained on H and those with director subspace $\langle \vec{v} \rangle$.
3. If $\mathcal{R} = \{P; (\overrightarrow{v_1}, \dots, \overrightarrow{v_{n-1}}, \vec{v})\}$ is an affine frame such that $P \in H$ and $\overrightarrow{v_1}, \dots, \overrightarrow{v_{n-1}} \in E$, then
$$M_{\mathcal{R}}(f) = \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & 0 \\ \vdots & \ddots & 1 & 0 & \vdots \\ 0 & \cdots & 0 & -1 & 0 \\ \hline 0 & \cdots & 0 & 0 & 1 \end{array}\right)$$
4. All reflections are similar and $\rho(f) = 0$.

Definition 4.130 (Projections). Let $\mathbb{A}$ be an affine space and H a hyperplane of $\mathbb{A}$ with director subspace E and let $\vec{v} \notin E$. The *projection over* H *in the direction of* $\vec{v}$ is the affinity $f : \mathbb{A} \to \mathbb{A}$ such that $f(P) = P$ for all $P \in H$ and $df(\vec{v}) = 0$.

Proposition 4.131 (Properties of projections). Let f be a projection over $H = P + E$ in the direction of $\vec{v}$. Then:

1. $\mathrm{Fix}(f) = H$.
2. Invariant lines are those contained on H.
3. If $\mathcal{R} = \{P; (\overrightarrow{v_1}, \dots, \overrightarrow{v_{n-1}}, \vec{v})\}$ is an affine frame such that $P \in H$ and $\overrightarrow{v_1}, \dots, \overrightarrow{v_{n-1}} \in E$, then
$$M_{\mathcal{R}}(f) = \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 0 \\ 0 & \ddots & \ddots & \vdots & 0 \\ \vdots & \ddots & 1 & 0 & \vdots \\ 0 & \cdots & 0 & 0 & 0 \\ \hline 0 & \cdots & 0 & 0 & 1 \end{array}\right)$$
4. All projections are similar and $\rho(f) = 0$.

Definition 4.132 (Homotheties). An *homothety* is an affinity $f : \mathbb{A} \to \mathbb{A}$ such that $df = \lambda id$, $\lambda \neq 0, 1$. This λ is called the *similitude ratio of the homothety*.

Proposition 4.133 (Properties of homotheties). Let f be an homothety of similitude ratio λ. Then:

1. f has a unique fixed.
2. If $\mathcal{R} = \{P; \mathcal{B}\}$ is an affine frame with $P \in \mathrm{Fix}(f)$ and $\mathcal{B}$ an arbitrary basis, then
$$M_{\mathcal{R}}(f) = \left(\begin{array}{ccc|c} & \lambda Id & & 0 \\ & & & \vdots \\ & & & 0 \\ \hline 0 & \cdots & 0 & 1 \end{array}\right)$$
3. Two homotheties are similar if and only if they have the same similitude ratio. Moreover, $\rho(f) = 0$.

Proposition 4.134. Let $T_{\overrightarrow{w}}$ be a translation and R a reflection with root $\vec{v}$ with respect to the hyperplane $H = P + E$. Let $f = T_{\overrightarrow{w}} \circ R$. We take an affine frame $\mathcal{R} = \{P; (\overrightarrow{v_1}, \dots, \overrightarrow{v_{n-1}}, \vec{v})\}$ such that $\overrightarrow{v_1}, \dots, \overrightarrow{v_{n-1}} \in E$. Then if $\vec{w} = (w_1, \dots, w_n)$ in this frame we have,
$$M_{\mathcal{R}}(f) = \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & w_1 \\ 0 & \ddots & \ddots & \vdots & w_2 \\ \vdots & \ddots & 1 & 0 & \vdots \\ 0 & \cdots & 0 & -1 & w_n \\ \hline 0 & \cdots & 0 & 0 & 1 \end{array}\right)$$
1. If $\vec{w} \in \langle \vec{v} \rangle \implies w_1 = \cdots = w_{n-1} = 0$ and therefore f is a reflection with mirror the hyperplane $2x_n = w_n$.
2. If $\vec{w} \notin \langle \vec{v} \rangle$ we say f is a *glide reflection*. In this case, if $\vec{w} = w_n \vec{v} + \vec{e}$ with $\vec{e} \in E$ and we take $\mathcal{R} = (P + \frac{w_n}{2}\vec{v}; (\vec{e}, \overrightarrow{e_2} \dots, \overrightarrow{e_{n-1}}, \vec{v}))$, then
$$M_{\mathcal{R}}(f) = \left(\begin{array}{cccc|c} 1 & 0 & \cdots & 0 & 1 \\ 0 & \ddots & \ddots & \vdots & 0 \\ \vdots & \ddots & 1 & 0 & \vdots \\ 0 & \cdots & 0 & -1 & 0 \\ \hline 0 & \cdots & 0 & 0 & 1 \end{array}\right).$$

The invariance level of glide reflections is $\rho(f) = 1$.

Fundamental theorem of affine geometry

Definition 4.135 (Simple ratio). Let $A, B, C \in \mathbb{A}$ be three different collinear points. The *simple ratio of* A, B, C is the unique scalar $\lambda := (A, B, C) \in k$ such that

$$\overrightarrow{AB} = \lambda \overrightarrow{AC}.$$

Theorem 4.136 (Fundamental theorem of affine geometry). Let $f : \mathbb{A} \to \mathbb{A}$ be a collineation of an affine space of dimension $n \geq 2$ over the field k with more than two elements. Then f is a semiaffinity.

Proposition 4.137. Two affinities $f, g : \mathbb{A} \to \mathbb{A}$ are similar if and only if

1. df and dg are similar.
2. $\rho(f) = \rho(g)$.

Theorem 4.138. Let $f : \mathbb{A} \to \mathbb{A}$ be an affinity and $P \in \mathbb{A}$ a point. Let $\vec{v} := \overrightarrow{Pf(P)}$. Then

$$\rho(f) = \min\{r : (df - id)^r(\vec{v}) \in \mathrm{Im}(df - id)^{r+1}\}.$$

Corollary 4.139. If f is a affinity and 1 is not an eigenvector of df, then $\rho(f) = 0$.

Euclidean affine spaces

Definition 4.140. An *Euclidean affine space* is an affine space such that the associated vector space is an Euclidean vector space[48].

Definition 4.141. Let $\mathbb{A}$ be an Euclidean affine space. We define the *distance between two points* $P, Q \in \mathbb{A}$ as

$$d(A, B) := \|\overrightarrow{AB}\|.$$

We define the *segment delimited by* A *and* B as

$$\{P \in \mathbb{A} : P = \lambda A + (1 - \lambda)B, \lambda \in [0, 1]\}.$$

Proposition 4.142. Let $\mathbb{A}$ be an Euclidean affine space. Then the following properties are satisfied:

1. $d(A, C) \leq d(A, B) + d(B, C)$ *(Triangular inequality).*

If ABC is a right triangle with right angle at A, then:

2. $d(B, C)^2 = d(A, B)^2 + d(A, C)^2$ *(Pythagorean theorem).*

Definition 4.143. Two subvarieties $L_1 = P_1 + F_1$, $L_2 = P_2 + F_2$ of an Euclidean affine space $\mathbb{A}$ are *orthogonal*, $L_1 \perp L_2$, if $F_1 \perp F_2$[49].

Definition 4.144. Let $L_1 = P_1 + F_1$, $L_2 = P_2 + F_2$ be two subvarieties of an Euclidean affine space $\mathbb{A}$. We define the *distance between two affine subvarieties* as

$$d(L_1, L_2) := \inf\{d(A_1, A_2) : A_1 \in L_1, A_2 \in L_2\}.$$

Theorem 4.145. Let $L_1 = P_1 + F_1$, $L_2 = P_2 + F_2$ be two subvarieties of an Euclidean affine space $\mathbb{A}$. Let $\vec{u} = \overrightarrow{u_1} + \overrightarrow{u_2} \in F_1 + F_2$ and $\vec{v} \in (F_1 + F_2)^\perp$ such that $\overrightarrow{P_1P_2} = \vec{u} + \vec{v}$. Then we have

$$d(L_1, L_2) = \|\vec{v}\| = d(P_1 + \overrightarrow{u_1}, P_2 - \overrightarrow{u_2}).$$

Euclidean motions

Definition 4.146. Let $\mathbb{A}$ be an Euclidean affine space. A map $f : \mathbb{A} \to \mathbb{A}$ is an *Euclidean motion* if

$$d(f(A), f(B)) = d(A, B) \quad \forall P, Q \in \mathbb{A}.$$

Proposition 4.147. Let $\mathbb{A}$ be an Euclidean affine space. $f : \mathbb{A} \to \mathbb{A}$ is an Euclidean motion if and only if f is a affinity and df is a isometry[50].

Proposition 4.148 (Examples of Euclidean motions).

- Any translation $T_{\vec{v}}$ is an Euclidean motion. Moreover, $T_{\vec{u}} \sim T_{\vec{v}}$ (as Euclidean motions) if and only if $\|\vec{u}\| = \|\vec{v}\|$.
- An homothety f of similitude ratio λ is an Euclidean motion if and only if $\lambda = -1$. Moreover, all homotheties are similar as Euclidean motions.
- A reflection f of mirror $H = Q + E$ and root $\vec{v}$ is an Euclidean motion if and only if $\langle \vec{v} \rangle \perp E$. These reflections are called *orthogonal reflections*. If $\vec{n}$ is a unit normal vector to the mirror, then the orthogonal reflection is given by
$$f(P) = P - 2\langle \overrightarrow{QP}, \vec{n} \rangle \vec{n}.$$
- *Glide orthogonal reflections* are Euclidean motions.
- A rotation on the affine plane is an Euclidean motion, whose differential is a rotation of an angle other than zero. This affinity has a unique fixed point and if we take this point as a reference, its matrix in this frame will be
$$\begin{pmatrix} \cos\alpha & -\sin\alpha & 0 \\ \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

[48] Remember definition 2.96.
[49] Remember definition 2.86.
[50] Remember definition 2.88. From this we deduce that if $A \in \mathcal{M}_n(k)$ is the matrix associated with an isometry, then $AA^t = I_n$.

Classification of Euclidean motions

Theorem 4.149 (Classification of isometries).

1. Two isometries are similar if and only if they have the same characteristic polynomial.
2. For any isometry, there exists an orthonormal basis in which the matrix associated with the isometry is of the form

$$\begin{pmatrix} I_r & & & & \\ & -I_s & & & \\ & & R_1 & & \\ & & & \ddots & \\ & & & & R_t \end{pmatrix}$$

where $r, s, t \geq 0$, I_m denote the identity matrix of size $m \times m$ and each R_i is a rotation with matrix

$$R_i = \begin{pmatrix} \cos\alpha_i & -\sin\alpha_i \\ \sin\alpha_i & \cos\alpha_i \end{pmatrix}.$$

with $\alpha_i \neq 0, \pi$ for $i = 1, \ldots, t$.

Definition 4.150. Let $P \in \mathbb{A}$ be a point of an Euclidean affine space and $f : \mathbb{A} \to \mathbb{A}$ be an Euclidean motion. Express the vector $\overrightarrow{Pf(P)}$ as

$$\overrightarrow{Pf(P)} = \vec{u} + \vec{v} \quad \vec{u} \in \ker(df - I), \vec{v} \in \mathrm{Im}(df - I).$$

Then $\overrightarrow{u_f} := \vec{u}$ is the *glide vector of* f.

Proposition 4.151. The glide vector $\overrightarrow{u_f}$ has the following properties:

- $df(\overrightarrow{u_f}) = \overrightarrow{u_f}$.
- $\overrightarrow{u_f}$ does not depend on the point P.
- If $\overrightarrow{u_f} = 0 \implies \rho(f) = 0$. Otherwise, $\rho(f) = 1$.

Theorem 4.152 (Classification of Euclidean motions). Two Euclidean motions $f, g : \mathbb{A} \to \mathbb{A}$ are similar (as Euclidean motions) if and only if $df \sim dg$ (as isometries) and $\|\overrightarrow{u_f}\| = \|\vec{u}_g\|$.

2.4.4 | Quadrics

Quadrics

Definition 4.153. Let $\mathbb{A}$ an affine space of dimension n over a field k. A *quadric in* $\mathbb{A}$ is a polynomial of degree 2 with n variables, $p(x_1, \ldots, x_n)$, and coefficients in the field k modulo the equivalence relation

$$p(x_1, \ldots, x_n) \sim \lambda p(x_1, \ldots, x_n) \quad \text{if } \lambda \in k, \lambda \neq 0.$$

The *points of the quadric* $p(x_1, \ldots, x_n)$ are

$$\{(a_1, \ldots, a_n) \in \mathbb{A} : p(a_1, \ldots, a_n) = 0\}.$$

Definition 4.154. A *conic* is a quadric in a 2-dimensional space.

Definition 4.155. Two quadrics p, q of an affine space $\mathbb{A}$ are *equivalent* if there exists a bijective affinity $f : \mathbb{A} \to \mathbb{A}$ such that $f(p) = q$.

Definition 4.156. Let $\mathcal{P}_n(k)$ be a projective space of dimension n over a field k. A *quadric in* $\mathcal{P}_n(k)$ is a homogeneous polynomial of degree 2 with $n+1$ variables, $p(x_1, \ldots, x_{n+1})$, and coefficients in the field k modulo the equivalence relation

$$p(x_1, \ldots, x_{n+1}) \sim \lambda p(x_1, \ldots, x_{n+1}) \quad \text{si } \lambda \in k, \lambda \neq 0.$$

The *points of the quadric* $p(x_1, \ldots, x_{n+1})$ are

$$\{(a_1, \ldots, a_{n+1}) \in \mathcal{P}_n(k) : p(a_1, \ldots, a_{n+1}) = 0\}.$$

Definition 4.157. Two quadrics p, q in $\mathcal{P}_n(k)$ are *equivalent* if there exists a homography $f : \mathcal{P}_n(k) \to \mathcal{P}_n(k)$ such that $f(p) = q$.

Theorem 4.158. There is a bijective correspondence between quadrics of k^n and quadrics of $\mathcal{P}_n(k)$ not divisible by x_{n+1}. Thus, the points of the affine quadric are the points of the projective quadric that are in the affine space[51].

Proposition 4.159. Let $\mathbb{A}$ be an affine space and $\mathcal{P}_n(k)$ a projective space, both of dimension n and over a field k. Let p be a quadric.

- *Homogenization*: If $p(x_1, \ldots, x_n) \in \mathbb{A}$, then

$$p(x_1, \ldots, x_n) \mapsto x_{n+1}^2 p\left(\frac{x_1}{x_{n+1}}, \ldots, \frac{x_n}{x_{n+1}}\right) \in \mathcal{P}_n(k).$$

- *Deshomogenization*: If $p(x_1, \ldots, x_{n+1}) \in \mathcal{P}_n(k)$, then

$$p(x_1, \ldots, x_{n+1}) \mapsto p(x_1, \ldots, x_n, 1) \in \mathbb{A}.$$

Four points of view of quadrics

Definition 4.160. We say a bilinear form is *anisotropic* or *elliptic* if the unique isotropic vector[52] is the null vector.

Theorem 4.161. There is, expect for equivalence, only one symmetric bilinear form of dimension 2 such that it is non-singular[53] and non-elliptic. We call this bilinear form *hyperbolic plane*.

Definition 4.162. Let $\varphi : V \times V \to k$ a symmetric bilinear form. We define the *quadratic form associated with* φ as

$$\begin{aligned} q : V &\longrightarrow k \\ \vec{u} &\longmapsto \varphi(\vec{u}, \vec{u}). \end{aligned}$$

This map, clearly satisfies:

[51] Nevertheless, observe that two equivalent projective quadrics as projective quadrics may not be equivalent as affine quadrics.
[52] Remember definition 2.80.
[53] Remember definition 2.82.

1. $q(\lambda\vec{u}) = \lambda^2\vec{u}$.
2. $\varphi(\vec{u},\vec{v}) = \frac{1}{2}\left(q(\vec{u}+\vec{v}) - q(\vec{u}) - q(\vec{v})\right)$.

Proposition 4.163. Two symmetric bilinear forms φ_1, φ_2 over V are equivalent if there exists an isomorphism $\phi : V \to V$ such that $\varphi_1(\vec{u},\vec{v}) = \varphi_2(\phi(\vec{u}),\phi(\vec{v}))\ \forall\vec{u},\vec{v} \in V$. Two quadratic forms q_1, q_2 over V are equivalent if there exists an isomorphism $\phi : V \to V$ such that $q_1(\vec{u}) = q_2(\phi(\vec{u}))\ \forall\vec{u} \in V$.

Theorem 4.164. Symmetric bilinear forms, quadratic forms, symmetric matrices and homogeneous polynomials of degree 2 are equivalents ways to study quadrics.

Definition 4.165. A quadric is *non-degenerate* if its associated quadratic form is non-singular.

Classification of quadratic forms and quadrics

Definition 4.166. A *quadratic space* is a pair (V, q) where V is a vector space over a field k and q is a quadratic form.

Definition 4.167. Let $E_1 = (V_1, q_1)$ and $E_2 = (V_2, q_2)$ be two quadratic spaces. An *isometry between E_1 and E_2*, $E_1 \cong E_2$, is an isomorphism $\phi : V_1 \to V_2$ such that $q_1(\vec{v}) = q_2(\phi(\vec{v}))\ \forall\vec{v} \in V$.

Definition 4.168. Let (V, q) be a quadratic space. (V, q) is *totally isotropic* if all its vectors are isotropic.

Definition 4.169. Let (V, q) be a quadratic space. We define the *rank of* (V, q) as

$$\rho(V) := \dim V - \dim \mathrm{Rad}(V)^{54}.$$

Theorem 4.170 (Witt's theorem). Let E be a quadratic space and suppose that $E = E_1 \perp F_1 = E_2 \perp F_2$. If $E_1 \cong E_2$, then $F_1 \cong F_2$.

Definition 4.171. Let (V, q) be a quadratic space. We define the *index of* (V, q) as

$$\iota(V) := \max\{\dim F : F \subseteq V \text{ and } F \text{ is totally isotropic}\}.$$

Theorem 4.172. Let $E \subseteq V$ a totally isotropic subspace of maximum dimension and $(\vec{e_1},\dots,\vec{e_r})$ a basis of E (therefore, $r = \iota(V)$). Then, there exist vectors $\vec{v_1},\dots,\vec{v_r} \in V$ such that each $H_i := \langle\vec{e_i},\vec{v_i}\rangle$ is an hyperbolic plane and $V = H_1 \perp \cdots \perp H_r \perp F$, where F is anisotropic.

Proposition 4.173. Let (V, q) be a quadratic space and M be the associated matrix of q. Then $\dim V$, $\rho(V)$, $\iota(V)$ and $\det M$ modulo squares[55] are invariant under isometries.

Theorem 4.174 (Classification of quadratic forms in $\mathbb{C}$). If $k = \mathbb{C}$, two quadratic forms are equivalent if and only if they have the same rank. All quadratic forms of rank r are equivalent to

$$x_1^2 + \cdots + x_r^2.$$

Theorem 4.175 (Classification of quadratic forms in $\mathbb{F}_q$). If $k = \mathbb{F}_q$ with q odd, all quadratic form of rank n are equivalent to either of these two diagonal forms:

$$x_1^2 + \cdots + x_n^2,$$
$$x_1^2 + \cdots + x_{n-1}^2 + \nu x_n^2,$$

where ν in not a square. Moreover, two quadratic forms are equivalent if and only if they have the same rank and determinant (modulo squares).

Theorem 4.176 (Classification of quadratic forms in $\mathbb{R}$). If $k = \mathbb{R}$, all quadratic forms of rang r are equivalent to the diagonal form

$$\pm x_1^2 \pm \cdots \pm x_r^2.$$

If we denote by r^+ the number of positive signs and by r^- the number of negative signs, then two quadratic forms are equivalents if and only if they have the same values (r^+, r^-).

Theorem 4.177 (Classification of projective quadrics in $\mathbb{C}$). If $k = \mathbb{C}$, two projective quadrics are equivalents if and only if they have the same rank.

Theorem 4.178 (Classification of projective quadrics in $\mathbb{F}_q$). If $k = \mathbb{F}_q$, there are (except of equivalence) this projective quadrics in each rank n:

- If n is odd:
$$x_1^2 + \cdots + x_n^2.$$
- If n is even:
$$x_1^2 + \cdots + x_n^2,$$
$$x_1^2 + \cdots + x_{n-1}^2 + \nu x_n^2,$$
where ν in not a square.

Theorem 4.179 (Classification of projective quadrics in $\mathbb{R}$). If $k = \mathbb{R}$, two projective quadrics are equivalents if they have the same rank and index.

Theorem 4.180 (Classification of affine quadrics). Let q_1, q_2 be two affine quadrics. $q_1 \sim q_2$ if and only if:

1. $q_1 \sim q_2$ as projective quadrics, that is, in $P_n(k)$.
2. $q_1^\infty \sim q_2^\infty$ as quadrics in $H \cong P_{n-1}(k)$[56].

[54] If A is the associated matrix of q, we have $\operatorname{rank} A = \rho(V)$.

[55] That is, if (V_1, q_1), (V_1, q_1) are two quadratic spaces and M_i, $i = 1, 2$, are the associated matrices to q_1, q_2, respectively, we have $\det M_1 = a^2 \det M_2$, for some $a \in k$.

[56] Here q_i^∞ is the quadric q_i restricted to the hyperplane "at infinity" H, that is, by agreement, to the hyperplane $x_{n+1} = 0$.

2.5 Mathematical analysis

2.5.1 | Numeric series

Series convergence

Definition 5.1. Let (a_n) be a sequence of real numbers. A *numeric series* is an expression of the form

$$\sum_{n=1}^{\infty} a_n.$$

We call a_n *general term of the series* and $S_N = \sum_{n=1}^{N} a_n$, for all $N \in \mathbb{N}$, *N-th partial sum of the series*[57].

Definition 5.2. We say the series $\sum a_n$ is *convergent* if the sequence of partial sums is convergent, that is, if $S = \lim_{N\to\infty} S_N$ exist and it's finite. In that case, S is called the *sum of the series*. If the previous limit doesn't exists or it is infinite we say the series is *divergent*[58].

Proposition 5.3. Let (a_n) be a sequence such that $\sum a_n < \infty$. Then $\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}$ such that

$$\left| \sum_{n=1}^{N} a_n - \sum_{n=1}^{\infty} a_n \right| < \varepsilon$$

if $N \geq n_0$.

Theorem 5.4 (Cauchy's test). Let (a_n) be a sequence. $\sum a_n < \infty$ if and only if $\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N}$ such that

$$\left| \sum_{n=N}^{M} a_n \right| < \varepsilon$$

if $M \geq N \geq n_0$.

Corollary 5.5. Changing a finite number of terms in a series has no effect on the convergence or divergence of the series.

Corollary 5.6. If $\sum a_n < \infty$, then $\lim_{n\to\infty} a_n = 0$.

Theorem 5.7 (Linearity). Let $\sum a_n, \sum b_n$ be two convergent series with sums A and B respectively and let λ be a real number. The series

$$\sum_{n=1}^{\infty} (a_n + \lambda b_n)$$

is convergent and has sum $A + \lambda B$.

Theorem 5.8 (Associative property). Let $\sum a_n$ be a convergent series with sum A. Suppose (n_k) is a strictly increasing sequence of natural numbers. The series $\sum b_n$, with $b_k = a_{n_{k-1}+1} + \cdots + a_{n_k}$ for all $i \in \mathbb{N}$, is convergent and its sum is A.

Non-negative terms series

Theorem 5.9. Let $\sum a_n$ be a series of non-negative terms $a_n \geq 0$[59]. The series converges if and only if the sequence (S_N) of partial sums is bounded.

Theorem 5.10 (Comparison test). Let $(a_n), (b_n) \geq 0$ be two sequences of real numbers. Suppose that exists a constant $C > 0$ and a number $n_0 \in \mathbb{N}$ such that $a_n \leq C b_n$ for all $n \geq n_0$.

1. If $\sum b_n < \infty \implies \sum a_n < \infty$.
2. If $\sum a_n = +\infty \implies \sum b_n = +\infty$.

Theorem 5.11 (Limit comparison test). Let (a_n), $(b_n) \geq 0$ be two sequences of real numbers. Suppose that the limit $\ell = \lim_{n\to\infty} \frac{a_n}{b_n}$ exists.

1. If $0 < \ell < \infty \implies \sum a_n < \infty \iff \sum b_n < \infty$.
2. If $\ell = 0$ and $\sum b_n < \infty \implies \sum a_n < \infty$.
3. If $\ell = \infty$ and $\sum a_n < \infty \implies \sum b_n < \infty$.

Theorem 5.12 (Root test). Let $(a_n) \geq 0$. Suppose that the limit $\ell = \lim_{n\to\infty} \sqrt[n]{a_n}$ exists.

1. If $\ell < 1 \implies \sum a_n < \infty$.
2. If $\ell > 1 \implies \sum a_n = +\infty$.

Theorem 5.13 (Ratio test). Let $(a_n) \geq 0$. Suppose that the limit $\ell = \lim_{n\to\infty} \frac{a_{n+1}}{a_n}$ exists.

1. If $\ell < 1 \implies \sum a_n < \infty$.
2. If $\ell > 1 \implies \sum a_n = +\infty$.

Theorem 5.14 (Raabe's test). Let $(a_n) \geq 0$. Suppose that the limit $\ell = \lim_{n\to\infty} n\left(1 - \frac{a_{n+1}}{a_n}\right)$ exists.

1. If $\ell > 1 \implies \sum a_n < \infty$.
2. If $\ell < 1 \implies \sum a_n = +\infty$.

Theorem 5.15 (Condensation test). Let $(a_n) \geq 0$ be a decreasing sequence. Then:

$$\sum a_n < \infty \iff \sum 2^n a_{2^n} < \infty.$$

[57] From now on we will write $\sum a_n$ to refer $\sum_{n=1}^{\infty} a_n$.

[58] We will use the notation $\sum a_n < \infty$ or $\sum a_n = +\infty$ to express that the series converges or diverges, respectively.

[59] Obviously the following results are also valid if the series is of non-positive terms or has a finite number of negative or positive terms.

Theorem 5.16 (Logarithmic test). Let $(a_n) \geq 0$. Suppose that the limit $\ell = \lim_{n\to\infty} \frac{\log \frac{1}{a_n}}{\log n}$ exists.

1. If $\ell > 1 \implies \sum a_n < \infty$.
2. If $\ell < 1 \implies \sum a_n = +\infty$.

Theorem 5.17 (Integral test). Let $f : [1, \infty) \to (0, \infty)$ be a decreasing function. Then:

$$\sum f(n) < \infty \iff$$
$$\iff \exists C > 0 \text{ such that } \int_1^n f(x)dx \leq C \ \forall n.$$

Alternating series

Definition 5.18. An *alternating series* is a series of the form $\sum(-1)^n a_n$, with $(a_n) \geq 0$.

Theorem 5.19 (Leibnitz's test). Let $(a_n) \geq 0$ be a decreasing sequence such that $\lim_{n\to\infty} a_n = 0$. Then, $\sum(-1)^n a_n$ is convergent.

Theorem 5.20 (Abel's summation formula). Let $(a_n), (b_n)$ be two sequences of real numbers. Then,

$$\sum_{n=N}^{M} a_n(b_{n+1} - b_n) = a_{M+1}b_{M+1} - a_N b_N - \\ - \sum_{n=N}^{M} b_{n+1}(a_{n+1} - a_n).$$

Theorem 5.21 (Dirichlet's test). Let $(a_n), (b_n)$ be two sequences of real numbers such that:

1. $\exists C > 0$ such that $\left|\sum_{n=1}^{N} a_n\right| \leq C$ for all $N \in \mathbb{N}$.
2. (b_n) is monotone and $\lim_{n\to\infty} b_n = 0$.

Then, $\sum a_n b_n$ is convergent.

Theorem 5.22 (Abel's test). Let $(a_n), (b_n)$ be two sequences of real numbers such that:

1. The series $\sum a_n$ is convergent.
2. (b_n) is monotone and bounded.

Then, $\sum a_n b_n$ is convergent.

Absolute convergence and rearrangement of series

Definition 5.23. We say a series $\sum a_n$ is *absolutely convergent* if $\sum |a_n|$ is convergent.

Theorem 5.24. If a series converges absolutely, it converges.

Definition 5.25. We say a sequence (b_n) is a *rearrangement of the sequence* (a_n) if exists a bijective map $\sigma : \mathbb{N} \to \mathbb{N}$ such that $b_n = a_{\sigma(n)}$. A *rearrangement of the series* $\sum a_n$ is the series $\sum a_{\sigma(n)}$ for some bijection $\sigma : \mathbb{N} \to \mathbb{N}$.

Definition 5.26. Let $x \in \mathbb{R}$. We define the *positive part of* x as

$$x^+ = \begin{cases} x & \text{si } x \geq 0 \\ 0 & \text{si } x < 0 \end{cases}$$

Analogously, we define the *negative part of* x as

$$x^- = \begin{cases} 0 & \text{si } x \geq 0 \\ -x & \text{si } x < 0 \end{cases}$$

Note that we can write $x = x^+ - x^-$ and $|x| = x^+ + x^-$.

Theorem 5.27. A series $\sum a_n$ is absolutely convergent if and only if positive and negative terms series, $\sum a_n^+$ and $\sum a_n^-$, converge. In this case,

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_n^+ - \sum_{n=1}^{\infty} a_n^-.$$

Theorem 5.28. Let $\sum a_n$ be an absolutely convergent series. Then, for all bijection $\sigma : \mathbb{N} \to \mathbb{N}$, the rearranged series $\sum a_{\sigma(n)}$ is absolutely convergent and $\sum a_n = \sum a_{\sigma(n)}$.

Theorem 5.29 (Riemann's theorem). Let $\sum a_n$ be a convergent series but not absolutely convergent. Then, $\forall \alpha \in \mathbb{R} \cup \{\infty\}$, there exists a bijective map $\sigma : \mathbb{N} \to \mathbb{N}$ such that $\sum a_{\sigma(n)}$ converges and $\sum a_{\sigma(n)} = \alpha$.

Theorem 5.30. A series $\sum a_n$ is absolutely convergent if and only if any rearranged series converges to the same value of $\sum a_n$.

2.5.2 | Sequences and series of functions

Sequences of functions

Definition 5.31. Let $D \subseteq \mathbb{R}$. A set

$$(f_n(x)) = \{f_1(x), f_2(x), \ldots, f_n(x), \ldots\}$$

is a *sequence of real functions* if $f_i : D \to \mathbb{R}$ is a real-valued function. In this case we say the sequence $(f_n(x))$, or simply (f_n), is well-defined on D.

Definition 5.32. Let (f_n) be a sequence of functions defined on $D \subseteq \mathbb{R}$ and $f : D \to \mathbb{R}$. We say (f_n) *converges pointwise to* f *on* D if $\forall x \in D$, $\lim_{n\to\infty} f_n(x) = f(x)$

Definition 5.33. Let (f_n) be a sequence of functions defined on $D \subseteq \mathbb{R}$ and $f : D \to \mathbb{R}$. We say (f_n) *converges uniformly to* f *on* D if $\forall \varepsilon > 0$, $\exists n_0 : |f_n(x) - f(x)| < \varepsilon$ $\forall n \geq n_0$ and $\forall x \in D$.

Lemma 5.34. Let (f_n) be an uniform convergent sequence of functions defined on $D \subseteq \mathbb{R}$ and let f be a function such that (f_n) converges pointwise to f. Then, (f_n) converges uniformly f on D.

Lemma 5.35. Let (f_n) be a sequence of functions defined on $D \subseteq \mathbb{R}$. (f_n) converges uniformly a f en D if and only if $\lim_{n\to\infty} \sup\{|f_n(x) - f(x)| : x \in D\} = 0$.

Corollary 5.36. A sequence of functions (f_n) converges uniformly to f on $D \subseteq \mathbb{R}$ if and only if there is a sequence (a_n), with $a_n \geq 0$ and $\lim_{n\to\infty} a_n = 0$, and a number $n_0 \in \mathbb{N}$ such that $\sup\{|f_n(x) - f(x)| : x \in D\} \leq a_n$, $\forall n \geq n_0$.

Theorem 5.37 (Cauchy's test). A sequence of functions (f_n) converges uniformly to f on $D \subseteq \mathbb{R}$ if and only if $\forall \varepsilon > 0\ \exists n_0 : \sup\{|f_n(x) - f_m(x)| : x \in D\} < \varepsilon$ if $n, m \geq n_0$.

Theorem 5.38. Let (f_n) be a sequence of continuous functions defined on $D \subseteq \mathbb{R}$. If (f_n) converges uniformly to f on D, then f is continuous on D, that is, for any $x_0 \in D$, it satisfies:

$$\lim_{n\to\infty} \left(\lim_{x\to x_0} f_n(x) \right) = \lim_{x\to x_0} f(x).$$

Theorem 5.39. Let (f_n) be a sequence of functions defined on $I = [a, b] \subseteq \mathbb{R}$. If (f_n) are Riemann-integrable on I and (f_n) converges uniformly to f on I, then f is Riemann-integrable on I and

$$\int_a^b \lim_{n\to\infty} f_n(x)dx = \lim_{n\to\infty} \int_a^b f_n(x)dx.$$

Theorem 5.40. Let (f_n) be a sequence of functions defined on $I = (a, b) \subset \mathbb{R}$. If (f_n) are derivable on I, $(f'_n(x))$ converges uniformly on I and $\exists x_0 \in I : \lim_{n\to\infty} f_n(x_0) \in \mathbb{R}$, then there is a function f such that (f_n) converges uniformly to f on I, f is derivable on I and $(f'_n(x))$ converges uniformly to f' on I.

Series of functions

Definition 5.41. Let (f_n) be a sequence of functions defined on $D \subseteq \mathbb{R}$. The expression

$$\sum_{n=1}^{\infty} f_n(x)$$

is the *series of functions associated with* (f_n).

Definition 5.42. A series of functions $\sum f_n(x)$ defined on $D \subseteq \mathbb{R}$ *converges pointwise on* D if the sequence of partials sums

$$F_N(x) = \sum_{n=1}^{N} f_n(x)$$

converges pointwise. If the pointwise limit of (F_N) is $F(x)$, we say F is the *sum of the series in a pointwise sense.*

Definition 5.43. A series of functions $\sum f_n(x)$ defined on $D \subseteq \mathbb{R}$ *converges uniformly on* D if the sequence of partials sums

$$F_N(x) = \sum_{n=1}^{N} f_n(x)$$

converges uniformly. If the uniform limit of (F_N) is $F(x)$, we say F is the *sum of the series in an uniform sense.*

Theorem 5.44 (Cauchy's test). A series of functions $\sum f_n(x)$ defined on $D \subseteq \mathbb{R}$ converges uniformly if and only if $\forall \varepsilon > 0\ \exists n_0$ such that

$$\sup\left\{ \left| \sum_{n=N}^{M} f_n(x) \right| : x \in D \right\} < \varepsilon$$

if $M \geq N \geq n_0$.

Corollary 5.45. If $\sum f_n(x)$ is a series of continuous functions on $D \subseteq \mathbb{R}$, then (f_n) converges uniformly to zero on D.

Theorem 5.46. If $\sum f_n(x)$ is uniformly convergent series of functions on $D \subseteq \mathbb{R}$, then its sum function is also continuous on D.

Theorem 5.47. Let (f_n) be a sequence of functions defined on $I = [a, b] \subseteq \mathbb{R}$. If (f_n) are Riemann-integrable on I and $\sum f_n(x)$ converges uniformly on I, then $\sum f_n(x)$ is Riemann-integrable on I and

$$\int_a^b \sum_{n=1}^{\infty} f_n(x)dx = \sum_{n=1}^{\infty} \int_a^b f_n(x)dx.$$

Theorem 5.48. Let (f_n) be a sequence of functions defined on $I = (a, b) \subset \mathbb{R}$. If (f_n) are derivable on I, $\sum f'_n(x)$ converges uniformly on I and $\exists c \in I : \sum f_n(c) < \infty$, then $\sum f_n(x)$ converges uniformly on I, $\sum f_n(x)$ is derivable on I and

$$\left(\sum_{n=1}^{\infty} f_n(x) \right)' = \sum_{n=1}^{\infty} f'_n(x).$$

Theorem 5.49 (Weierstraß M-test). Let (f_n) be a sequence of functions defined on $D \subseteq \mathbb{R}$ such that $|f_n(x)| \leq M_n\ \forall x \in D$ and suppose that $\sum M_n$ is a convergent numeric series. Then, $\sum f_n(x)$ is converges uniformly on D.

Theorem 5.50 (Dirichlet's test). Let $(f_n), (g_n)$ be two sequences of functions defined on $D \subseteq \mathbb{R}$. Suppose:

1. $\exists C > 0 : \sup\left\{ \left| \sum_{n=1}^{N} f_n(x) \right| : x \in D \right\} \leq C$, $\forall N$.
2. $(g_n(x))$ is a monotone sequence for all $x \in D$ and $\lim_{n\to\infty} \sup\{|g_n(x)| : x \in D\} = 0$.

Then, $\sum f_n(x)g_n(x)$ converges uniformly on D.

Theorem 5.51 (Abel's test). Let $(f_n), (g_n)$ be two sequences of functions defined on $D \subseteq \mathbb{R}$. Suppose:

1. The series $\sum f_n(x)$ converges uniformly on D.
2. $(g_n(x))$ is a monotone and bounded sequence for all $x \in D$.

Then, $\sum f_n(x)g_n(x)$ converges uniformly on D.

Power series

Definition 5.52. Let (a_n) be a sequence of real numbers and $x_0 \in \mathbb{R}$. A *power series centred on* x_0 is a series of functions of the form

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n.$$

Proposition 5.53. Let $\sum a_n(x-x_0)^n$ be a power series. Suppose there exists an $x_1 \in \mathbb{R}$ such that $\sum a_n(x_1-x_0)^n < \infty$. Then, $\sum a_n(x-x_0)^n$ converges uniformly on any closed interval $I \subset A = \{x \in \mathbb{R} : |x-x_0| < |x_1-x_0|\}$.

Theorem 5.54. Let $\sum a_n(x-x_0)^n$ be a power series and consider

$$R = \left(\limsup_{n\to\infty} \sqrt[n]{|a_n|}\right)^{-1} \in [0,\infty].$$

Then:

1. If $|x-x_0| < R \implies \sum a_n(x-x_0)^n$ converges absolutely.
2. If $0 \le r < R \implies \sum a_n(x-x_0)^n$ converges uniformly on $[x_0-r, x_0+r]$.
3. If $|x-x_0| > R \implies \sum a_n(x-x_0)^n$ diverges.

The number R is called *radius of convergence of the power series.*

Theorem 5.55 (Abel's theorem). Let $\sum a_n x^n$ be a power series[60] with radius of convergence R satisfying $\sum a_n R^n < \infty$. Then the series $\sum a_n x^n$ converges uniformly on $[0,R]$. In particular, if $f(x) = \sum a_n x^n$,

$$\lim_{x\to R^-} f(x) = \sum_{n=0}^{\infty} a_n R^n.$$

Corollary 5.56. Let f be the sum function of a power series $\sum a_n x^n$. Then f is continuous on the domain of convergence of the series.

Corollary 5.57. If the series $\sum a_n x^n$ has radius of convergence $R \neq 0$ and f is its sum function, then f is Riemann-integrable on any closed subinterval on the domain of convergence of the series. In particular, for $|x| < R$,

$$\int_0^x f(t)dt = \sum_{n=0}^{\infty} a_n \frac{x^{n+1}}{n+1}^{61}.$$

Corollary 5.58. Let f be the sum function of the power series $\sum a_n x^n$. Then f is derivable within the domain of convergence of the series and

$$f'(x) = \sum_{n=0}^{\infty} n a_n x^{n-1}.$$

Corollary 5.59. Any function f defined as a sum of a power series $\sum a_n x^n$ is indefinitely derivable within the domain of convergence of the series and

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdots(n-k+1)a_n x^{n-k},$$

for all $k \in \mathbb{N} \cup \{0\}$. In particular $f^{(k)}(0) = k!a_k$.

Definition 5.60. A function is *analytic* if it can be expressed locally as a power series.

Stone-Weierstraß approximation theorem

Definition 5.61. Let f be a real-valued function. We say f has *compact support*[62] if exists an $M > 0$ such that $f(x) = 0$ for all $x \in \mathbb{R} \setminus [-M, M]$.

Definition 5.62. Let f, g be real-valued functions with compact support. We define the convolution of f with g as

$$(f * g)(x) = \int_{\mathbb{R}} f(t)g(x-t)dt^{63}.$$

Definition 5.63. We sap a sequence of functions (ϕ_ε) with compact support is an *approximation of unity* if

1. $\phi_\varepsilon \ge 0$.
2. $\displaystyle\int_{\mathbb{R}} \phi_\varepsilon = 1$.
3. For all $\delta > 0$, $\phi_\varepsilon(t)$ converges uniformly to zero when $\varepsilon \to 0$ if $|t| > \delta$.

Lemma 5.64. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function with compact support. Let (ϕ_ε) be an approximation of unity. Then $(f * \phi_\varepsilon)$ converges uniformly to f on $\mathbb{R}$ when $\varepsilon \to 0$.

Theorem 5.65 (Stone-Weierstraß theorem). Let $f : [a,b] \to \mathbb{R}$ be a continuous function. Then, there exists polynomials $p_n \in \mathbb{R}[x]$ such that the sequence (p_n) converge uniformly to f on $[a,b]$.

[60] From now on we will suppose, for simplicity, $x_0 = 0$.
[61] The formula is also valid for $|x| = R$ if the series $\sum a_n R^n$ (or $\sum a_n(-R)^n$) is convergent.
[62] In general, the support of a function is the adherence of the set of points which are not mapped to zero.
[63] Alternatively if f, g are Riemann-integrable functions on $[a,b]$ we can define the convolution of f and g as

$$(f * g)(x) = \int_a^b f(t)g(x-t)dt.$$

2.5.3 | Improper integrals

Locally integrable functions

Definition 5.66. Let $f : [a, b) \to \mathbb{R}$, with $b \in \mathbb{R} \cup \{\infty\}$. We say f is *locally integrable on* $[a, b)$ if f is Riemann-integrable on $[a, x]$ for all $a \leq x < b$.

Definition 5.67. Let $f : [a, b) \to \mathbb{R}$ be a locally integrable function. If there exists

$$\lim_{x \to b^-} \int_a^x f$$

and it's finite, we say that the *improper integral of* f *on* $[a, b)$, $\int_a^b f$, is *convergent*.

Theorem 5.68 (Cauchy's test). Let $f : [a, b) \to \mathbb{R}$ be a locally integrable function. The improper integral $\int_a^b f$ is convergent if and only if $\forall \varepsilon > 0$ $\exists b_0$, $a < b_0 < b$, such that

$$\left| \int_{x_1}^{x_2} f \right| < \varepsilon$$

if $b_0 < x_1 < x_2 < b$.

Improper integrals of non-negative functions

Theorem 5.69. Let $f : [a, b) \to \mathbb{R}$ be a locally integrable non-negative function. A necessary and sufficient condition for $\int_a^b f$ to be convergent is that the function

$$F(x) = \int_a^x f(t)dt$$

must be bounded for all $x < b$.

Theorem 5.70 (Comparison test). Let $f, g : [a, b) \to [0, +\infty)$ be two locally integrable non-negative functions. Then:

1. If $\exists C > 0$ such that $f(x) \leq Cg(x)$ $\forall x$ on a neighborhood of b and $\int_a^b g < \infty \implies \int_a^b f < \infty$.
2. Suppose the limit $\ell = \lim_{x \to b} \frac{f(x)}{g(x)}$ exists. Then,
   i) If $\ell \in (0, \infty) \implies \int_a^b f < \infty \iff \int_a^b g < \infty$.
   ii) If $\ell = 0$ and $\int_a^b g < \infty \implies \int_a^b f < \infty$.
   iii) If $\ell = \infty$ and $\int_a^b f < \infty \implies \int_a^b g < \infty$.

Theorem 5.71 (Integral test). Let $f : [1, \infty) \to (0, \infty)$ be a locally integrable decreasing function. Then:

$$\sum f(n) < \infty \iff \int_1^\infty f(x)dx < \infty$$[64].

[64] This is another way of formulating theorem 5.17.

Absolute convergence of improper integrals

Definition 5.72. Let $f : [a, b) \to (0, \infty)$ be a locally integrable function. We say $\int_a^b f$ *converges absolutely* if $\int_a^b |f|$ is convergent.

Theorem 5.73 (Dirichlet's test). Let $f, g : [a, b) \to \mathbb{R}$ be two locally integrable functions Suppose:

1. $\exists C > 0$ such that $\left|\int_a^x f(t)dt\right| \leq C$ for all $x \in [a, b)$.
2. g is monotone and $\lim_{x \to b} g(x) = 0$.

Then, $\int_a^b fg$ is convergent.

Theorem 5.74 (Abel's test). Let $f, g : [a, b) \to \mathbb{R}$ be two locally integrable functions. Suppose:

1. $\int_a^b f$ is convergent.
2. g is monotone and bounded.

Then, $\int_a^b fg$ is convergent.

Differentiation under integral sign

Theorem 5.75. Let $f : [a, b] \times [c, d] \to \mathbb{R}$ be a continuous function on $[a, b] \times [c, d]$. Consider the function $F(y) = \int_a^b f(x, y)dx$ defined on $[c, d]$. Then, F is continuous, that is, if $c < y_0 < d$,

$$\lim_{y \to y_0} F(y) = \lim_{y \to y_0} \int_a^b f(x, y)dx = \int_a^b \lim_{y \to y_0} f(x, y)dx = \\ = \int_a^b f(x, y_0)dx = F(y_0).$$

Theorem 5.76. Let $f : [a, b] \times [c, d] \to \mathbb{R}$ be a Riemann-integrable function and let $F(y) = \int_a^b f(x, y)dx$. If f is differentiable with respect to y and $\partial f / \partial y$ is continuous on $[a, b] \times [c, d]$, then $F(y)$ is derivable on (c, d) and its derivative is

$$F'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y)dx,$$

for all $y \in (c, d)$.

Theorem 5.77. Let $f : [a,b] \times [c,d] \to \mathbb{R}$ be a continuous function on $[a,b] \times [c,d]$. Let $a, b : [c,d] \to \mathbb{R}$ be to differentiable functions satisfying $a \leq a(y) \leq b(y) \leq b$ for every $y \in [c,d]$. Suppose that $\partial f/\partial y$ is continuous on $\{(x,y) \in \mathbb{R}^2 : a(y) \leq x \leq b(y),\ c \leq y \leq d\}$. Then $F(y) = \displaystyle\int_{a(y)}^{b(y)} f(x,y)dx$ is derivable on (c,d) and its derivative is

$$F'(y) = b'(y)f(b(y),y) - a'(y)f(a(y),y) + \\ + \int_{a(y)}^{b(y)} \frac{\partial f}{\partial y}(x,y)dx,$$

for all $y \in (c,d)$.

Theorem 5.78. Let $f : [a,b) \times [c,d] \to \mathbb{R}$ be a continuous function on $[a,b) \times [c,d]$. We consider $F(y) = \displaystyle\int_a^b f(x,y)dx$. Suppose that:

1. $\dfrac{\partial f}{\partial y}$ is continuous on $[a,b) \times [c,d]$.
2. Given $y_0 \in [c,d]$, $\exists \delta > 0$ such that the integral
$$\int_a^b \sup\left\{\left|\frac{\partial f}{\partial y}(x,y)\right| : y \in (y_0-\delta, y_0+\delta)\right\} dx$$
exists and it's finite on $[a,b)$.

Then, $F(y)$ is derivable at y_0 and

$$F'(y_0) = \int_a^b \frac{\partial f}{\partial y}(x,y_0)dx.$$

Theorem 5.79. Let $f : [a,b) \times [c,d] \to \mathbb{R}$ be a continuous function on $[a,b) \times [c,d]$. Let $a, b : [c,d] \to \mathbb{R}$ be two differentiable functions satisfying $a \leq a(y) \leq b(y) \leq b$ for every $y \in [c,d]$. We consider $F(y) = \displaystyle\int_{a(y)}^{b(y)} f(x,y)dx$. Suppose that:

1. $\dfrac{\partial f}{\partial y}$ is continuous on $\{(x,y) \in \mathbb{R}^2 : a(y) \leq x \leq b(y),\ c \leq y \leq d\}$.
2. Given $y_0 \in [c,d]$, $\exists \delta > 0$ such that the integral
$$\int_{a(y)}^{b(y)} \sup\left\{\left|\frac{\partial f}{\partial y}(x,y)\right| : y \in (y_0-\delta, y_0+\delta)\right\} dx$$
exists and it's finite on $[a,b)$.

The, $F(y)$ is derivable at y_0 and

$$F'(y_0) = b'(y_0)f(b(y_0),y_0) - a'(y_0)f(a(y_0),y_0) + \\ + \int_{a(y_0)}^{b(y_0)} \frac{\partial f}{\partial y}(x,y_0)dx.$$

Gamma function

Definition 5.80. For $x > 0$, *Gamma function* is defined as

$$\Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

Theorem 5.81. Gamma function is a generalization of the factorial. In fact, for $x > 0$ we have

$$\Gamma(x+1) = x\Gamma(x).$$

In particular, $\Gamma(n+1) = n!$ for all $n \in \mathbb{N}$.

Theorem 5.82. Gamma function satisfies:

$$\lim_{x\to\infty} \frac{\Gamma(x+1)}{(x/e)^x\sqrt{2\pi x}} = 1.$$

Corollary 5.83 (Stirling's formula).

$$\lim_{n\to\infty} \frac{n!}{n^n e^{-n}\sqrt{2\pi n}} = 1.$$

2.5.4 | Fourier series

Periodic functions

Definition 5.84. Let $f : \mathbb{R} \to \mathbb{C}$ be a function. We say that f is *T-periodic*, or is *periodic with period T*, being $T > 0$, if $f(x+T) = f(x)$ for all $x \in \mathbb{R}$.

Lemma 5.85. Let $f : \mathbb{R} \to \mathbb{C}$ be a T-periodic function. Then $f(x+T') = f(x)$ for all $x \in \mathbb{R}$ if and only if $T' = kT$ for some $k \in \mathbb{Z}$.

Proposition 5.86. Let $f : \mathbb{R} \to \mathbb{C}$ be a T-periodic function. Then

$$\int_a^{a+T} f(x)dx = \int_0^T f(x)dx,$$

where $a \in \mathbb{R}$. In particular,

$$\int_a^{a+kT} f(x)dx = k\int_0^T f(x)dx.$$

Lemma 5.87. Let $f : \mathbb{R} \to \mathbb{C}$ be a T-periodic continuous function. Then, $|f|$ is bounded.

Proposition 5.88. Given a T-periodic function f, there is no power series uniformly convergent to f on $\mathbb{R}$.

Orthogonal systems

Definition 5.89. Let $f : \mathbb{R} \to \mathbb{C}$ be a function. Then $f \in L^p(I)$, $p \geq 1$, if

$$\|f\|_p := \left(\int_I |f(t)|^p dt\right)^{1/p} < \infty.$$

Definition 5.90. Let $f, g : [a,b] \to \mathbb{C}$ be Riemann-integrable functions. We define the *inner product of f and g* as

$$\langle f, g\rangle := \int_a^b f(x)\overline{g(x)}dx,$$

where $\overline{g}$ is the complex conjugate of g. Now, it's natural to define the *norm of f* as

$$\|f\| := \langle f, f\rangle^{1/2} = \left(\int_a^b |f(x)|^2 dx\right)^{1/2} = \|f\|_2.$$

And the *distance between f and g* as

$$d(f,g) := \|f - g\|.$$

Proposition 5.91. Let $f, g : [a,b] \to \mathbb{C}$ be Riemann-integrable functions and let $\alpha \in \mathbb{C}$. Then we have:

1. $\langle f, f\rangle \geq 0$.
2. $\langle f + h, g\rangle = \langle f, g\rangle + \langle h, g\rangle$ and $\langle f, g + h\rangle = \langle f, g\rangle + \langle f, h\rangle$.
3. $\langle f, g\rangle = \overline{\langle g, f\rangle}$.
4. $\langle \alpha f, g\rangle = \alpha\langle f, g\rangle$ and $\langle f, \alpha g\rangle = \overline{\alpha}\langle f, g\rangle$.

Theorem 5.92 (Cauchy–Schwarz inequality). Let $f, g : [a,b] \to \mathbb{C}$ be Riemann-integrable functions. Then,

$$|\langle f, g\rangle| \leq \|f\| \cdot \|g\|,$$

which can be written as

$$\int_a^b f\overline{g} \leq \left(\int_a^b |f|^2\right)^{1/2}\left(\int_a^b |g|^2\right)^{1/2}.$$

Theorem 5.93 (Minkowski inequality). Let $f, g : [a,b] \to \mathbb{C}$ be Riemann-integrable functions. Then,

$$\|f + g\| \leq \|f\| + \|g\|.$$

Definition 5.94. Let $f, g : [a,b] \to \mathbb{C}$ be Riemann-integrable functions with $f \neq g$. We say f and g are *orthogonal* if $\langle f, g\rangle = 0$. We say f and g are *orthonormal* if they are orthogonal and $\|f\| = \|g\| = 1$.

Definition 5.95. Let $S = \{\phi_0, \phi_1, \ldots\}$ be a collection of Riemann-integrable functions on $[a,b]$. We say S is an *orthonormal system* if $\|\phi_n\| = 1$ $\forall n$ and $\langle \phi_n, \phi_m\rangle = 0$ $\forall n \neq m$.

Proposition 5.96. Let

$$S_1 = \left\{\frac{1}{T}e^{\frac{2\pi i n x}{T}}, n \in \mathbb{Z}\right\},$$

$$S_2 = \left\{\frac{1}{\sqrt{T}}, \frac{\cos\left(\frac{2\pi n x}{T}\right)}{\sqrt{T/2}}, \frac{\sin\left(\frac{2\pi m x}{T}\right)}{\sqrt{T/2}}, n, m \in \mathbb{N}\right\}.$$

Then S_1 and S_2 orthonormal systems on $[-T/2, T/2]$.

Definition 5.97. A collection of functions $S = \{\phi_0, \phi_1, \ldots, \phi_n\}$ is *linearly dependent* on $[a,b]$ if there exist $c_0, c_1, \ldots, c_n \in \mathbb{R}$ not all zero, such that

$$c_0\phi_0 + c_1\phi_1 + \cdots + c_n\phi_n = 0, \quad \forall x \in [a,b].$$

Otherwise we say S is *linearly independent*. If the collection S has an infinity number of functions, we say S is linearly independent on $[a,b]$ if any finite subset of S is linearly independent on $[a,b]$.

Theorem 5.98. Let $S = \{\phi_0, \phi_1, \ldots\}$ be an orthonormal system on $[a,b]$. Suppose that $\sum c_n\phi_n(x)$ converges uniformly to a function f on $[a,b]$. Then, f is Riemann-integrable on $[a,b]$ and, moreover,

$$c_n = \langle f, \phi_n\rangle = \int_a^b f(x)\overline{\phi_n(x)}dx, \quad \forall n \geq 0.$$

Fourier coefficients and Fourier series

Definition 5.99. Let $S = \left\{e^{\frac{2\pi i n x}{T}}/T, n \in \mathbb{Z}\right\}$ be an orthonormal system on $[-T/2, T/2]$ and let $f \in L^1([-T/2, T/2])$[65] be a T-periodic function[66]. We define the *n-th Fourier coefficient of f* as

$$\widehat{f}(n) = \left\langle f, e^{\frac{2\pi i n x}{T}}/T\right\rangle = \frac{1}{T}\int_{-T/2}^{T/2} f(x)e^{-\frac{2\pi i n x}{T}}dx,$$

for all $n \in \mathbb{Z}$.

Proposition 5.100. Let $f, g \in L^1([-T/2, T/2])$. The following properties are satisfied:

1. For all $\lambda, \mu \in \mathbb{C}$,
$$\widehat{\lambda f + \mu g}(n) = \lambda\widehat{f}(n) + \mu\widehat{g}(n).$$
2. Let $\tau \in \mathbb{R}$. We define $f_\tau(x) = f(x - \tau)$. Then,
$$\widehat{f_\tau}(n) = e^{-\frac{2\pi i n \tau}{T}}\widehat{f}(n).$$
3. If f is even, then $\widehat{f}(n) = \widehat{f}(-n)$, $\forall n \in \mathbb{Z}$.
If f is odd, then $\widehat{f}(n) = -\widehat{f}(-n)$, $\forall n \in \mathbb{Z}$.
4. If $f \in \mathcal{C}^k$, then
$$\widehat{f^{(k)}}(n) = \left(\frac{2\pi i n}{T}\right)^k \widehat{f}(n).$$
5. $\widehat{(f * g)}(n) = \widehat{f}(n)\widehat{g}(n)$.

Definition 5.101. Let $f \in L^1([-T/2, T/2])$. We define the *Fourier series of f* as

$$Sf(x) = \sum_{n\in\mathbb{Z}} \widehat{f}(n)e^{\frac{2\pi i n x}{T}}.$$

[65] Saying that $f \in L^1([-T/2, T/2])$ is equivalent to say that f is integrable on $[-T/2, T/2$.
[66] From now on, we will work only with functions defined on $[-T/2, T/2]$ and extended periodically on $\mathbb{R}$.

Definition 5.102. Let $f \in L^1([-T/2, T/2])$ and Sf be the Fourier series of f. The *N-th partial sum of Sf* is

$$S_N f(x) = \sum_{n=-N}^{N} \widehat{f}(n) e^{\frac{2\pi i n x}{T}}.$$

Proposition 5.103. Let $f \in L^1([-T/2, T/2])$. Then

$$Sf(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n x}{T}\right) + b_n \sin\left(\frac{2\pi n x}{T}\right),$$

where

$$a_n = \frac{2}{T}\int_{-T/2}^{T/2} f(x)\cos\left(\frac{2\pi n x}{T}\right) dx,$$
$$b_n = \frac{2}{T}\int_{-T/2}^{T/2} f(x)\sin\left(\frac{2\pi n x}{T}\right) dx,$$

for $n \geq 0$[67]. In particular, if f is even we have

$$Sf(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n x}{T}\right),$$

and if f is odd we have

$$Sf(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi n x}{T}\right).$$

Definition 5.104. Let $f : (0, L) \to \mathbb{C}$ be a function. We define the *even extension of f* as

$$\tilde{f}(x) = \begin{cases} f(x) & \text{si} \quad x \in (0, L) \\ f(-x) & \text{si} \quad x \in (-L, 0) \end{cases}$$

Analogously, we define the *odd extension of f* as

$$\hat{f}(x) = \begin{cases} f(x) & \text{si} \quad x \in (0, L) \\ -f(-x) & \text{si} \quad x \in (-L, 0) \end{cases}$$

Proposition 5.105. Let $f \in L^1([0, T/2])$. If we make the even extension of f[68], then

$$Sf(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{2\pi n x}{T}\right),$$

where $a_n = \dfrac{4}{T}\displaystyle\int_0^{T/2} f(x)\cos\left(\frac{2\pi n x}{T}\right) dx$ for $n \geq 0$. If we make the odd extension of f, then

$$Sf(x) = \sum_{n=1}^{\infty} b_n \sin\left(\frac{2\pi n x}{T}\right),$$

where $b_n = \dfrac{4}{T}\displaystyle\int_0^{T/2} f(x)\sin\left(\frac{2\pi n x}{T}\right) dx$ for $n \geq 1$.

[67] The relation between a_n, b_n and $\widehat{f}(n)$ is given by:

$$a_n = \widehat{f}(n) + \widehat{f}(-n) \quad \text{and} \quad b_n = i\left[\widehat{f}(n) - \widehat{f}(-n)\right], \quad \forall n \in \mathbb{N} \cup \{0\}.$$

[68] For simplicity, when we have a function f and make its even or odd extension, we will still call its even or odd extension f instead of $\tilde{f}$ or $\hat{f}$.

Pointwise convergence

Definition 5.106 (Dirichlet kernel). We define the *Dirichlet kernel of order N* as

$$D_N(t) = \frac{1}{T}\sum_{n=-N}^{N} e^{\frac{2\pi i n t}{T}} = \frac{1}{T}\frac{\sin\left(\frac{(2N+1)\pi t}{T}\right)}{\sin\left(\frac{\pi t}{T}\right)}.$$

Proposition 5.107. The Dirichlet kernel has the following properties:

1. D_N is a T-periodic and even function.
2. $\displaystyle\int_0^T D_N(t)dt = 1, \ \forall N.$

Proposition 5.108. Let $f \in L^1([-T/2, T/2])$. Then

$$S_N f(x) = (f * D_N)(x) = \int_{-T/2}^{T/2} f(x-t) D_N(t) dt = \\ = \int_0^{T/2} [f(x+t) + f(x-t)] D_N(t) dt.$$

Lemma 5.109 (Riemann-Lebesgue lemma). Let $f \in L^1([-T/2, T/2])$ and $\lambda \in \mathbb{R}$. Then:

$$\lim_{\lambda\to\infty} \int_{-T/2}^{T/2} f(t)\sin(\lambda t)dt = \lim_{\lambda\to\infty} \int_{-T/2}^{T/2} f(t)\cos(\lambda t)dt = 0.$$

In particular, $\displaystyle\lim_{|n|\to\infty} \widehat{f}(n) = 0$.

Theorem 5.110. Let $f \in L^1([-T/2, T/2])$ be a function left and right differentiable at x_0, that is, there exists the following limits

$$f'(x_0^+) = \lim_{t\to 0^+} \frac{f(x_0+t) - f(x_0^+)}{t},$$
$$f'(x_0^-) = \lim_{t\to 0^-} \frac{f(x_0+t) - f(x_0^-)}{t},$$

(supposing the existence of left- and right-sided limits). Then,

$$\lim_{N\to\infty} S_N f(x_0) = \frac{f(x_0^+) + f(x_0^-)}{2}.$$

Theorem 5.111 (Dini's theorem). Let $f \in L^1([-T/2, T/2])$, $x_0 \in (-T/2, T/2)$ and $\ell \in \mathbb{R}$ such that

$$\int_0^{\delta} \frac{|f(x_0+t) + f(x_0-t) - 2\ell|}{t} dt < \infty$$

for some $\delta > 0$. Then $\displaystyle\lim_{N\to\infty} S_N f(x_0) = \ell$.

Theorem 5.112 (Lipschitz's theorem). Let $f \in L^1([-T/2,T/2])$ such that at a point $x_0 \in (-T/2,T/2)$ it satisfies

$$|f(x_0+t)-f(x_0)| \leq k|t|$$

for some constant $k \in \mathbb{R}$ and for $|t| < \delta$. Then $\lim_{N\to\infty} S_N f(x_0) = f(x_0)$.

Uniform convergence

Definition 5.113. Let $\sum a_n$ be a series with partial sums S_k. The series $\sum a_n$ is called *Cesàro summable* with sum S if

$$\lim_{N\to\infty} \frac{S_1+\cdots+S_N}{N} = S.$$

Definition 5.114 (Fejer kernel). We define the *Fejer kernel of order* N as

$$K_N(t) = \frac{1}{N+1}\sum_{k=0}^{N} D_k(t) = \frac{1}{T(N+1)}\frac{\sin^2\left(\frac{(N+1)\pi t}{T}\right)}{\sin^2\left(\frac{\pi t}{T}\right)},$$

being $D_k(t)$ the Dirichlet kernel of order k, $0 \leq k \leq N$.

Proposition 5.115. The Fejer kernel has the following properties:

1. K_N is a T-periodic, even and non-negative function.
2. $\int_{-T/2}^{T/2} K_N(t)dt = 1$, $\forall N$.
3. $\forall \delta > 0$, $\lim_{N\to\infty} \sup\{|K_N(t)| : \delta \leq |t| \leq T/2\} = 0$.

Definition 5.116. Let $f \in L^1([-T/2,T/2])$. We define the *Fejér means* $\sigma_N f$, for all $N \in \mathbb{N}$, as

$$\sigma_N f(x) = \frac{S_0 f(x)+\cdots+S_N f(x)}{N+1}.$$

Proposition 5.117. Let $f \in L^1([-T/2,T/2])$. Then

$$\sigma_N f(x) = (f*K_N)(x) = \int_{-T/2}^{T/2} f(x-t)K_N(t)dt = \\ = \int_0^{T/2} [f(x+t)+f(x-t)]K_N(t)dt.$$

Theorem 5.118 (Fejér's theorem). Let $f \in L^1([-T/2,T/2])$ be a function having left- and right-sided limits at point x_0. Then,

$$\lim_{N\to\infty} \sigma_N f(x_0) = \frac{f(x_0^+)+f(x_0^-)}{2}.$$

In particular, if f is continuous at x_0, $\lim_{N\to\infty} \sigma_N f(x_0) = f(x_0)$.

Theorem 5.119 (Fejér's theorem). Let f be a continuous function on $[-T/2,T/2]$. Then $\sigma_N f$ converges uniformly to f on $[-T/2,T/2]$.

Corollary 5.120. Let f be a continuous function on $[-T/2,T/2]$. Then there exists a sequence of trigonometric polynomials that converge uniformly to f on $[-T/2,T/2]$. In fact,

$$\sigma_N f(x) = \sum_{k=-N}^{N}\left(1-\frac{|k|}{N+1}\right)\widehat{f}(k)e^{2\pi ikx}.$$

Corollary 5.121. Let f and g be continuous functions on $[-T/2,T/2]$ such that $Sf(x) = Sg(x)$. Then $f = g$.

Convergence in norm

Definition 5.122. We say a sequence (f_N) *converge to* f *in norm* L^p if $\lim_{N\to\infty} \|f_N - f\|_p = 0$.

Theorem 5.123. Let $f \in L^2([-T/2,T/2])$. Then, $\lim_{N\to\infty} \|\sigma_N f - f\| = 0$.

Corollary 5.124. Let $f \in L^1([-T/2,T/2])$. Then $\lim_{N\to\infty} \|\sigma_N f - f\|_1 = 0$.

Corollary 5.125. Let $f,g \in L^1([-T/2,T/2])$ be functions such that $Sf(x) = Sg(x)$. Then $\lim_{N\to\infty} \|g-f\|_1 = 0$.

Theorem 5.126 (Bessel's inequality). Let $f \in L^2(I)$, where I is any interval on the real line. Then:

$$T\sum_{n=-N}^{N} |\widehat{f}(n)|^2 \leq \|f\|^2,$$

$$\frac{T}{2}\left(\frac{|a_0|^2}{2}+\sum_{n=1}^{N}|a_n|^2+|b_n|^2\right) \leq \|f\|^2,$$

for all $N \in \mathbb{N}$.

Theorem 5.127. $S_N f$ is the trigonometric polynomial of degree N that best approximates f in norm L^2.

Corollary 5.128. Let $f \in L^2([-T/2,T/2])$. Then, $\lim_{N\to\infty} \|S_N f - f\| = 0$.

Theorem 5.129 (Parseval's identity). Let $f,g \in L^2([-T/2,T/2])$ be bounded functions. Then

$$\langle f,g\rangle = T\sum_{n\in\mathbb{Z}} \widehat{f}(n)\overline{\widehat{g}(n)}.$$

In particular, if $f = g$:

$$\|f\|^2 = T\sum_{n\in\mathbb{Z}} |\widehat{f}(n)|^2,$$

$$\|f\|^2 = \frac{T}{2}\left(\frac{|a_0|^2}{2}+\sum_{n=1}^{\infty}|a_n|^2+|b_n|^2\right).$$

Applications of Fourier series

Theorem 5.130 (Wirtinger's inequality). Let f be a function such that $f(0) = f(T)$, $f' \in L^2([0,T])$ and $\int_a^b f(t)dt = 0$. Then,

$$\int_0^T |f(x)|^2 dx \leq \frac{T^2}{4\pi^2} \int_0^T |f'(x)|^2 dx,$$

with equality if and only if

$$f(x) = A\cos\left(\frac{2\pi x}{T}\right) + B\sin\left(\frac{2\pi x}{T}\right).$$

Theorem 5.131 (Wirtinger's inequality). Let $f \in \mathcal{C}^1([a,b])$ with $f(a) = f(b) = 0$. Then,

$$\int_a^b |f(x)|^2 dx \leq \frac{(b-a)^2}{\pi^2} \int_a^b |f'(x)|^2 dx.$$

Theorem 5.132 (Isoperimetric inequality). Let c be a simple and closed curve of class $\mathcal{C}^1$ whose length is L. If A_c is the area enclosed by c, then

$$A_c \leq \frac{L^2}{4\pi},$$

with equality if and only if c is a circle.

2.6 Numerical methods

2.6.1 | Errors

Floating-point representation

Theorem 6.1. Let $b \in \mathbb{N}$, $b \geq 2$. Any real number $x \in \mathbb{R}$ can be represented of the form

$$x = s\left(\sum_{i=1}^{\infty} \alpha_i b^{-i}\right) b^q,$$

where $s \in \{-1, 1\}$, $q \in \mathbb{Z}$ and $\alpha_i \in \{0, 1, \ldots, b-1\}$. Moreover, this representation is unique if $\alpha_1 \neq 0$ and $\forall i_0 \in \mathbb{N}$, $\exists i \geq i_0 : \alpha_i \neq b-1$. We will write

$$x = s(0.\alpha_1\alpha_2\cdots)_b b^q,$$

where the subscript b in the parenthesis indicates that the number $0.\alpha_1\alpha_2\alpha_3\cdots$ is in base b.

Definition 6.2 (Floating-point representation). Let x be a real number. Then the floating-point representation of x is

$$x = s\left(\sum_{i=1}^{t} \alpha_i b^{-i}\right) b^q.$$

Here s is called the *sign*; $\sum_{i=1}^{t} \alpha_i b^{-i}$, the *significant* or *mantissa*, and q, the *exponent*, limited to a prefixed range $q_{\min} \leq q \leq q_{\max}$. So, the floating-point representation of x is

$$x = smb^q = s(0.\alpha_1\alpha_2\cdots\alpha_t)_b b^q.$$

Finally we say a floating-point number is *normalized* if $\alpha_1 \neq 0$.

Definition 6.3. Let $x \in \mathbb{R}$ be such that $x = s(0.\alpha_1\alpha_2\cdots)_b b^q$ with $q_{\min} \leq q \leq q_{\max}$. We say the *floating-point representation by truncation of* x is

$$fl_T(x) = s(0.\alpha_1\alpha_2\cdots\alpha_t)_b b^q.$$

We say the *floating-point representation by rounding of* x is

$$fl_R(x) = \begin{cases} s(0.\alpha_1\cdots\alpha_t)_b b^q & \text{if } 0 \leq \alpha_{t+1} < \dfrac{b}{2} \\ s(0.\alpha_1\cdots\alpha_{t-1}(\alpha_t+1))_b b^q & \text{if } \dfrac{b}{2} \leq \alpha_{t+1} \leq b-1. \end{cases}$$

Definition 6.4. Given a value $x \in \mathbb{R}$ and an approximation $\tilde{x}$ of x, the *absolute error* is

$$\Delta x := |x - \tilde{x}|.$$

If $x \neq 0$, the *relative error* is

$$\delta x := \frac{|x - \tilde{x}|}{x}.$$

If x is unknown, we take

$$\delta x \approx \frac{|x - \tilde{x}|}{\tilde{x}}.$$

Definition 6.5. Let $\tilde{x}$ be an approximation of x. If $\Delta x \leq \frac{1}{2}10^{-t}$, we say $\tilde{x}$ *has* t *correct decimal digits*. If $x = sm10^q$ with $0.1 \leq m < 1$, $\tilde{x} = s\tilde{m}10^q$ and

$$u := \max\{i \in \mathbb{Z} : |m - \tilde{m}| \leq \frac{1}{2}10^{-i}\},$$

then we say that $\tilde{x}$ *has* u *significant digits*.

Proposition 6.6. Let $x \in \mathbb{R}$ be such that $x = s(0.\alpha_1\alpha_2\cdots)_b b^q$ with $\alpha_1 \neq 0$ and $q_{\min} \leq q \leq q_{\max}$. Then, its floating-point representation in base b and with t digits satisfy:

$$|fl_T(x) - x| \leq b^{q-t}, \qquad |fl_R(x) - x| \leq \frac{1}{2}b^{q-t}.$$
$$\left|\frac{fl_T(x) - x}{x}\right| \leq b^{1-t}, \qquad \left|\frac{fl_R(x) - x}{x}\right| \leq \frac{1}{2}b^{1-t}.$$

Definition 6.7. The *machine epsilon* ϵ is defined as

$$\epsilon := \min\{\varepsilon > 0 : fl(1+\varepsilon) \neq 1\}.$$

Proposition 6.8. For a machine working by truncation, $\epsilon = b^{1-t}$. For a machine working by rounding, $\epsilon = \frac{1}{2}b^{1-t}$.

Propagation of errors

Proposition 6.9 (Propagation of absolute errors). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathcal{C}^2$. If Δx_j is the absolute error of the variable x_j and $\Delta f(x)$ is the absolute error of the function f evaluated at the point $x = (x_1, \ldots, x_n)$, we have

$$|\Delta f(x)| \lesssim \sum_{j=1}^{n} \left|\frac{\partial f}{\partial x_j}(x)\right| |\Delta x_j|^{69}.$$

The coefficients $\left|\frac{\partial f}{\partial x_j}(x)\right|$ are called *absolute condition numbers of the problem.*

Proposition 6.10 (Propagation of relative errors). Let $f : \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathcal{C}^2$. If δx_j is the relative error of the variable x_j and $\delta f(x)$ is the relative error of the function f evaluated at the point $x = (x_1, \ldots, x_n)$, we have

$$|\delta f(x)| \lesssim \sum_{j=1}^{n} \frac{\left|\frac{\partial f}{\partial x_j}(x)\right| |x_j|}{|f(x)|} |\delta x_j|.$$

The coefficients $\frac{\left|\frac{\partial f}{\partial x_j}(x)\right||x_j|}{|f(x)|}$ are called *relative condition numbers of the problem.*

[69] The symbol $\lesssim$ means that we are omitting terms of order $\Delta x_j \Delta x_k$ and higher.

Numerical stability of algorithms

Definition 6.11. An algorithm is said to be *numerically stable* if errors in the input lessen in significance as the algorithm executes, having little effect on the final output. On the other hand, an algorithm is said to be *numerically unstable* if errors in the input cause a considerably larger error in the final output.

Definition 6.12. A problem with a low condition number is said to be *well-conditioned*. Conversely, a problem with a high condition number is said to be *ill-conditioned*.

2.6.2 | Zeros of functions

Definition 6.13. Let $f : \mathbb{R} \to \mathbb{R}$ be a function. We say α is a *zero* or a *solution to the equation* $f(x) = 0$ if $f(\alpha) = 0$.

Definition 6.14. Let $f : \mathbb{R} \to \mathbb{R}$ be a sufficiently differentiable function. We say α is a *zero of multiplicity* $m \in \mathbb{N}$ if

$$f(\alpha) = f'(\alpha) = \cdots = f^{(m-1)}(\alpha) = 0 \quad \text{and} \quad f^{(m)}(\alpha) \neq 0.$$

If $m = 1$, the zero is called *simple*; if $m = 2$, *double*; if $m = 3$, *triple*...

Root-finding methods

For the following methods consider a continuous function $f : I \subseteq \mathbb{R} \to \mathbb{R}$ with an unknown zero $\alpha \in I$. Given $\varepsilon > 0$, we want to approximate α with $\tilde{\alpha}$ such that $|\alpha - \tilde{\alpha}| < \varepsilon$.

Theorem 6.15 (Bisection method). Suppose $I = [a_0, b_0]$. For each step $n \geq 0$ of the algorithm we will approximate α by

$$c_n = \frac{a_n + b_n}{2}.$$

If $f(c_n) = 0$ we are done. If not, let

$$[a_{n+1}, b_{n+1}] = \begin{cases} [a_n, c_n] & \text{if} \quad f(a_n)f(c_n) < 0, \\ [c_n, b_n] & \text{if} \quad f(a_n)f(c_n) > 0. \end{cases}$$

and iterate the process again[70]. Observe the length of the interval $[a_n, b_n]$ is $\frac{b_0 - a_0}{2^n}$ and therefore:

$$|\alpha - c_n| < \frac{b_0 - a_0}{2^{n+1}} < \varepsilon \iff n > \frac{\log\left(\frac{b_0 - a_0}{\varepsilon}\right)}{\log 2} - 1.$$

Theorem 6.16 (*Regula falsi* method). Suppose $I = [a_0, b_0]$. For each step $n \geq 0$ of the algorithm we will approximate α by

$$c_n = b_n - f(b_n)\frac{b_n - a_n}{f(b_n) - f(a_n)} = \frac{a_n f(b_n) - b_n f(a_n)}{f(b_n) - f(a_n)}.$$

If $f(c_n) = 0$ we are done. If not, let

$$[a_{n+1}, b_{n+1}] = \begin{cases} [a_n, c_n] & \text{if} \quad f(a_n)f(c_n) < 0, \\ [c_n, b_n] & \text{if} \quad f(a_n)f(c_n) > 0, \end{cases}$$

and iterate the process again.

Theorem 6.17 (Secant method). Suppose $I = \mathbb{R}$ and that we have two different initial approximations x_0, x_1. Then for each step $n \geq 0$ of the algorithm we obtain a new approximation x_{n+2}, given by:

$$x_{n+2} = x_{n+1} - f(x_{n+1})\frac{x_{n+1} - x_n}{f(x_{n+1}) - f(x_n)}.$$

Theorem 6.18 (Newton-Raphson method). Suppose $I = \mathbb{R}$, $f \in \mathcal{C}^1$ and that we have an initial approximation x_0. Then for each step $n \geq 0$ we obtain a new approximation x_{n+1}, given by:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}.$$

Theorem 6.19 (Newton-Raphson modified method). Suppose $I = \mathbb{R}$, $f \in \mathcal{C}^1$ and that we have an initial approximation x_0 of a zero α of multiplicity m. Then for each step $n \geq 0$ we obtain a new approximation x_{n+1}, given by:

$$x_{n+1} = x_n - m\frac{f(x_n)}{f'(x_n)}.$$

Theorem 6.20 (Chebyshev method). Suppose $I = \mathbb{R}$, $f \in \mathcal{C}^2$ and that we have an initial approximation x_0. Then for each step $n \geq 0$ we obtain a new approximation x_{n+1}, given by:

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} - \frac{1}{2}\frac{[f(x_n)]^2 f''(x_n)}{[f'(x_n)]^3}.$$

Fixed-point iterations

Definition 6.21. Let $g : [a, b] \to [a, b] \subset \mathbb{R}$ be a function. A point $\alpha \in [a, b]$ is *n-periodic* if $g^n(\alpha) = \alpha$ and $g^j(\alpha) \neq \alpha$ for $j = 1, \ldots, n-1$[71].

Theorem 6.22 (Fixed-point theorem). Let (M, d) be a complete metric space and $g : M \to M$ be a contraction[72]. Then g has a unique fixed point $\alpha \in M$ and for every $x_0 \in M$,

$$\lim_{n \to \infty} x_n = \alpha, \quad \text{where } x_n = g(x_{n-1}) \quad \forall n \in \mathbb{N}.$$

Proposition 6.23. Let (M, d) be a metric space and $g : M \to M$ be a contraction of constant k. Then if

[70] Note that bisection method only works for zeros of odd multiplicity.
[71] Note that 1-periodic points are the fixed points of f.
[72] Remember definitions 3.1, 3.36 and 3.52.

we want to approximate a fixed point α by the iteration $x_n = g(x_{n-1})$, we have:

$$d(x_n, \alpha) \leq \frac{k^n}{1-k} d(x_1, x_0) \qquad \text{(a priori estimation)}$$

$$d(x_n, \alpha) \leq \frac{k}{1-k} d(x_n, x_{n-1}) \qquad \text{(a posteriori estimation)}$$

Corollary 6.24. Let $g : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^1$. Suppose α is a fixed point of g and $|g'(\alpha)| < 1$. Then, there exists $\varepsilon > 0$ and $I_\varepsilon := [\alpha - \varepsilon, \alpha + \varepsilon]$ such that $g(I_\varepsilon) \subseteq I_\varepsilon$ and g is a contraction on I_ε. In particular, if $x_0 \in I_\varepsilon$, the iteration $x_{n+1} = g(x_n)$ converges to α.

Definition 6.25. Let $g : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^1$ and α be a fixed point of g. We say α is an *attractor fixed point* if $|g'(\alpha)| < 1$. In this case, any iteration $x_{n+1} = g(x_n)$ in I_ε converges to α. If $|g'(\alpha)| > 1$, we say α is a *repulsor fixed point*. In this case, $\forall x_0 \in I_\varepsilon$ the iteration $x_{n+1} = g(x_n)$ doesn't converge to α.

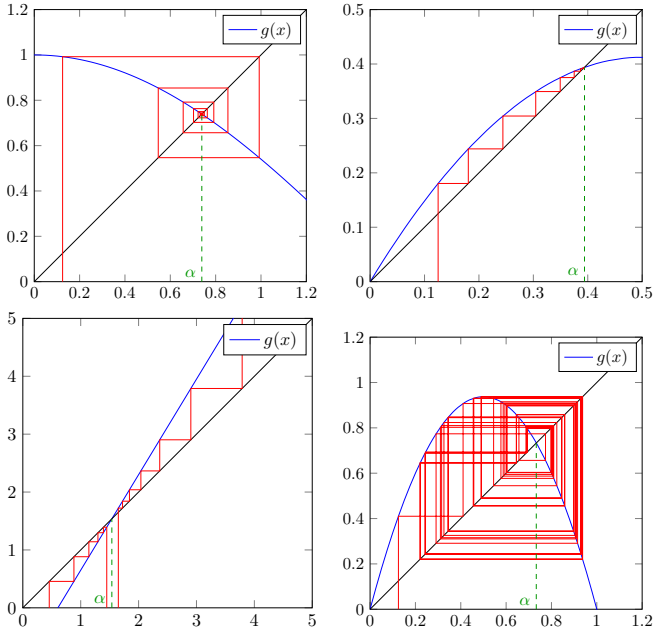

Figure 2.9: Cobweb diagrams. In the figures at the top, α is a attractor point, that is, $|g'(\alpha)| < 1$. More precisely, the figure at the top left occurs when $-1 < g'(\alpha) \leq 0$ and the figure at the top right when $0 \leq g'(\alpha) < 1$. In the figure at bottom left, α is a repulsor point. Finally, in the figure at bottom right the iteration $x_{n+1} = g(x_n)$ has no limit. It is said that to have a *chaotic behavior*.

Order of convergence

Definition 6.26 (Order of convergence). Let (x_n) be a sequence of real numbers that converges to $\alpha \in \mathbb{R}$. We say (x_n) has *order of convergence* $p \in \mathbb{R}^+$ if exists $C > 0$ such that:

$$\lim_{n\to\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = C.$$

The constant C is called *asymptotic error constant*. For the case $p = 1$, we need $C < 1$. In this case the convergence is called *linear convergence*; for $p = 2$, is called *quadratic convergence*; for $p = 3$, *cubic convergence*... If it's satisfied that

$$\lim_{n\to\infty} \frac{|x_{n+1} - \alpha|}{|x_n - \alpha|^p} = 0$$

for some $p \in \mathbb{R}^+$, we say the sequence has *order of convergence at least p*.

Theorem 6.27. Let $g : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^p$ and let α be a fixed point of g. Suppose

$$g'(\alpha) = g''(\alpha) = \cdots = g^{(p-1)}(\alpha) = 0$$

with $|g'(\alpha)| < 1$ if $p = 1$. Then the iteration $x_{n+1} = g(x_n)$, with x_0 sufficiently close to α, has order of convergence at least p. If, moreover, $g^{(p)}(\alpha) \neq 0$, then the previous iteration has order of convergence p with asymptotic error constant $C = \frac{|g^{(p)}(\alpha)|}{p!}$.

Theorem 6.28. Let $f : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^3$ and α be a simple zero of f. If $f''(\alpha) \neq 0$, then Newton-Raphson method for finding α has quadratic convergence with asymptotic error constant $C = \frac{1}{2}\left|\frac{f''(\alpha)}{f'(\alpha)}\right|$.
If $f \in \mathcal{C}^{m+2}$, and α is a zero of multiplicity $m > 1$, then Newton-Raphson method has linear convergence but Newton-Raphson modified method has at least quadratic convergence.

Theorem 6.29. Let $f : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^3$ and let α be a simple zero of f. Then Chebyshev's method for finding α has at least cubic convergence.

Definition 6.30. We define the *computational efficiency of an algorithm* as a function $E(p, t)$, where t is the time taken for each iteration of the method and p is the order of convergence of the method. $E(p, t)$ must satisfy the following properties:

1. $E(p, t)$ is increasing with respect to the variable p and decreasing with respect to t.
2. $E(p, t) = E(p^m, mt)\ \forall m \in \mathbb{R}$.

Examples of such functions are the following:

$$E(p, t) = \frac{\log p}{t}, \quad E(p, t) = p^{1/t}.$$

Sequence acceleration

Definition 6.31 (Aitken's Δ^2 method). Let (x_n) be a sequence of real numbers. We denote:

$$\Delta x_n := x_{n+1} - x_n,$$
$$\Delta^2 x_n := \Delta x_{n+1} - \Delta x_n = x_{n+2} - 2x_{n+1} + x_n.$$

Aitken's Δ^2 method is the transformation of the sequence (x_n) into a sequence y_n, defined as:

$$y_n := x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n} = x_n - \frac{(x_{n+1} - x_n)^2}{x_{n+2} - 2x_{n+1} + x_n},$$

with $y_0 = x_0$.

Theorem 6.32. Let (x_n) be a sequence of real numbers such that $\lim_{n\to\infty} x_n = \alpha$, $x_n \neq \alpha\ \forall n \in \mathbb{N}$ and $\exists C$, $|C| < 1$, satisfying

$$x_{n+1} - \alpha = (C + \delta_n)(x_n - \alpha), \quad \text{with } \lim_{n\to\infty} \delta_n = 0.$$

Then the sequence (y_n) obtained from Aitken's Δ^2 process is well-defined and

$$\lim_{n\to\infty} \frac{y_n - \alpha}{x_n - \alpha} = 0^{73}.$$

Theorem 6.33 (Steffensen's method). Let $g : \mathbb{R} \to \mathbb{R}$ be a continuous function and suppose we have an iterative method $x_{n+1} = g(x_n)$. Then for each step n we can consider a new iteration y_{n+1}, with $y_0 = x_0$, given by:

$$y_{n+1} = y_n - \frac{(g(y_n) - y_n)^2}{g(g(y_n)) - 2g(y_n) + y_n}.$$

Proposition 6.34. Let $f : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^2$ and α be a simple zero of f. Then Steffensen's method for finding α has at least quadratic convergence[74].

Zeros of polynomials

Lemma 6.35. Let $p(z) = a_0 + a_1 z + \cdots + a_n z^n \in \mathbb{C}[x]$ with $a_n \neq 0$. We define

$$\lambda := \max\left\{ \left\| \frac{a_i}{a_n} \right\| : i = 0, 1, \ldots, n-1 \right\}.$$

Then if $p(\alpha) = 0$ for some $\alpha \in \mathbb{C}$, $\|\alpha\| \leq \lambda + 1$.

Definition 6.36 (Strum's sequence). Let (f_i), $i = 0, \ldots, n$, be a sequence of continuous functions defined on $[a, b] \subset \mathbb{R}$ and $f : [a, b] \to \mathbb{R}$ be a function of class $\mathcal{C}^1$ such that $f(a)f(b) \neq 0$. We say (f_n) is a *Sturm's sequence* if:

1. $f_0 = f$.
2. If $\alpha \in [a, b]$ satisfies $f_0(\alpha) = 0 \implies f_0'(\alpha) f_1(\alpha) > 0$.
3. For $i = 1, \ldots, n-1$, if $\alpha \in [a, b]$ satisfies $f_i(\alpha) = 0 \implies f_{i-1}(\alpha) f_{i+1}(\alpha) < 0$.
4. $f_n(x) \neq 0\ \forall x \in [a, b]$.

Definition 6.37. Let (a_i), $i = 0, \ldots, n$, be a sequence. We define $\nu(a_i)$ as the number of sign variations of the sequence

$$\{a_0, a_1, \ldots, a_n\},$$

without taking into account null values.

Theorem 6.38 (Sturm's theorem). Let $f : [a, b] \to \mathbb{R}$ be a function of class $\mathcal{C}^1$ such that $f(a)f(b) \neq 0$ and with a finite number of zeros. Let (f_i), $i = 0, \ldots, n$, be a Sturm sequence defined on $[a, b]$. Then the number of zeros of f on $[a, b]$ is

$$\nu(f_i(a)) - \nu(f_i(b)).$$

Lemma 6.39. Let $p \in \mathbb{C}[x]$ be a polynomial. Then the polynomial $q = \dfrac{p}{\gcd(p, p')}$ has the same roots as p but all of them are simple.

Proposition 6.40. Let $p \in \mathbb{R}[x]$ be a polynomial with $\deg p = m$. We define $f_0 = \dfrac{p}{\gcd(p, p')}$ and $f_1 = f_0'$. If $\deg f_0 = n$, then for $i = 0, 1, \ldots, n-2$, we define f_{i+2} as

$$f_i(x) = q_{i+1}(x) f_{i+1}(x) - f_{i+2}(x),$$

(similarly to the euclidean division between f_i and f_{i+1}). Then f_n is constant and hence the sequence (f_i), $i = 0, \ldots, n$, is a Sturm sequence.

Theorem 6.41 (Budan-Fourier theorem). Let $p \in \mathbb{R}[x]$ be a polynomial with $\deg p = n$. Consider the sequence $(p^{(i)})$, $i = 0, \ldots, n$. If $p(a)p(b) \neq 0$, the number of zeros of p on $[a, b]$ is

$$\nu\left(p^{(i)}(a)\right) - \nu\left(p^{(i)}(b)\right) - 2k, \quad \text{for some } k \in \mathbb{N} \cup \{0\}.$$

Corollary 6.42 (Descartes' rule of signs). Let $p = a_0 + a_1 x + \cdots + a_n x^n \in \mathbb{R}[x]$ be a polynomial. If $p(0) \neq 0$, the number of zeros of p on $[0, \infty)$ is

$$\nu(a_i) - 2k, \quad \text{for some } k \in \mathbb{N} \cup \{0\}^{75}.$$

Theorem 6.43 (Greshgorin theorem). Let $A = (a_{ij}) \in \mathcal{M}_n(\mathbb{C})$ be a complex matrix and λ be an eigenvalue of A. For all $i, j \in \{1, 2, \ldots, n\}$ we define:

$$r_i = \sum_{\substack{k=1 \\ k\neq i}}^{n} |a_{ik}|, \quad R_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq r_i\},$$

$$c_j = \sum_{\substack{k=1 \\ k\neq j}}^{n} |a_{kj}|, \quad C_j = \{z \in \mathbb{C} : |z - a_{jj}| \leq c_j\}.$$

Then $\lambda \in \bigcup_{i=1}^n R_i$ and $\lambda \in \bigcup_{j=1}^n C_j$. Moreover in each connected component of $\bigcup_{i=1}^n R_i$ (respectively $\bigcup_{j=1}^n C_j$) there are as many eigenvalues (taking into account the multiplicity) as disks R_i (respectively C_i).

Corollary 6.44. Let $p(z) = a_0 + a_1 z + \cdots + a_n z^n + z^{n+1} \in \mathbb{C}[x]$. We define

$$r = \sum_{i=1}^{n-1} |a_i|, \quad c = \max\{|a_0|, |a_1| + 1, \ldots, |a_{n-1}| + 1\}.$$

Then if $p(\alpha) = 0$ for some $\alpha \in \mathbb{C}$,

$$\alpha \in (B(0, 1) \cup B(-a_n, r)) \cap (B(-a_n, 1) \cup B(0, c)).$$

[73] This means that Aitken's Δ^2 method produces an acceleration of the convergence of the sequence (x_n).

[74] Note that the advantage of Steffensen's method over Newton-Raphson method is that in the former we don't need the differentiability of the function whereas in the latter we do.

[75] Note that making the change of variable $t = -x$ one can obtain the number of zeros on $(-\infty, 0]$ of p by considering the polynomial $p(t)$.

2.6.3 | Interpolation

Definition 6.45. Suppose we have a family of real valued functions $\mathfrak{C}$ and a set of points $\{(x_i, y_i)\}_{i=0}^n := \{(x_i, y_i) : x_j \neq x_k \iff j \neq k, i = 0, \dots, n\}$. These points $\{(x_i, y_i)\}_{i=0}^n$ are called *support points*. The *interpolation problem* consists in finding a function $f \in \mathfrak{C}$ such that $f(x_i) = y_i$ for $i = 0, \dots, n$[76].

Polynomial interpolation

Definition 6.46. Given a set of support points $\{(x_i, y_i)\}_{i=0}^n$, *Lagrange's interpolation problem* consists in finding a polynomial $p_n \in \mathbb{R}[x]$ such that $\deg p_n \leq n$ and $p_n(x_i) = y_i$.

Proposition 6.47. Lagrange's interpolation problem has a unique solution and this is:

$$p_n(x) = \sum_{k=0}^{n} y_k \frac{\omega_n(x)}{\omega_n'(x)}, \quad \text{where } \omega_n(x) := \prod_{j=0}^{n}(x - x_j).$$

Proposition 6.48 (Neville's algorithm). Let $P_{i_1,\dots,i_k}(x) \in \mathbb{R}[x]$ be such that $\deg P_{i_0,\dots,i_k} \leq k$ and $P_{i_1,\dots,i_k}(x_{i_j}) = y_{i_j}$ for $j = 0, \dots, k$. Then, it is satisfied that:

1. $P_i(x) = y_i$.
2. $$P_{i_0,\dots,i_k}(x) = \frac{\begin{vmatrix} P_{i_1,\dots,i_k}(x) & x - x_{i_k} \\ P_{i_0,\dots,i_{k-1}}(x) & x - x_{i_0} \end{vmatrix}}{x_{i_k} - x_{i_0}}$$

Definition 6.49. Let $f : \mathbb{R} \to \mathbb{R}$ be a function and $\{x_i\}_{i=0}^k \subset \mathbb{R}$ be pairwise distinct points. We define the *divided difference of order k of f applied to $\{x_i\}_{i=0}^k$*, denoted by $f[x_0, \dots, x_k]$, as the coefficient of x^k of the interpolating polynomial with support points $\{(x_i, f(x_i))\}_{i=0}^k$

Proposition 6.50. Let $f : \mathbb{R} \to \mathbb{R}$ be a function and $\{x_i\}_{i=0}^k \subset \mathbb{R}$ be different points. Lagrange interpolating polynomial with support points $\{(x_i, f(x_i))\}_{i=0}^k$ is

$$p_n(x) = \sum_{j=0}^{n} f[x_j]\omega_{j-1}(x),$$

assuming $\omega_{-1} := 1$.

Proposition 6.51 (Newton's divided differences method). Let $f : \mathbb{R} \to \mathbb{R}$ be a function. For $x \in \mathbb{R}$, we have $f[x] = f(x)$. And if $\{x_i\}_{i=0}^n \subset \mathbb{R}$ are different points, then

$$f[x_0, \dots, x_n] = \frac{f[x_1, \dots, x_n] - f[x_0, \dots, x_{n-1}]}{x_n - x_0}.$$

Theorem 6.52. Let $f : [a, b] \to \mathbb{R}$ be a function of class $\mathcal{C}^{n+1}$, $\{x_i\}_{i=0}^n \subset \mathbb{R}$ be different points and $p_n \in \mathbb{R}[x]$ be the interpolating polynomial with support points $\{(x_i, f(x_i))\}_{i=0}^n$. Then $\forall x \in [a, b]$:

$$f(x) - p_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}\omega_n(x),$$

where $\xi_x \in \langle x_0, \dots, x_n, x\rangle$[77].

Lemma 6.53. Let $f : [a, b] \to \mathbb{R}$ be a function of class $\mathcal{C}^{n+1}$ and $\{x_i\}_{i=0}^n \subset \mathbb{R}$ be pairwise distinct points. Then $\exists \xi \in \langle x_0, \dots, x_n\rangle$ such that:

$$f[x_0, \dots, x_n] = \frac{f^{(n)}(\xi)}{n!}.$$

Proposition 6.54. Let $f : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^{n+1}$, $\{x_i\}_{i=0}^n \subset \mathbb{R}$ be pairwise distinct points and $\sigma \in S_n$. Then

$$f[x_0, \dots, x_n] = f[x_{\sigma(0)}, \dots, x_{\sigma(n)}]$$

Definition 6.55. Let $\{(x_i, y_i)\}_{i=0}^n$ be support points. The x-axis points $\{(x_i)\}_{i=0}^n$ are *equally-spaced* if

$$x_i = x_0 + ih, \quad \text{for } i = 0, \dots, n \quad \text{with} \quad h := \frac{x_n - x_0}{n}.$$

Definition 6.56. Let $f : [a, b] \to \mathbb{R}$ be a function and $\{x_i\}_{i=0}^n \subset \mathbb{R}$ be equally-spaced points. We define:

$$\Delta f(x) := f(x + h) - f(x),$$
$$\Delta^{n+1} f(x) := \Delta(\Delta^n f(x)).$$

Lemma 6.57. Let $f : [a, b] \to \mathbb{R}$ be a function and $\{x_i\}_{i=0}^n \subset \mathbb{R}$ be equally-spaced points. Then,

$$\Delta^n f(x_0) = n! h^n f[x_0, \dots, x_n].$$

Corollary 6.58. Let $f \in \mathbb{R}[x]$ with $\deg f = n$. Suppose we interpolate f with equally-spaced nodes. Then, $\Delta^n f(x) \equiv$ constant.

Hermite interpolation

Definition 6.59. Given a sets of points $\{(x_i)\}_{i=0}^m \subset \mathbb{R}$, $\{(n_i)\}_{i=0}^m \subset \mathbb{N}$ and $\{(y_i^{(k)} : k = 0, \dots, n_i - 1\}_{i=0}^m \subset \mathbb{R}$ *Hermite interpolation problem* consists in finding a polynomial $h_n \in \mathbb{R}[x]$ such that $\deg h_n \leq n$, $\sum_{i=0}^m n_i = n + 1$ and

$$h_n^{(k)}(x_i) = y_i^{(k)} \text{ for } i = 0, \dots, m \text{ and } k = 0, \dots, n_i - 1.$$

Proposition 6.60. Hermite interpolation problem has a unique solution.

[76] Types of interpolation are for example polynomial interpolation, trigonometric interpolation, Padé interpolation, Hermite interpolation and spline interpolation.

[77] The interval $\langle a_1, \dots, a_k\rangle$ is defined as $\langle a_1, \dots, a_k\rangle := (\min(a_1, \dots, a_k), \max(a_1, \dots, a_k))$.

Theorem 6.61. Let $f : [a, b] \to \mathbb{R}$ be a function of class $\mathcal{C}^{n+1}$, $\{x_i\}_{i=0}^m \subset \mathbb{R}$ be different points, $\{(n_i)\}_{i=0}^m \subset \mathbb{N}$ be such that $\sum_{i=0}^m n_i = n + 1$. Let h_n be the Hermite interpolating polynomial of f with nodes $\{x_i\}_{i=0}^m \subset \mathbb{R}$, that is,

$$h_n^{(k)}(x_i) = f^{(k)}(x_i) \text{ for } i = 0, \dots, m \text{ and } k = 0, \dots, n_i - 1.$$

Then $\forall x \in [a, b]$ $\exists \xi_x \in \langle x_0, \dots, x_n, x\rangle$ such that:

$$f(x) - h_n(x) = \frac{f^{(n+1)}(\xi_x)}{(n+1)!}(x - x_0)^{n_0} \cdots (x - x_m)^{n_m}.$$

Spline interpolation

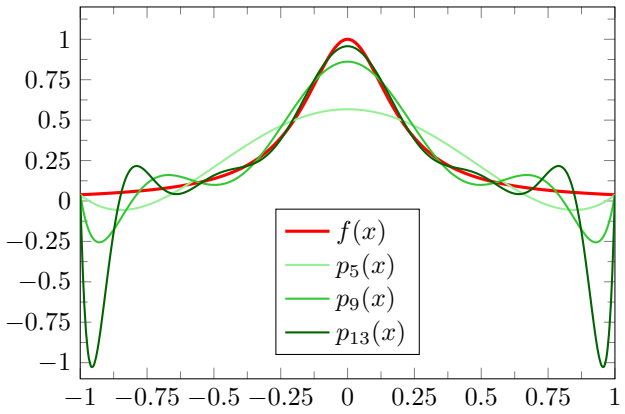

Figure 2.10: Runge's phenomenon. In this case $f(x) = \frac{1}{1+25x^2}$. $p_5(x)$ is the 5th-order Lagrange interpolating polynomial with equally-spaced interpolating points; $p_9(x)$, the 9th-order Lagrange interpolating polynomial with equally-spaced interpolating points, and $p_{13}(x)$, the 13th-order Lagrange interpolating polynomial with equally-spaced interpolating points.

Definition 6.62 (Spline). Let $\{(x_i, y_i)\}_{i=0}^n$ be support points of an interval $[a, b]$. A *spline of degree* p is a function $s : [a, b] \to \mathbb{R}$ of class $\mathcal{C}^{p-1}$ satisfying:

$$s_{|[x_i, x_{i+1}]} \in \mathbb{R}[x], \quad \deg s_{|[x_i, x_{i+1}]} = p, \quad s(x_i) = y_i,$$

for $i = 0, \dots, n - 1$. The most common case are splines of degree $p = 3$ or *cubic spline*. In this case we can impose two more conditions on their definition in one of the following ways:

1. *Natural cubic spline*:
$$s''(x_0) = s''(x_n) = 0.$$
2. *Cubic Hermite spline*: Given $y_0', y_n' \in \mathbb{R}$,
$$s'(x_0) = y_0', \quad s'(x_n) = y_n'.$$
3. *Cubic periodic spline*:
$$s'(x_0) = s'(x_n), \quad s''(x_0) = s''(x_n)$$

Definition 6.63. Let $f : [a, b] \to \mathbb{R}$ a function of class $\mathcal{C}^2$. We define the *seminorm*[78] *of* f as

$$\|f\|^2 = \int_a^b (f''(x))^2 dx.$$

Proposition 6.64. Let $f : [a, b] \to \mathbb{R}$ a function of class $\mathcal{C}^2$ interpolating the support points $\{(x_i, y_i)\}_{i=0}^n \subset \mathbb{R}^2$, $a \le x_0 < \cdots < x_n \le b$. If s is the natural cubic spline associated with $\{(x_i, y_i)\}_{i=0}^n$, then:

$$\|f-s\|^2 = \|f\|^2 - \|s\|^2 - 2(f'-s)s''\Big|_{x_0}^{x_n} + 2\sum_{i=1}^{n}(f-s)s'''\Big|_{x_{i-1}^+}^{x_i^-}.$$

Theorem 6.65. Let $f : [a, b] \to \mathbb{R}$ a function of class $\mathcal{C}^2$ interpolating the support points $\{(x_i, y_i)\}_{i=0}^n \subset \mathbb{R}^2$, $a \le x_0 < \cdots < x_n \le b$. If s is the natural cubic spline associated with $\{(x_i, y_i)\}_{i=0}^n$, then

$$\|s\| \le \|f\|.$$

2.6.4 | Numerical differentiation and integration

Differentiation

Theorem 6.66 (Intermediate value theorem). Let $f : [a, b] \to \mathbb{R}$ be a continuous function, $\xi_0, \dots, \xi_n \in [a, b]$ and $\alpha_0, \dots, \alpha_n \ge 0$. Then, $\exists \eta \in [a, b]$ such that:

$$\sum_{i=0}^{n} \alpha_i f(\xi_i) = \left(\sum_{i=0}^{n} \alpha_i\right) f(\eta).$$

Theorem 6.67 (Forward and backward difference formula of order 1). Let $f : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^2$. Then, forward difference formula of order 1 is:

$$f'(a) = \frac{f(a+h) - f(a)}{h} - \frac{f''(\xi)}{2}h,$$

where $\xi \in \langle a, a + h\rangle$. Analogously, backward difference formula of order 1 is:

$$f'(a) = \frac{f(a) - f(a-h)}{h} + \frac{f''(\eta)}{2}h,$$

where $\eta \in \langle a - h, a\rangle$.

Theorem 6.68 (Symmetric difference formula of order 1). Let $f : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^3$. Then, symmetric difference formula of order 1:

$$f'(a) = \frac{f(a+h) - f(a-h)}{2h} - \frac{f^{(3)}(\xi)}{6}h^2,$$

where $\xi \in \langle a - h, a + h\rangle$.

[78] The term *seminorm* has been used instead of *norm* to emphasize that not all properties of a norm are satisfied with this definition.

Theorem 6.69 (Symmetric difference formula of order 2). Let $f : \mathbb{R} \to \mathbb{R}$ be a function of class $\mathcal{C}^4$. Then, symmetric difference formula of order 2:

$$f''(a) = \frac{f(a+h) - 2f(a) + f(a-h)}{h^2} - \frac{f^{(4)}(\xi)}{12}h^2,$$

where $\xi \in \langle a-h, a, a+h \rangle$.

Richardson extrapolation

Theorem 6.70 (Richardson extrapolation). Suppose we have a function f that approximate a value α with an error that depends on a small quantity h. That is:

$$f(h) = \alpha + a_1 h^{k_1} + a_2 h^{k_2} + \cdots,$$

with $k_1 < k_2 < \cdots$ and a_i are unknown constants. Given $q > 0$, we define

$$D_1(h) = f(h), \quad D_{n+1}(h) = \frac{q^{k_n} D_n(h/q) - D_n(h)}{q^{k_n} - 1}.$$

And we can observe that $\alpha = D_{n+1}(h) + O(h^{k_{n+1}})$.

Integration

Definition 6.71. Let $f : [a,b] \to \mathbb{R}$ be a continuous function, $\{x_i\}_{i=0}^n \subseteq [a,b]$ be a set of nodes and P_n be the Lagrange interpolating polynomial with support points $\{(x_i, f(x_i))\}_{i=0}^n$. We define the *integration formula base on interpolation* as

$$I(f) = \int_a^b P_n(x)dx \tag{2.2}$$

Lemma 6.72. Let $f : [a,b] \to \mathbb{R}$ be a continuous function $\{x_i\}_{i=0}^n \subseteq [a,b]$ be a set of nodes. Then,

$$I(f) = \sum_{i=1}^n A_i f(x_i), \quad \text{where } A_i = \int_a^b \prod_{\substack{j=0 \\ j \neq i}}^n \frac{x - x_j}{x_i - x_j} dx.$$

Lemma 6.73. Let $p \in \mathbb{R}[x]$ be a polynomial defined on an interval $[a,b]$ such that $\deg p \leq n$ and let $\{x_i\}_{i=0}^n \subseteq [a,b]$ be a set of nodes. Then, $I(p) = \int_a^b p(x)dx$.

Lemma 6.74. Let $p \in \mathbb{R}[x]$ be a polynomial defined on an interval $[a,b]$ such that $\deg p \leq n$ and let $\{x_i\}_{i=0}^n \subseteq [a,b]$ be a set of nodes. Then,

$$I(p) = \int_a^b p(x)dx \iff I(x^k) = \int_a^b x^k dx, \text{ for } 0 \leq k \leq n.$$

Newton-Cotes formulas

Chapter 3

Third year

Chapter 4

Fourth year