

Notes

Logic, probability and machine learning

Reijo Jaakkola

Tampere University, Finland

October 2026

Preface

These notes of mine collect results and worked exercises in probability, stochastic processes, information theory, machine learning, model theory, algebra, and logic. Each note begins with a statement of the result followed by a proof.

Notation and conventions

Notation	Meaning
$[n]$	$\{1, \dots, n\}$.
$\Pr, \mathbb{E}, \text{Var}, \text{Cov}$	Probability, expectation, variance, and covariance.
1_A	The indicator of A .
$G(n, p)$	The random graph with independent edge probability p .
a.s.; w.h.p.	Almost surely; with probability tending to one.
$\bar{a}$	A finite tuple, with length determined by context.
$\mathfrak{A} \models \varphi$	Satisfaction of φ in $\mathfrak{A}$.
FO; SO	First-order logic; second-order logic with full semantics.
ESO; USO	Existential and universal second-order logic.
LFP	Least fixed-point logic.

Contents

Preface	i
1 Probability and analysis	1
1.1 Probability and measure	1
1.2 Random discrete structures	7
1.3 Limit theorems and random walks	15
1.4 Dynamics and Fourier analysis	17
1.5 Metric geometry	21
2 Stochastic processes	24
2.1 Brownian motion	24
2.2 Conditional expectation and stochastic integration	28
2.3 Linear stochastic differential equations	34
2.4 Exponential waiting times and Poisson processes	37
3 Information, randomization and cryptography	40
3.1 Entropy and coding	40
3.2 Randomized computation	42
3.3 Perfect secrecy	44
4 Learning and optimization	47
4.1 Estimation and moments	47
4.2 Probability distributions on Boolean types	52
4.3 Representation and convex optimization	57
4.4 Attention and classification objectives	62
5 Model theory and algebra	66
5.1 Cardinality, order and infinite games	66
5.2 Compactness and elementary equivalence	68
5.3 Fields and definability	74
5.4 Groups and Boolean algebras	81
6 Logic and computation	86
6.1 Computability and arithmetic	86
6.2 Fragments, games and model checking	90
6.3 Fixed points and expressive power	96
Note index	102

1. Probability and analysis

1.1 Probability and measure

Note 35. Functions of independent discrete variables

Statement. Let X, Y be independent discrete random variables, and let f, g be functions on their respective ranges. Then $f(X)$ and $g(Y)$ are independent.

Proof.

$$\begin{aligned}\Pr[f(X) = x \wedge g(Y) = y] &= \sum_{x' \in f^{-1}(\{x\})} \sum_{y' \in g^{-1}(\{y\})} \Pr[X = x' \wedge Y = y'] \\ &= \sum_{x' \in f^{-1}(\{x\})} \sum_{y' \in g^{-1}(\{y\})} \Pr[X = x'] \Pr[Y = y'] \\ &= \left(\sum_{x' \in f^{-1}(\{x\})} \Pr[X = x'] \right) \left(\sum_{y' \in g^{-1}(\{y\})} \Pr[Y = y'] \right) \\ &= \Pr[f(X) = x] \Pr[g(Y) = y]\end{aligned}$$

□

Note 115. Measurable functions preserve independence

Statement. If X, Y are independent real-valued random variables and $f, g : \mathbb{R} \rightarrow \mathbb{R}$ are measurable, then $f(X)$ and $g(Y)$ are independent.

Proof. Let $A, B \subseteq \mathbb{R}$ be Borel sets. Now

$$\begin{aligned}\Pr[f(X) \in A, g(Y) \in B] &= \Pr[X \in f^{-1}(A), Y \in g^{-1}(B)] \\ &= \Pr[X \in f^{-1}(A)] \Pr[Y \in g^{-1}(B)] \\ &= \Pr[f(X) \in A] \Pr[g(Y) \in B]\end{aligned}$$

proving the claim.

□

Note 116. Factoring expectations of independent variables

Statement. Let X, Y be independent random variables and f, g measurable functions with $\mathbb{E}|f(X)|, \mathbb{E}|g(Y)| < \infty$. Then $f(X)g(Y)$ is integrable and

$$\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)].$$

Proof. Since X, Y are independent, their law $F_{X,Y}$ over $\mathbb{R}^2$ is the product measure of F_X and F_Y . Furthermore, by Tonelli's theorem, we have that

$$\int_{\mathbb{R}^2} |f(x)g(y)| dF_{X,Y}(x, y) = \left(\int_{\mathbb{R}} |f(x)| dF_X(x) \right) \left(\int_{\mathbb{R}} |g(y)| dF_Y(y) \right) < \infty.$$

Hence it follows from Fubini's theorem that

$$\int_{\mathbb{R}^2} f(x)g(y) dF_{X,Y}(x, y) = \left(\int_{\mathbb{R}} f(x) dF_X(x) \right) \left(\int_{\mathbb{R}} g(y) dF_Y(y) \right),$$

In other words $\mathbb{E}[f(X)g(Y)] = \mathbb{E}[f(X)]\mathbb{E}[g(Y)]$.

□

Note 73. Conditioning on a disjoint union

Statement. Let A, B, C be events with $B \cap C = \emptyset$ and $\Pr(B), \Pr(C) > 0$. Then

$$\Pr(A \mid B \cup C) \leq \max\{\Pr(A \mid B), \Pr(A \mid C)\}.$$

The inequality can fail when B and C are not disjoint.

Proof. Aiming for a contradiction, suppose that

$$\Pr[A \mid B \cup C] > \Pr[A \mid B], \Pr[A \mid C]$$

Since $B \cap C = \emptyset$ and $\Pr(B), \Pr(C) > 0$, we can rewrite the left-hand-side as

$$\frac{\Pr[A \cap B] + \Pr[A \cap C]}{\Pr[B] + \Pr[C]}$$

Consider then the inequality

$$\frac{\Pr[A \cap B] + \Pr[A \cap C]}{\Pr[B] + \Pr[C]} > \frac{\Pr[A \cap B]}{\Pr[B]}$$

Multiply both sides with $\Pr[B]$ and $\Pr[B] + \Pr[C]$ to obtain

$$\begin{aligned} \Pr[B] \Pr[A \cap B] + \Pr[B] \Pr[A \cap C] &> \Pr[B] \Pr[A \cap B] + \Pr[C] \Pr[A \cap B] \\ \Leftrightarrow \Pr[B] \Pr[A \cap C] &> \Pr[C] \Pr[A \cap B] \\ \Leftrightarrow \frac{\Pr[A \cap C]}{\Pr[C]} &> \frac{\Pr[A \cap B]}{\Pr[B]} \end{aligned}$$

On the other hand, if we start with the inequality

$$\frac{\Pr[A \cap B] + \Pr[A \cap C]}{\Pr[B] + \Pr[C]} > \frac{\Pr[A \cap C]}{\Pr[C]}$$

and multiply both sides with $\Pr[C]$ and $\Pr[B] + \Pr[C]$, we obtain that

$$\begin{aligned} \Pr[C] \Pr[A \cap B] + \Pr[C] \Pr[A \cap C] &> \Pr[B] \Pr[A \cap C] + \Pr[C] \Pr[A \cap C] \\ \Leftrightarrow \Pr[C] \Pr[A \cap B] &> \Pr[B] \Pr[A \cap C] \\ \Leftrightarrow \frac{\Pr[A \cap B]}{\Pr[B]} &> \frac{\Pr[A \cap C]}{\Pr[C]}, \end{aligned}$$

which is a clear contradiction.

The above result does not hold in general if we drop the assumption $B \cap C = \emptyset$ and $\Pr(B), \Pr(C) > 0$. Consider for example the case where $A = \{1, 3\}, B = \{1, 2\}, C = \{2, 3\}$ and the underlying probability distribution is the uniform probability distribution over $\{1, 2, 3\}$. Then

$$\Pr[A \mid B \cup C] = \frac{|\{1, 3\}|}{|\{1, 2, 3\}|} = \frac{2}{3}$$

but

$$\Pr[A \mid B] = \frac{|\{1\}|}{|\{1, 2\}|} = \frac{1}{2}$$

and

$$\Pr[A \mid C] = \frac{|\{3\}|}{|\{2, 3\}|} = \frac{1}{2}$$

□

Note 74. Expectation under a push-forward law

Statement. Let $X : \Omega \rightarrow \mathbb{R}^d$ be a random variable with law P_X . If $g : \mathbb{R}^d \rightarrow \mathbb{R}^d$ is measurable and $g(X)$ is integrable, then

$$\mathbb{E}[g(X)] = \int_{\mathbb{R}^d} g(x) dP_X(x).$$

Proof. This can be established very briefly by using change-of-variables:

$$\begin{aligned} \mathbb{E}[g(X)] &= \int_{\Omega} (g \circ X)(\omega) dP(\omega) \\ &= \int_{\mathbb{R}^d} g(x) dP_X(x) \end{aligned}$$

□

Note 103. Bayesian inference from two bags of fruit

Statement. Bag 1 contains four mangoes and two apples; bag 2 contains four mangoes and four apples. Choose bag 2 with probability $2/5$ and bag 1 with probability $3/5$, then choose a fruit uniformly from the selected bag. Conditional on obtaining a mango, the probability that bag 2 was selected is $1/3$.

Proof. Let A be the event that bag 2 was chosen and B the event that the fruit is a mango. By total probability,

$$\Pr(B) = \frac{1}{2} \cdot \frac{2}{5} + \frac{4}{6} \cdot \frac{3}{5} = \frac{3}{5}.$$

Bayes' theorem now gives

$$\Pr(A \mid B) = \frac{\Pr(B \mid A) \Pr(A)}{\Pr(B)} = \frac{(1/2)(2/5)}{3/5} = \frac{1}{3}.$$

Thus the mango is twice as likely to have come from bag 1 as from bag 2.

□

Note 117. Scaling Lebesgue integrals

Statement. Let μ be Lebesgue measure on $\mathbb{R}$, let $f : \mathbb{R} \rightarrow [0, \infty]$ be measurable, and let $\lambda > 0$. Then

$$\int_{\mathbb{R}} f(\lambda x) \, d\mu(x) = \frac{1}{\lambda} \int_{\mathbb{R}} f(x) \, d\mu(x).$$

Proof. First, consider the function $g(x) := \lambda x$. Now

$$\int_{\mathbb{R}} f(\lambda x) \, d\mu(x) = \int_{\mathbb{R}} (f \circ g)(x) \, d\mu(x) = \int_{\mathbb{R}} f \, d(g_*\mu)(x)$$

by change-of-variables. By definition,

$$(g_*\mu)(A) = \mu(g^{-1}(A)) = \mu\left(\frac{1}{\lambda}A\right).$$

Since μ is the Lebesgue measure, we have that

$$\mu\left(\frac{1}{\lambda}A\right) = \frac{1}{\lambda} \mu(A).$$

Now it is easy to show using the monotone convergence theorem that

$$\int_{\mathbb{R}} f \, d(g_*\mu)(x) = \frac{1}{\lambda} \int_{\mathbb{R}} f(x) \, d\mu(x)$$

and we are done. □

Note 118. The tail integral formula for expectation

Statement. For every nonnegative random variable X , allowing an infinite expectation,

$$\mathbb{E}[X] = \int_0^\infty \Pr(X > \lambda) \, d\lambda.$$

Proof. Consider first the case where X is a simple function, i.e., a sum of the form

$$\sum_{i=1}^m c_i 1_{A_i},$$

where $A_i \cap A_j = \emptyset$ for each $i \neq j$, and $0 = c_0 \leq c_1 < \dots < c_m$. Now

$$\begin{aligned}
 \int_0^\infty \Pr(X > \lambda) d\lambda &= \int_0^{c_m} \Pr(X > \lambda) d\lambda \\
 &= \sum_{i=1}^m \int_{c_{i-1}}^{c_i} \Pr(X > \lambda) d\lambda \\
 &= \sum_{i=1}^m (c_i - c_{i-1}) \mu\left(\bigcup_{j=i}^m A_j\right) \\
 &= \sum_{i=1}^m (c_i - c_{i-1}) \sum_{j=i}^m \mu(A_j) \\
 &= \sum_{j=1}^m \mu(A_j) \sum_{i=1}^j (c_i - c_{i-1}) \\
 &= \sum_{j=1}^m c_j \mu(A_j) = \mathbb{E}[X]
 \end{aligned}$$

For general non-negative random variables the claim follows from the monotone convergence theorem. □

Note 119. Images of inverse images under a surjection

Statement. If $f : X \rightarrow Y$ is surjective, then $f(f^{-1}(A)) = A$ for every subset $A \subseteq Y$.

Proof. First, if $y \in A$, then there exists $x \in X$ such that $f(x) = y$. Now $x \in f^{-1}(A)$ and hence $y = f(x) \in f(f^{-1}(A))$. Secondly, if $y \in f(f^{-1}(A))$, then there exists $x \in f^{-1}(A)$ such that $f(x) = y$. Now $y = f(x) \in A$. □

Note 126. The law of total expectation

Statement. If X is integrable and Y is a random variable on the same probability space, then

$$\mathbb{E}[\mathbb{E}(X \mid Y)] = \mathbb{E}[X].$$

Proof. By the defining property of conditional expectation, its integral over the whole probability space equals the integral of X . The identity follows. In the case of a joint density $p(x, y) = p(x \mid y)p_Y(y)$, Fubini's theorem gives

$$\begin{aligned}
 \mathbb{E}[\mathbb{E}(X \mid Y)] &= \int_{\mathbb{R}} \left(\int_{\mathbb{R}} x p(x \mid y) dx \right) p_Y(y) dy \\
 &= \int_{\mathbb{R}} x \left(\int_{\mathbb{R}} p(x, y) dy \right) dx \\
 &= \int_{\mathbb{R}} x p_X(x) dx = \mathbb{E}[X].
 \end{aligned}$$

□

Note 134. Total probability conditional on an event

Statement. Let $\Pr(B) > 0$, and let $(C_i)_{i \in I}$ be a countable measurable partition of the sample space. Then

$$\Pr(A | B) = \sum_{\substack{i \in I \\ \Pr(B \cap C_i) > 0}} \Pr(A | B \cap C_i) \Pr(C_i | B).$$

Proof.

$$\begin{aligned} \Pr(A|B) &= \frac{\Pr(A \cap B)}{\Pr(B)} \\ &= \sum_i \frac{\Pr(A \cap B \cap C_i)}{\Pr(B)} \\ &= \sum_i \Pr(A|B \cap C_i) \frac{\Pr(B \cap C_i)}{\Pr(B)} \\ &= \sum_i \Pr(A|B \cap C_i) \Pr(C_i|B) \end{aligned}$$

□

Note 177. Symmetric difference bounds probability differences

Statement. For any two events A, B ,

$$|\Pr(A) - \Pr(B)| \leq \Pr(A \triangle B),$$

where $A \triangle B$ denotes their symmetric difference.

Proof. Writing

$$\Pr(A) = \Pr(A \cap B) + \Pr(A \cap B^c)$$

and

$$\Pr(B) = \Pr(A \cap B) + \Pr(A^c \cap B)$$

we get

$$\begin{aligned} |\Pr(A) - \Pr(B)| &= |\Pr(A \cap B) + \Pr(A \cap B^c) - \Pr(A \cap B) - \Pr(A^c \cap B)| \\ &= |\Pr(A \cap B^c) - \Pr(A^c \cap B)| \\ &\leq \Pr(A \cap B^c) + \Pr(A^c \cap B) \\ &= \Pr(A \triangle B). \end{aligned}$$

□

Note 178. Bayesian confirmation is reciprocal

Statement. *Let A, B be events with $\Pr(A), \Pr(B) > 0$. If $\Pr(A \mid B) > \Pr(A)$, then $\Pr(B \mid A) > \Pr(B)$.*

Proof.

$$\begin{aligned}\Pr(B \mid A) &= \frac{\Pr(A \mid B) \Pr(B)}{\Pr(A)} \\ &> \frac{\Pr(A) \Pr(B)}{\Pr(A)} \\ &= \Pr(B)\end{aligned}$$

Putting this in terms of abductive logic, if the theory B being true increases the likelihood of event A , then the occurrence of A makes theory B more likely. □

Note 72. Borel sets in a product of separable metric spaces

Statement. *If M is a separable metric space, then the Borel σ -algebra of its product satisfies*

$$\mathfrak{B}(M \times M) = \mathfrak{B}(M) \otimes \mathfrak{B}(M).$$

Proof. We start with the left-to-right inclusion. First, consider an open set $U \subseteq M \times M$. Since M is a separable metric space, it is also second-countable. Thus $M \times M$ is also second-countable with the product topology. Hence it is Lindelöf and thus U can be represented in the form

$$U = \bigcup_{n=0}^{\infty} (V_n \times W_n),$$

where $V_n, W_n \subseteq M$ are open sets. Now $(V_n \times W_n) \in \mathfrak{B}(M) \otimes \mathfrak{B}(M)$, for every n , and hence $U \in \mathfrak{B}(M) \otimes \mathfrak{B}(M)$. Thus $\mathfrak{B}(M \times M) \subseteq \mathfrak{B}(M) \otimes \mathfrak{B}(M)$, since the former is the smallest σ -algebra that contains all the open sets of $M \times M$.

Next we consider the right-to-left inclusion. Since $\mathfrak{B}(M) \otimes \mathfrak{B}(M)$ is the smallest σ -algebra that contains sets of the form $(B_1 \times B_2)$, where $B_1, B_2 \in \mathfrak{B}(M)$, it suffices to show that $\mathfrak{B}(M \times M)$ contains all such sets. In fact, it suffices to show that if $B \in \mathfrak{B}(M)$, then $B \times M \in \mathfrak{B}(M \times M)$. First, we note that if $U \subseteq M$ is open, then $U \times M \in \mathfrak{B}(M \times M)$. Secondly, we note that if $B \times M \in \mathfrak{B}(M \times M)$, then $B^c \times M = (B \times M)^c \in \mathfrak{B}(M \times M)$. Finally, if $(B_n \times M) \in \mathfrak{B}(M \times M)$, for every $n \in \mathbb{N}$, then

$$\left(\bigcup_{n=0}^{\infty} B_n \times M \right) = \bigcup_{n=0}^{\infty} (B_n \times M) \in \mathfrak{B}(M \times M)$$

Thus $\mathfrak{B}(M \times M)$ contains all sets of the form $(B_1 \times B_2)$, where $B_1, B_2 \in \mathfrak{B}(M)$. □

1.2 Random discrete structures

Note 2. Variable-length random clauses

Statement. Form a clause on n variables by including each of the $2n$ literals independently with a fixed probability $p \in (0, 1]$. The probability that the clause contains a complementary pair is $1 - (1 - p^2)^n$. Consequently, a CNF formula consisting of polynomially many independently generated clauses is a tautology with probability tending to one as $n \rightarrow \infty$.

Proof. For a moment fix a set of variables $\{x_1, \dots, x_n\}$. Now

$$\begin{aligned} & \Pr[\exists x : x \in C \text{ and } \neg x \in C] \\ &= 1 - \Pr[\forall x : \neg(x \in C \text{ and } \neg x \in C)] \\ &= 1 - \prod_{i=1}^n \Pr[\neg(x_i \in C \text{ and } \neg x_i \in C)] \end{aligned}$$

Note that $\Pr[\neg(x_i \in C \text{ and } \neg x_i \in C)]$ is the same as

$$\begin{aligned} & \Pr[x_i \in C \text{ and } \neg x_i \notin C] + \Pr[x_i \notin C \text{ and } \neg x_i \in C] + \Pr[x_i \notin C \text{ and } \neg x_i \notin C] \\ &= 2p(1 - p) + (1 - p)^2 \end{aligned}$$

Thus the probability is $1 - (1 - p^2)^n$, which tends to 1 for fixed $p > 0$.

From this it follows immediately that if we want to generate, say, $\text{poly}(n)$ -many clauses with variables from the set $\{x_1, \dots, x_n\}$, then w.h.p. there is no clause C which would not contain both x_i and $\neg x_i$, for some $1 \leq i \leq n$. In particular, the resulting formula is w.h.p. valid. □

Note 3. Specified letters in a random word

Statement. Let Σ be an alphabet of size N , let $\ell_1, \dots, \ell_t$ be distinct letters, and let $m \geq t$. A uniformly random word $x \in \Sigma^m$ satisfies

$$\Pr(\ell_1, \dots, \ell_t \text{ all occur in } x) \leq \frac{m(m-1) \dots (m-t+1)}{N^t}.$$

Proof. Let $\Sigma_{m;\ell_1, \dots, \ell_t}$ denote the set of all words over Σ of length m in which each of the letters ℓ_i occurs. Given pairwise distinct positions $j_1, \dots, j_t \in [m]$, we let $\Sigma_{m;\ell_1, \dots, \ell_t}^{j_1, \dots, j_t}$ denote the set of all words over Σ of length m such that for every $1 \leq i \leq t$ the letter ℓ_i occurs (at least) in position j_i . Clearly

$$\begin{aligned} |\Sigma_{m;\ell_1, \dots, \ell_t}| &\leq \sum_{\substack{j_1, \dots, j_t \in [m] \\ \text{pairwise distinct}}} |\Sigma_{m;\ell_1, \dots, \ell_t}^{j_1, \dots, j_t}| \\ &= \sum_{\substack{j_1, \dots, j_t \in [m] \\ \text{pairwise distinct}}} N^{m-t} \\ &= m(m-1) \dots (m-t+1) N^{m-t} \end{aligned}$$

Hence

$$\begin{aligned} \Pr_{x \in \Sigma^m} [\forall 1 \leq i \leq t : \ell_i \in x] &= \frac{|\Sigma_{m;\ell_1, \dots, \ell_t}|}{N^m} \\ &\leq \frac{m(m-1) \dots (m-t+1) N^{m-t}}{N^m} \\ &= \frac{m(m-1) \dots (m-t+1)}{N^t} \end{aligned}$$

□

Note 4. Connectivity under the pure literal rule

Statement. Let ϕ be a CNF formula that can be satisfied by successive applications of the pure literal rule until no clauses remain. Its satisfying assignments form a connected subgraph of the Boolean cube, where two assignments are adjacent when they differ in one coordinate.

Proof. By assumption, there exists a sequence of literals $\ell_{i_1}, \dots, \ell_{i_m}$ such that (1) setting all of them true guarantees that ϕ is also true, (2) ℓ_{i_1} is pure and (3) for each $1 \leq j < m$ we have that once the literals $\ell_{i_1}, \dots, \ell_{i_j}$ are set to be true and ϕ is simplified accordingly the literal $\ell_{i_{j+1}}$ is pure (its complement does not occur in the formula).

For each $1 \leq j \leq m$ we let b_j denote the value which makes ℓ_{i_j} true; we have for every $x \in \{0, 1\}^n$ that if $x_{i_j} = b_j$, for every $1 \leq j \leq m$, then $x \in S(\phi)$. We now claim that for every $x \in S(\phi)$ there exists a path to an assignment $y \in S(\phi)$ such that $y_{i_j} = b_j$, for every $1 \leq j \leq m$, which will then imply that $S(\phi)$ is connected.

Pick some $x \in S(\phi)$. Let y^1 denote the assignment obtained from x by replacing x_{i_1} with b_1 . Clearly $y^1 \in S(\phi)$, since ℓ_{i_1} is pure. Indeed, if x makes ϕ true, then the restriction of x to indices $[n] \setminus \{i_1\}$ also makes true the formula obtained from ϕ by removing all the clauses in which ℓ_{i_1} occurs.

Suppose now that we are considering an assignment $y^j \in S(\phi)$, where $j < m$, such that $y_{i_k}^j = b_k$, for every $k \leq j$. We let y^{j+1} denote the assignment obtained from y^j by replacing $y_{i_{j+1}}^j$ with b_{j+1} . Since $\ell_{i_{j+1}}$ becomes pure after $\ell_{i_1}, \dots, \ell_{i_j}$ have been set to true, a reasoning similar to the case of ℓ_{i_1} shows that $y^{j+1} \in S(\phi)$. After deleting consecutive repetitions, $x, y^1, \dots, y^m$ is the desired path.

The assignments with $y_{i_j} = b_j$ for all j form a subcube contained in $S(\phi)$, and this subcube is connected. Hence $S(\phi)$ is connected. □

Note 5. Cycles in sparse random graphs

Statement. If $p = p(n)$ satisfies $np \rightarrow 0$, then $G(n, p)$ is acyclic with probability tending to one.

Proof. Let X count the cycles in $G(n, p)$. By Markov's inequality,

$$\begin{aligned} \Pr[X \geq 1] &\leq \mathbb{E}[X] \\ &= \sum_{k=3}^n \frac{n!}{2k(n-k)!} p^k \\ &\leq \frac{1}{6} \sum_{k=3}^n (np)^k \\ &\leq \frac{(np)^3}{6(1-np)} \rightarrow 0, \end{aligned}$$

where the last bound holds for all sufficiently large n .

Remark 5.1. The weaker assumption $np \leq \varepsilon < 1$ only yields a constant upper bound; it does not imply that the probability of a cycle tends to zero. □

Note 6. Common neighbours in dense random graphs

Statement. For fixed $p \in (0, 1]$, with probability tending to one every pair of distinct vertices of $G(n, p)$ has a common neighbour. In particular, the graph is connected and has diameter at most two with high probability.

Proof. For every distinct $i, j \in [n]$ we introduce the indicator random variable $X_{i,j}$ for the event “ i and j do not have a common neighbor”. Clearly

$$\mathbb{E}(X_{i,j}) = (1 - p^2)^{(n-2)}$$

Let $X = \sum_{i \neq j} X_{i,j}$. Markov’s inequality gives

$$\Pr[X \geq 1] \leq \mathbb{E}(X) = \sum_{i \neq j} \mathbb{E}(X_{i,j}) = n(n-1)(1 - p^2)^{(n-2)}$$

which goes to zero as $n \rightarrow \infty$. □

Note 7. The threshold for isolated vertices

Statement. Let X count the isolated vertices of $G(n, p)$. If $p \gg (\log n)/n$, then $\Pr(X = 0) \rightarrow 1$. If $p \ll (\log n)/n$, then $\Pr(X \geq 1) \rightarrow 1$. Thus $(\log n)/n$ is a threshold for the absence of isolated vertices.

Proof. Suppose first that $p \gg \ln n/n$. Let X denote the random variable which counts the number of isolated vertices in $G(n, p)$. Then Markov’s inequality entails that

$$\begin{aligned} \Pr[X \geq 1] &\leq \mathbb{E}(X) \\ &= n(1 - p)^{n-1} \\ &\leq n \left(1 - \frac{c \ln n}{n}\right)^{n-1} \\ &\sim n \cdot e^{-c \ln n} \\ &= n \cdot n^{-c} \rightarrow 0 \end{aligned}$$

as long as $c > 1$.

Suppose then that $p \ll \ln n/n$. Now

$$\Pr[X \geq 1] \geq \frac{(\mathbb{E}(X))^2}{\mathbb{E}(X^2)}$$

We already saw that

$$(\mathbb{E}(X))^2 = n^2(1 - p)^{2(n-1)}$$

Furthermore

$$\begin{aligned}
\mathbb{E}(X^2) &= \mathbb{E}\left(\sum_i X_i\right)^2 \\
&= \sum_{i,j} \mathbb{E}(X_i X_j) \\
&= \sum_{i \neq j} \mathbb{E}(X_i X_j) + \mathbb{E}(X) \\
&= n(n-1)(1-p)^{2n-3} + \mathbb{E}(X) \\
&= \frac{1-1/n}{1-p} (\mathbb{E}(X))^2 + \mathbb{E}(X)
\end{aligned}$$

Hence

$$\begin{aligned}
\frac{(\mathbb{E}(X))^2}{\mathbb{E}(X^2)} &= \frac{(\mathbb{E}(X))^2}{\frac{1-1/n}{1-p} (\mathbb{E}(X))^2 + \mathbb{E}(X)} \\
&= \frac{1}{\frac{1-1/n}{1-p} + \underbrace{(\mathbb{E}(X))^{-1}}_{\rightarrow 0}} \rightarrow 1
\end{aligned}$$

since $p \ll \ln n/n$.

□

Note 50. Coincidences of independent binomial variables

Statement. If X, Y are independent $\text{Binom}(n, 1/2)$ random variables, then

$$\Pr(X = Y) = 4^{-n} \binom{2n}{n} \sim \frac{1}{\sqrt{\pi n}}.$$

In particular, $\Pr(X \neq Y) \rightarrow 1$.

Proof. First, a direct computation gives us that

$$\begin{aligned}
\Pr[X = Y] &= \sum_{k=0}^n \Pr[X = k \wedge Y = k] \\
&= \sum_{k=0}^n \Pr[X = k] \cdot \Pr[Y = k] \\
&= \frac{1}{2^{2n}} \sum_{k=0}^n \binom{n}{k}^2
\end{aligned}$$

Now we have that

$$\sum_{k=0}^n \binom{n}{k}^2 = \binom{2n}{n} \sim \frac{2^{2n}}{\sqrt{\pi n}},$$

where the last approximation follows from Stirling's approximation. Thus

$$\Pr[X = Y] \sim \frac{1}{2^{2n}} \cdot \frac{2^{2n}}{\sqrt{\pi n}} = \frac{1}{\sqrt{\pi n}} \rightarrow 0,$$

which is what we wanted to show. □

Note 51. Separation of independent binomial variables

Statement. Let X, Y be independent $\text{Binom}(n, 1/2)$ random variables. For every fixed integer $m \geq 0$,

$$\Pr(|X - Y| > m) \rightarrow 1 \quad (n \rightarrow \infty).$$

Proof. Clearly it suffices to show that $\Pr(|X - Y| = m) \rightarrow 0$, for any fixed integer $m \geq 0$. The case $m = 0$ follows from Note 50. For $m > 0$ we first perform the following direct computation

$$\begin{aligned} \Pr[|X - Y| = m] &= \sum_{k=0}^{n-m} \Pr[X = k + m \wedge Y = k] + \sum_{k=0}^{n-m} \Pr[X = k \wedge Y = k + m] \\ &= \sum_{k=0}^{n-m} \Pr[X = k + m] \cdot \Pr[Y = k] + \sum_{k=0}^{n-m} \Pr[X = k] \cdot \Pr[Y = k + m] \\ &= \frac{1}{2^{2n-1}} \sum_{k=0}^{n-m} \binom{n}{n-k} \binom{n}{k+m} \\ &= \frac{1}{2^{2n-1}} \binom{2n}{n+m} \end{aligned}$$

We now use Stirling's approximation to approximate the binomial coefficient.

$$\begin{aligned} \binom{2n}{n+m} &\sim \sqrt{\frac{n}{\pi(n+m)(n-m)}} \cdot \frac{2^{2n} n^{2n}}{(n+m)^{(n+m)}(n-m)^{n-m}} \\ &\sim \frac{1}{\sqrt{\pi n}} \cdot \frac{2^{2n} n^{2n}}{(n+m)^{(n+m)}(n-m)^{n-m}} \\ &\sim \frac{2^{2n}}{\sqrt{\pi n}} \end{aligned}$$

Thus

$$\Pr[|X - Y| = m] \sim \frac{2}{2^{2n}} \cdot \frac{2^{2n}}{\sqrt{\pi n}} = \frac{2}{\sqrt{\pi n}} \rightarrow 0,$$

which is what we wanted to show. □

Note 53. The parity of a Bernoulli sum

Statement. For independent Bernoulli random variables $X_1, \dots, X_k$ with $\Pr(X_i = 1) = 1 - \delta$,

$$\Pr\left(\sum_{i=1}^k X_i \equiv 0 \pmod{2}\right) = \frac{1}{2} + \frac{(2\delta - 1)^k}{2}.$$

Proof. We use induction with respect to k . When $k = 1$ the claim is clear. Suppose that it holds for k and we want to establish it for $k + 1$. Then

$$\begin{aligned} & \Pr\left[\sum_{i=1}^{k+1} X_i \equiv 0 \pmod{2}\right] \\ &= \Pr\left[\sum_{i=1}^{k+1} X_i \equiv 0 \pmod{2} \mid \sum_{i=1}^k X_i \equiv 0 \pmod{2}\right] \cdot \Pr\left[\sum_{i=1}^k X_i \equiv 0 \pmod{2}\right] \\ &+ \Pr\left[\sum_{i=1}^{k+1} X_i \equiv 0 \pmod{2} \mid \sum_{i=1}^k X_i \equiv 1 \pmod{2}\right] \cdot \Pr\left[\sum_{i=1}^k X_i \equiv 1 \pmod{2}\right] \\ &= \Pr[X_{k+1} = 0] \cdot \Pr\left[\sum_{i=1}^k X_i \equiv 0 \pmod{2}\right] + \Pr[X_{k+1} = 1] \cdot \Pr\left[\sum_{i=1}^k X_i \equiv 1 \pmod{2}\right] \\ &= \delta \cdot \left(\frac{1}{2} + \frac{(2\delta - 1)^k}{2}\right) + (1 - \delta) \cdot \left(1 - \left(\frac{1}{2} + \frac{(2\delta - 1)^k}{2}\right)\right) \\ &= \frac{\delta}{2} + \frac{\delta(2\delta - 1)^k}{2} + \frac{(1 - \delta)}{2} - \frac{(1 - \delta)(2\delta - 1)^k}{2} \\ &= \frac{1}{2} + \frac{\delta(2\delta - 1)^k - (1 - \delta)(2\delta - 1)^k}{2} \\ &= \frac{1}{2} + \frac{(\delta - 1 + \delta)(2\delta - 1)^k}{2} \\ &= \frac{1}{2} + \frac{(2\delta - 1)^{k+1}}{2} \end{aligned}$$

which is what we wanted to show. □

Note 57. Monotone paths in a random grid

Statement. Retain each edge of the $n \times n$ grid independently with probability $p(n) \leq 1/2$. The probability that the resulting graph has a path from $(1, 1)$ to (n, n) using only up and right steps tends to zero.

Proof. Let X count the paths from $(1, 1)$ to (n, n) that always go up or right. There are exactly $\binom{2n-2}{n-1}$ such paths, each using $2n - 2$ edges. Therefore, by Markov's inequality and Stirling's approximation,

$$\begin{aligned} \Pr[X \geq 1] &\leq \mathbb{E}[X] = \binom{2n-2}{n-1} p(n)^{2n-2} \\ &\leq \frac{1}{4^{n-1}} \binom{2n-2}{n-1} \sim \frac{1}{\sqrt{\pi(n-1)}} \rightarrow 0. \end{aligned}$$

Thus, with high probability, no such path exists. □

Note 58. A threshold for selecting a perfect square

Statement. Let $\mathcal{N}(N, p)$ be the random subset of $[N]$ obtained by including each integer independently with probability $p = p(N)$. The probability that this set contains a perfect square tends to zero if $p\sqrt{N} \rightarrow 0$, and tends to one if $p\sqrt{N} \rightarrow \infty$. Hence $N^{-1/2}$ is a threshold for this property.

Proof. Given a perfect square s in $[N]$, we let 1_s denote the indicator variable for the event that $\mathcal{N}(N, p)$ contains s . Let $X = \sum_s 1_s$. By Markov's inequality we have that

$$\Pr[X \geq 1] \leq \mathbb{E}(X) = \sum_s \Pr[1_s = 1] = p \sum_s 1 \leq p\sqrt{N}$$

Here $\sum_s 1 \leq \sqrt{N}$ follows from the fact that $[N]$ has $\lfloor \sqrt{N} \rfloor$ many perfect squares. Thus, as long as $\lim_{N \rightarrow \infty} \sqrt{N}p(N) = 0$, we have that $\Pr[X \geq 1] \rightarrow 0$. Thus we might guess that $1/\sqrt{N}$ is a threshold for this property.

Suppose then that $\lim_{N \rightarrow \infty} \sqrt{N}p(N) = \infty$, in which case $\mathbb{E}(X) \rightarrow \infty$. We show that with high probability $\mathcal{N}(N, p)$ contains a perfect square. It suffices to show that $\text{Var}(X) = o(\mathbb{E}^2(X))$. By calculating that

$$\begin{aligned} \mathbb{E}(X^2) &= \sum_s \mathbb{E}(1_s^2) + \sum_{s \neq s'} \mathbb{E}(1_s 1_{s'}) \\ &= \sum_s \mathbb{E}(1_s) + \sum_{s \neq s'} \mathbb{E}(1_s) \mathbb{E}(1_{s'}) \\ &\leq \mathbb{E}(X) + \mathbb{E}^2(X) \end{aligned}$$

we see that

$$\text{Var}(X) \leq \mathbb{E}(X)$$

By assumption we have that $\mathbb{E}(X) \rightarrow \infty$, which implies that $\mathbb{E}(X) = o(\mathbb{E}^2(X))$. Thus with high probability we have that $\mathcal{N}(N, p)$ contains a perfect square. □

Note 130. The expected image size of a random function

Statement. Let $f : [n] \rightarrow A$ be uniformly random among all functions, where $|A| = n$. Then

$$\mathbb{E}|\text{im}(f)| = n \left(1 - \left(1 - \frac{1}{n} \right)^n \right) \sim (1 - e^{-1})n.$$

Proof. For each $a \in A$ we introduce a random variable X_a such that

$$X_a(\omega) := \begin{cases} 1, & \text{if } a \in \text{im}(f) \\ 0, & \text{otherwise} \end{cases}$$

We are interested in the quantity

$$\mathbb{E} \left[\sum_a X_a \right] = \sum_a \mathbb{E}[X_a] = \sum_a \Pr[a \in \text{im}(f)].$$

Now

$$\Pr[a \in \text{im}(f)] = 1 - \Pr[a \notin \text{im}(f)] = 1 - \frac{(n-1)^n}{n^n} = 1 - \left(1 - \frac{1}{n}\right)^n,$$

where

$$\left(1 - \frac{1}{n}\right)^n \approx \frac{1}{e}.$$

So roughly a fraction of $(1 - 1/e) \approx 0.632$ of the elements of A have a pre-image. $\square$

Note 162. Chromatic number forces edges

Statement. Every finite graph G has at least $\binom{\chi(G)}{2}$ edges, where $\chi(G)$ is its chromatic number.

Proof. Let $k := \chi(G)$ and consider a coloring $\chi : G \rightarrow [k]$. It suffices to show that for each $i, j \in [k]$, where $i \neq j$, there exists an edge in G whose endpoints receive the colors i and j . Aiming for a contradiction, suppose that this is not the case. Thus there exist distinct $i, j \in [k]$ such that there is no edge whose endpoints receive the colors i and j .

Consider now the coloring $\chi_j : G \rightarrow [k] \setminus \{j\}$ defined by setting

$$\chi_j(v) := \begin{cases} i & \text{if } \chi(v) = j \\ \chi(v) & \text{otherwise} \end{cases}$$

We claim that χ_j is a valid coloring, which would be a contradiction because k was the chromatic number of G . Aiming for a contradiction, assume that there exists an edge $\{v, u\}$ such that $\chi_j(v) = \chi_j(u)$. Since χ was a valid coloring, either v or u received color j under χ but not both. Let's say $\chi(v) = j$. Then $\chi_j(u) = \chi(u)$.

But $i = \chi_j(v) = \chi_j(u) = \chi(u)$ implies that u received color i and v received color j . But we assumed that there were no edges in G whose endpoints received colors i and j . Thus no such u, v exist and hence χ_j is a valid coloring. $\square$

1.3 Limit theorems and random walks

Note 8. Oscillation of a symmetric random walk

Statement. Let $X_1, X_2, \dots$ be independent random variables with $\Pr(X_i = 1) = \Pr(X_i = -1) = 1/2$, and put $S_n = \sum_{i=1}^n X_i$. Almost surely,

$$\limsup_{n \rightarrow \infty} S_n = \infty, \quad \liminf_{n \rightarrow \infty} S_n = -\infty.$$

Proof. The events

$$A_+ := \{\limsup_{n \rightarrow \infty} S_n = \infty\}, \quad A_- := \{\liminf_{n \rightarrow \infty} S_n = -\infty\}$$

are tail events: changing finitely many increments changes all sufficiently late partial sums by a fixed constant. Kolmogorov's zero-one law gives $\Pr(A_+), \Pr(A_-) \in \{0, 1\}$. By symmetry, these probabilities are equal.

If both probabilities were zero, the sequence (S_n) would be bounded almost surely. However, for each fixed positive integer L , disjoint blocks of L increments are all equal to 1 independently with probability $2^{-L} > 0$. By the second Borel–Cantelli lemma, infinitely many such blocks occur almost surely. Taking the intersection over all L , we obtain arbitrarily long runs of 1's almost surely. A bounded sequence of partial sums cannot have this property. Therefore

$$\Pr\left(\limsup_{n \rightarrow \infty} S_n = \infty\right) = \Pr\left(\liminf_{n \rightarrow \infty} S_n = -\infty\right) = 1.$$

□

Note 27. Typical empirical letter frequencies

Statement. Let $X_1, X_2, \dots$ be independent samples from $\Sigma = \{a_1, \dots, a_r\}$, with $\Pr(X_i = a_s) = p_s$, and set $Y_s = \sum_{j=1}^N 1_{\{X_j = a_s\}}$. For every $\delta > 0$,

$$\Pr\left(\max_{1 \leq s \leq r} \left| \frac{Y_s}{N} - p_s \right| < \delta\right) \longrightarrow 1.$$

Proof. This is an immediate consequence of the weak law of large numbers and the union bound.

□

Note 31. Monte Carlo integration and the strong law

Statement. If $f : [0, 1] \rightarrow \mathbb{R}$ is continuous and $X_1, X_2, \dots$ are independent uniform samples from $[0, 1]$, then almost surely

$$\frac{1}{N} \sum_{i=1}^N f(X_i) \longrightarrow \int_0^1 f(x) dx.$$

Proof. This follows almost directly from the strong law of large numbers. Indeed, since we have that if $X \sim \mathcal{U}[0, 1]$, then

$$\mathbb{E}(f(X)) = \int_{-\infty}^{\infty} f(x) f_X(x) dx = \int_0^1 f(x) \frac{1}{1-0} dx = \int_0^1 f(x) dx,$$

the strong law of large numbers implies that $\frac{1}{N} \sum_{i=1}^N f(X_i)$ converges almost surely to $\int_0^1 f(x) dx$.

□

Note 78. The variance of a biased random walk

Statement. Let $X_1, \dots, X_N$ be independent, with $\Pr(X_i = -1) = p$ and $\Pr(X_i = 1) = 1 - p$. For $S_N = \sum_{i=1}^N X_i$,

$$\mathbb{E}[S_N] = (1 - 2p)N, \quad \text{Var}(S_N) = 4Np(1 - p).$$

Proof. First, we know that

$$\mathbb{E}[(S_N - \mathbb{E}[S_N])^2] = \mathbb{E}[S_N^2] - \mathbb{E}[S_N]^2 = \mathbb{E}[S_N^2] - (1 - 2p)^2 N^2$$

To calculate $\mathbb{E}[S_N^2]$ we first note that

$$\begin{aligned} \mathbb{E}[S_N^2] &= \mathbb{E}[S_{N-1}^2] + 2\mathbb{E}[S_{N-1} \cdot X_N] + \mathbb{E}[X_N^2] \\ &= \mathbb{E}[S_{N-1}^2] + 2(1 - 2p)\mathbb{E}[S_{N-1}] + 1 \\ &= \mathbb{E}[S_{N-1}^2] + 2(1 - 2p)^2(N - 1) + 1 \end{aligned}$$

Solving this equation gives us that

$$\mathbb{E}[S_N^2] = (1 - 2p)^2(N - 1)N + N$$

Thus

$$\begin{aligned} \mathbb{E}[(S_N - \mathbb{E}[S_N])^2] &= (1 - 2p)^2(N - 1)N + N - (1 - 2p)^2 N^2 \\ &= (1 - 2p)^2 N^2 - (1 - 2p)^2 N + N - (1 - 2p)^2 N^2 \\ &= N - (1 - 2p)^2 N \\ &= (1 - (1 - 2p)^2)N \end{aligned}$$

Observe that the deviations are the largest when $p = 1/2$, since $(1 - 2p)^2$ is symmetric around $p = 1/2$. □

Note 160. Rational operations on irrational numbers

Statement. *If r is irrational and q is rational, then $r + q$ is irrational. If also $q \neq 0$, then rq is irrational.*

Proof. For the product assertion, assume $q \neq 0$. If it were rational, say p , then $r = p/q$. So r would be rational, a contradiction. Similar argument shows that sum of an irrational number and a rational number is irrational. For, if $r + q = p$, then $r = p - q$. □

1.4 Dynamics and Fourier analysis

Note 18. Invariant sets and indicator functions

Statement. *Let T be a measurable transformation of a measure space (X, ν) , and let A be measurable. The identity $1_A \circ T = 1_A$ almost everywhere is equivalent to*

$$\nu(A \setminus T^{-1}(A)) = \nu(T^{-1}(A) \setminus A) = 0,$$

or, equivalently, $\nu(A \Delta T^{-1}(A)) = 0$.

Proof. Suppose first that A is T -invariant. First, we have that

$$(A \setminus T^{-1}(A)) = \{x \mid x \in A \text{ and } T(x) \notin A\} \subseteq \{x \mid 1_A(x) \neq 1_A(T(x))\}$$

Secondly, we have that

$$(T^{-1}(A) \setminus A) = \{x \mid x \notin A \text{ and } T(x) \in A\} \subseteq \{x \mid 1_A(x) \neq 1_A(T(x))\}$$

Thus we can conclude that

$$\nu(A \setminus T^{-1}(A)), \nu(T^{-1}(A) \setminus A) \leq \nu(\{x \mid 1_A(x) \neq 1_A(T(x))\}) = 0$$

Suppose then that $\nu(A \setminus T^{-1}(A)) = 0$ and $\nu(T^{-1}(A) \setminus A) = 0$. Now, as demonstrated above, we have that

$$\{x \mid 1_A(x) \neq 1_A(T(x))\} = (A \setminus T^{-1}(A)) \cup (T^{-1}(A) \setminus A)$$

and hence

$$\begin{aligned} \nu(\{x \mid 1_A(x) \neq 1_A(T(x))\}) &= \nu((A \setminus T^{-1}(A)) \cup (T^{-1}(A) \setminus A)) \\ &\leq \nu(A \setminus T^{-1}(A)) + \nu(T^{-1}(A) \setminus A) \\ &= 0, \end{aligned}$$

i.e., A is T -invariant. □

Note 19. Periodic points of rational rotations

Statement. On the unit circle, let $T_\alpha(e^{i\theta}) = e^{i(\theta+\alpha)}$. If $\alpha/(2\pi) = p/q$ in lowest terms with $q > 0$, then every point has least period q under T_α .

Proof. For $\theta \in [0, 2\pi)$ we note that

$$\theta + q\alpha = \theta + 2\pi p,$$

which then implies that $T_\alpha^q(e^{i\theta}) = e^{i(\theta+q\alpha)} = e^{i\theta}$. Hence every point in $\mathbb{T}$ is periodic for T_α (and more precisely the period of every point has length q). If $T_\alpha^m(z) = z$, then mp/q is an integer. Since p, q are coprime, $q \mid m$, proving minimality. □

Note 20. Periodic points of the doubling map

Statement. For the doubling map $D(z) = z^2$ on the unit circle, the periodic points are precisely

$$\left\{ \exp\left(\frac{2\pi im}{2^n - 1}\right) : n \geq 1, m \in \mathbb{Z} \right\}.$$

This set is dense in the circle, but not every point is periodic.

Proof. We start by noting that

$$\begin{aligned} D(e^{i\frac{\pi}{2}}) &= e^{i\pi} \\ D(e^{i\pi}) &= e^{i2\pi} \\ D(e^{i2\pi}) &= e^{i4\pi} = e^{i2\pi} \end{aligned}$$

and hence there are points in $\mathbb{T}$ that are not periodic for D . On the other hand, if $e^{i\theta}$ is periodic for D , then there must exist $n \in \mathbb{N}$ and $m \in \mathbb{Z}_+$ such that $2^n\theta = \theta + 2\pi m$. Thus every point in the set

$$P_D := \left\{ e^{i\theta} \mid \theta = \frac{2\pi m}{2^n - 1}, n \in \mathbb{N}, m \in \mathbb{Z}_+ \right\}$$

is periodic for D . Since $\{\frac{m}{2^n - 1} \mid n \in \mathbb{N}, m \in \mathbb{Z}_+\}$ is dense in $[0, 1)$, we have that $\{\frac{2\pi m}{2^n - 1} \mid n \in \mathbb{N}, m \in \mathbb{Z}_+\}$ is dense in $[0, 2\pi)$. Hence P_D is dense in $\mathbb{T}$, i.e., the set of points in $\mathbb{T}$ which are periodic for D are dense in $\mathbb{T}$. □

Note 21. The tent map preserves Lebesgue measure

Statement. *The tent map*

$$T(x) = \begin{cases} 2x, & 0 \leq x \leq 1/2, \\ 2 - 2x, & 1/2 < x \leq 1 \end{cases}$$

preserves Lebesgue measure on $[0, 1]$.

Proof. Let $A \subseteq [0, 1]$ be a Lebesgue measurable set. Now

$$\nu(T^{-1}(A)) = \nu\left(T^{-1}(A) \cap \left[0, \frac{1}{2}\right]\right) + \nu\left(T^{-1}(A) \cap \left(\frac{1}{2}, 1\right]\right)$$

The mapping $T \upharpoonright [0, \frac{1}{2}]$ is just the mapping $x \mapsto 2x$, so

$$\nu\left(T^{-1}(A) \cap \left[0, \frac{1}{2}\right]\right) = \frac{1}{2}\nu(A)$$

The mapping $T \upharpoonright (\frac{1}{2}, 1]$ is a composition of a translation, a reflection, and the scaling $x \mapsto 2x$, and hence

$$\nu\left(T^{-1}(A) \cap \left(\frac{1}{2}, 1\right]\right) = \frac{1}{2}\nu(A)$$

In total we have that $\nu(T^{-1}(A)) = \nu(A)$. □

Note 22. Pointwise recurrence on the unit interval

Statement. Let $T : [0, 1] \rightarrow [0, 1]$ preserve Lebesgue measure. For almost every x , there is a strictly increasing sequence (n_i) of natural numbers such that $T^{n_i}(x) \rightarrow x$.

Proof. Fix a positive integer $i \geq 1$ and consider the following family $\{]\frac{j-1}{i}, \frac{j}{i}[\mid 1 \leq j \leq i \}$ of sub-intervals of $[0, 1]$. As each of the intervals $] \frac{j-1}{i}, \frac{j}{i}[$ is measurable, by Poincaré's recurrence theorem we have that almost all the points $x \in] \frac{j-1}{i}, \frac{j}{i}[$ are recurrent. Hence, for almost all the points $x \in [0, 1]$ there exists infinitely many natural numbers n such that $|x - T^n(x)| < \frac{1}{i}$. Let R_i denote the set of all such points.

Since each set R_i has measure one, also $\bigcap_{i \in \mathbb{Z}_+} R_i$ has measure one. Hence almost all the points $x \in [0, 1]$ are such that for every $i \in \mathbb{Z}_+$ there exists infinitely many natural numbers n such that $|x - T^n(x)| < \frac{1}{i}$. Now we can construct a sequence $(n_i)_{i \in \mathbb{Z}_+}$ such that $T^{n_i}(x) \rightarrow x$ as $i \rightarrow \infty$ as follows. Suppose that we have constructed a sequence $(n_1, \dots, n_i)$ such that $|x - T^{n_j}(x)| < \frac{1}{j}$, for every $1 \leq j \leq i$. Now we know that there exists $n > n_i$ such that $|x - T^n(x)| < \frac{1}{i+1}$.

We then simply set $n_{i+1} := n$. It is clear that this gives us the desired sequence. □

Note 23. Sine coefficients of an even function

Statement. If $f \in L^2[-\pi, \pi]$ is even almost everywhere, then all of its sine Fourier coefficients vanish:

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx = 0 \quad (n \geq 1).$$

Proof. This can be seen as follows.

$$\begin{aligned} \int_{-\pi}^{\pi} f(x) \sin(nx) \, dx &= \int_{-\pi}^0 f(x) \sin(nx) \, dx + \int_0^{\pi} f(x) \sin(nx) \, dx \\ &= \int_0^{\pi} f(-x) \sin(-nx) \, dx + \int_0^{\pi} f(x) \sin(nx) \, dx \\ &= - \int_0^{\pi} f(x) \sin(nx) \, dx + \int_0^{\pi} f(x) \sin(nx) \, dx \\ &= 0 \end{aligned}$$
□

Note 24. An integral bound from Cauchy–Schwarz

Statement. For $f \in L^2[-\pi, \pi]$ and every integer $n \geq 0$,

$$\left(\int_{-\pi}^{\pi} |x^n f(x)| \, dx \right)^2 \leq \frac{2\pi^{2n+1}}{2n+1} \int_{-\pi}^{\pi} |f(x)|^2 \, dx.$$

Proof. This follows from Cauchy–Schwarz as follows

$$\begin{aligned} \left(\int_{-\pi}^{\pi} |x^n f(x)| dx \right)^2 &\leq \left(\int_{-\pi}^{\pi} |x|^{2n} dx \right) \left(\int_{-\pi}^{\pi} |f(x)|^2 dx \right) \\ &= \left(\frac{2}{2n+1} x^{2n+1} \Big|_{-\pi}^{\pi} \right) \left(\int_{-\pi}^{\pi} |f(x)|^2 dx \right) \\ &= \frac{2\pi^{2n+1}}{2n+1} \int_{-\pi}^{\pi} |f(x)|^2 dx \end{aligned}$$

□

Note 25. The Fourier series of the identity function

Statement. For $f(x) = x$ on $[-\pi, \pi]$, the Fourier coefficients are $A_0 = A_n = 0$ and $B_n = 2(-1)^{n+1}/n$, for $n \geq 1$. Thus

$$x \sim 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx).$$

The series converges to x for $-\pi < x < \pi$ and to zero at the endpoints.

Proof. The constant and cosine coefficients vanish because their integrands are odd. Integration by parts gives the sine coefficients:

$$\begin{aligned} B_n &= \frac{1}{\pi} \int_{-\pi}^{\pi} x \sin(nx) dx \\ &= \frac{1}{\pi} \left(-\frac{x \cos(nx)}{n} \Big|_{-\pi}^{\pi} + \frac{1}{n} \int_{-\pi}^{\pi} \cos(nx) dx \right) \\ &= \frac{2(-1)^{n+1}}{n}. \end{aligned}$$

Thus the Fourier series is

$$x \sim 2 \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} \sin(nx).$$

It converges to x for $-\pi < x < \pi$ and to 0 at the endpoints.

□

1.5 Metric geometry

Note 37. Covering numbers in doubling metric spaces

Statement. Let (X, ρ) be a nonempty bounded metric space of doubling dimension d : every ball of radius r is covered by at most 2^d balls of radius $r/2$. If $r_X > \text{diam}(X)$, then for $0 < r \leq r_X$, the space X is covered by at most $(2r_X/r)^d$ balls of radius r .

Proof. Since (X, ρ) is bounded, there exists $r_X > 0$ such that $\rho(x, x') < r_X$, for any $x, x' \in X$. Thus X itself can be covered by a single ball with radius r_X . Applying iteratively our assumption that (X, ρ) has doubling dimension d , we see that for any $n \geq 0$ we have that X can be covered by 2^{nd} -balls with radius $r_X/2^n$.

We claim that one can take $C_d := (2r_X)^d$. To see this, fix $r > 0$ and let $n = \lceil \log_2(r_X/r) \rceil$, where $\log$ is a binary logarithm. Then

$$\frac{r_X}{2^n} \leq r$$

Hence X can be covered by 2^{nd} -many balls with radius r . Since

$$2^{nd} \leq 2^{d(\log_2(r_X/r)+1)} = (2r_X)^d r^{-d} = C_d r^{-d},$$

we are done.

The constant depends on the scale of X , as well as on d . □

Note 38. Lower mass bounds for doubling measures

Statement. *Let μ be a measure on a bounded metric space X , with $0 < \mu(X) < \infty$, and suppose that*

$$\mu(B(x, r)) \leq 2^d \mu(B(x, r/2))$$

for every $x \in \text{supp}(\mu)$ and $r > 0$. If $r_X > \text{diam}(X)$, then

$$\mu(B(x, r)) \geq \mu(X) \left(\frac{r}{2r_X} \right)^d \quad (0 < r \leq r_X).$$

Proof. Since (X, ρ) is bounded, there exists $r_X > 0$ such that $\rho(x, x') < r_X$, for any $x, x' \in X$. Thus X itself can be covered by a single ball with radius r_X . By applying iteratively the assumption that μ has doubling with exponent d , one obtains that for every $x \in \text{supp}(\mu)$ and $n \in \mathbb{N}$ we have that $\mu(X) \leq 2^{nd} \cdot \mu(B(x, r_X/2^n))$, which gives us that $2^{-nd} \cdot \mu(X) \leq \mu(B(x, r_X/2^n))$. By setting $n = \lceil \log_2(r_X/r) \rceil$, where $\log$ is the binary logarithm, we get that $(2r_X)^{-d} r^d \mu(X) \leq \mu(B(x, r))$.

Setting $C_d := (2r_X)^{-d} \mu(X)$ gives us the desired result. □

Note 39. From a doubling measure to a doubling metric

Statement. *Suppose a metric space X carries a doubling measure μ with full support, exponent d , and positive finite measure on every ball. Every ball of radius r can then be covered by at most 2^{4d} balls of radius $r/2$. In particular, the metric space has doubling dimension at most $4d$.*

Proof. Fix $r > 0$ and $x \in X$. By assumption, we have that $\mu(B(x, r)) \leq 2^d \cdot \mu(B(x, r/2))$. We construct recursively a sequence of points as follows. First, we pick some $x_0 \in B(x, r)$. Suppose then that we have picked $x_0, \dots, x_{k-1}$.

If possible, we pick any point $x_k \in B(x, r) \setminus \bigcup_{\ell=0}^{k-1} B(x_\ell, r/2)$; otherwise we stop picking points. We want to bound the number of points we had to pick in terms of d , the exponent of μ . If we managed to pick the points $x_0, \dots, x_{k-1}$, then we have the inequality

$$\sum_{\ell=0}^{k-1} \mu(B(x_\ell, r/4)) = \mu\left(\bigcup_{\ell=0}^{k-1} B(x_\ell, r/4)\right) \leq \mu(B(x, 3r/2))$$

Furthermore, for each $0 \leq \ell < k$ we have that $\mu(B(x, 3r/2)) \leq \mu(B(x_\ell, 5r/2))$. Applying the fact that μ has exponent d we obtain that

$$\mu(B(x, 3r/2)) \leq 2^{4d} \mu(B(x_\ell, r/4)).$$

Combining this with what we already got we get that $k \leq 2^{4d} = 2^{\mathcal{O}(d)}$. Thus $B(x, r)$ can be covered using $2^{\mathcal{O}(d)}$ balls of radius $r/2$. □

Note 75. Packing and covering numbers

Statement. Let $N(\varepsilon, X, \rho)$ be the minimum number of closed ε -balls needed to cover a metric space X , and let $M(\varepsilon, X, \rho)$ be its packing number for pairwise distances greater than ε . Allowing infinite values, for every $\varepsilon > 0$,

$$M(2\varepsilon, X, \rho) \leq N(\varepsilon, X, \rho) \leq M(\varepsilon, X, \rho).$$

Proof. We start with the second inequality. Let $\{\theta_1, \dots, \theta_M\}$ be an ε -packing of (X, ρ) of size $M := M(\varepsilon, X, \rho)$. We claim that it is also an ε -cover of (X, ρ) . Indeed, if this were not the case, there would exist a point $\theta \in X$ such that $\rho(\theta, \theta_i) > \varepsilon$, for every $1 \leq i \leq M$. But then $\{\theta_1, \dots, \theta_M, \theta\}$ would also be an ε -packing of (X, ρ) , which would contradict the maximality of M .

We then consider the first inequality. Let $\{\theta_1, \dots, \theta_N\}$ be an ε -cover of (X, ρ) of size $N := N(\varepsilon, X, \rho)$. We claim that (X, ρ) has no 2ε -packing of size $N + 1$. Aiming for a contradiction, suppose that $\{\theta_1^*, \dots, \theta_{N+1}^*\}$ is such a packing. Now, for every $1 \leq i \leq N + 1$ there exists $1 \leq j \leq N$ such that $\rho(\theta_i^*, \theta_j) \leq \varepsilon$. Pigeonhole principle implies that there exists $1 \leq j \leq N$ and two distinct indices $1 \leq i_1 < i_2 \leq N + 1$ such that $\rho(\theta_{i_1}^*, \theta_j), \rho(\theta_{i_2}^*, \theta_j) \leq \varepsilon$.

Applying triangle inequality, we get that $\rho(\theta_{i_1}^*, \theta_{i_2}^*) \leq 2\varepsilon$, which contradicts our assumption that $\{\theta_1^*, \dots, \theta_{N+1}^*\}$ was a 2ε -packing. □

2. Stochastic processes

Throughout this chapter, $W(0) = 0$, and W is standard Brownian motion with its usual filtration. The notation M_T^2 denotes admissible adapted processes f with $\mathbb{E} \int_0^T |f(t)|^2 dt < \infty$.

2.1 Brownian motion

Note 9. The mean of Brownian motion

Statement. *If W is standard Brownian motion, then $\mathbb{E}[W(t)] = 0$ for every $t \geq 0$.*

Proof. For $t = 0$ the conclusion is immediate. For $t > 0$, the Gaussian density gives

$$\begin{aligned} \mathbb{E}(W(t)) &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} x e^{-\frac{x^2}{2t}} dx \\ &= \frac{-t}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} \underbrace{-\frac{x}{t} e^{-\frac{x^2}{2t}}}_{=f'(x)e^{f(x)}} dx \\ &= \frac{-t}{\sqrt{2\pi t}} e^{-\frac{x^2}{2t}} \Big|_{-\infty}^{\infty} \\ &= 0 \end{aligned}$$

□

Note 10. The variance of Brownian motion

Statement. *For standard Brownian motion and every $t \geq 0$,*

$$\text{Var}(W(t)) = \mathbb{E}[W(t)^2] = t.$$

Proof. For $t = 0$ the conclusion is immediate. For $t > 0$, we compute it from the Gaussian density.

$$\begin{aligned} \text{Var}(W(t)) &= \mathbb{E}((W(t) - \underbrace{\mathbb{E}(W(t))}_{=0})^2) \\ &= \mathbb{E}(W(t)^2) \\ &= \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2t}} dx \end{aligned}$$

Integrating by parts we can compute the integral as follows.

$$\begin{aligned}
 & \int_{-\infty}^{\infty} x^2 e^{-\frac{x^2}{2t}} dx \\
 &= -t \int_{-\infty}^{\infty} x \cdot \underbrace{-\frac{x}{t} e^{-\frac{x^2}{2t}}}_{=\frac{d}{dx} e^{-\frac{x^2}{2t}}} dx \\
 &= -t \left(\underbrace{xe^{-\frac{x^2}{2t}}}_{=0} \Big|_{-\infty}^{\infty} - \underbrace{\int_{-\infty}^{\infty} e^{-\frac{x^2}{2t}} dx}_{=\sqrt{2\pi t}} \right) \\
 &= t\sqrt{2\pi t}
 \end{aligned}$$

Hence

$$\text{Var}(W(t)) = t$$

□

Note 11. The covariance of Brownian motion

Statement. For standard Brownian motion and $s, t \geq 0$,

$$\text{Cov}(W(s), W(t)) = \mathbb{E}[W(s)W(t)] = \min\{s, t\}.$$

Proof. By symmetry, assume $s \geq t$. The increment $W(s) - W(t)$ is independent of $W(t)$ and has mean zero. Therefore

$$\begin{aligned}
 \mathbb{E}[W(s)W(t)] &= \mathbb{E}[W(t)^2] + \mathbb{E}[(W(s) - W(t))W(t)] \\
 &= t + \mathbb{E}[W(s) - W(t)]\mathbb{E}[W(t)] \\
 &= t.
 \end{aligned}$$

□

Note 12. Mean-square Brownian increments

Statement. The increments of standard Brownian motion satisfy

$$\mathbb{E}[|W(t) - W(s)|^2] = |t - s| \quad (s, t \geq 0).$$

Proof. Using the covariance formula in Note 11, we obtain the second moment of a Brownian increment:

$$\begin{aligned}
 \mathbb{E}[|W(t) - W(s)|^2] &= \mathbb{E}(W^2(t) - 2W(t)W(s) + W^2(s)) \\
 &= \mathbb{E}(W^2(t)) + \mathbb{E}(W^2(s)) - 2\mathbb{E}(W(t)W(s)) \\
 &= t + s - 2\min\{s, t\} \\
 &= |t - s|
 \end{aligned}$$

□

Note 13. The Gaussian characteristic function

Statement. For standard Brownian motion, $t \geq 0$, and $\lambda \in \mathbb{R}$,

$$\mathbb{E}[e^{i\lambda W(t)}] = e^{-t\lambda^2/2}.$$

Proof. For $t > 0$, let $\chi_t(\lambda) := \mathbb{E}[e^{i\lambda W(t)}]$. Then

$$\chi_t(\lambda) = \frac{1}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{i\lambda x} e^{-x^2/(2t)} dx.$$

Differentiating under the integral and integrating by parts gives

$$\begin{aligned} \chi_t'(\lambda) &= \frac{i}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} x e^{i\lambda x} e^{-x^2/(2t)} dx \\ &= -\frac{it}{\sqrt{2\pi t}} \int_{-\infty}^{\infty} e^{i\lambda x} \frac{d}{dx} e^{-x^2/(2t)} dx \\ &= -t\lambda \chi_t(\lambda). \end{aligned}$$

The boundary term vanishes. Since $\chi_t(0) = 1$, solving this differential equation yields

$$\mathbb{E}[e^{i\lambda W(t)}] = e^{-t\lambda^2/2}.$$

The same identity holds for $t = 0$. □

Note 14. Brownian motion as a martingale

Statement. Standard Brownian motion $(W_t)_{t \geq 0}$ is a martingale with respect to its usual filtration $(\mathcal{F}_t)$. In particular,

$$\mathbb{E}[W_t \mid \mathcal{F}_s] = W_s \quad \text{almost surely whenever } 0 \leq s \leq t.$$

Proof. This follows from the following chain of equalities.

$$\begin{aligned} \mathbb{E}(W_t \mid \mathcal{F}_s) &= \mathbb{E}((W_t - W_s) + W_s \mid \mathcal{F}_s) \\ &= \mathbb{E}(W_t - W_s \mid \mathcal{F}_s) + \mathbb{E}(W_s \mid \mathcal{F}_s) \\ &= \mathbb{E}(W_t - W_s) + W_s \\ &= W_s, \end{aligned}$$

where we used the fact that $W_t - W_s$ is independent of $\mathcal{F}_s$ and that $W_t - W_s \sim \mathcal{N}(0, t - s)$. □

Note 15. Quadratic variation in mean square

Statement. Fix $T \geq 0$, and put $t_i^n = iT/n$ and $\Delta_i^n W = W(t_{i+1}^n) - W(t_i^n)$. Then

$$\mathbb{E} \left[\left(\sum_{i=0}^{n-1} (\Delta_i^n W)^2 - T \right)^2 \right] = \frac{2T^2}{n} \longrightarrow 0.$$

Thus Brownian quadratic variation along the uniform partitions converges to T in L^2 .

Proof. First, we have that

$$\begin{aligned}\mathbb{E}\left(\left[\sum_{i=0}^{n-1}(\Delta_i^n W)^2 - T\right]^2\right) &= \mathbb{E}\left(\left[\sum_{i=0}^{n-1}(\Delta_i^n W)^2\right]^2 - 2T \sum_{i=0}^{n-1}(\Delta_i^n W)^2 + T^2\right) \\ &= \mathbb{E}\left(\left[\sum_{i=0}^{n-1}(\Delta_i^n W)^2\right]^2\right) - 2T \sum_{i=0}^{n-1} \mathbb{E}((\Delta_i^n W)^2) + T^2\end{aligned}$$

Now

$$\mathbb{E}((\Delta_i^n W)^2) = \mathbb{E}((W(t_{i+1}^n) - W(t_i^n))^2) = \frac{T}{n}$$

and hence

$$2T \sum_{i=0}^{n-1} \mathbb{E}((\Delta_i^n W)^2) = 2T \sum_{i=0}^{n-1} \frac{T}{n} = 2T^2$$

Next, we have that

$$\begin{aligned}\mathbb{E}\left(\left[\sum_{i=0}^{n-1}(\Delta_i^n W)^2\right]^2\right) &= \sum_{i=0}^{n-1} \mathbb{E}((\Delta_i^n W)^4) + 2 \sum_{i < j} \mathbb{E}((\Delta_i^n W)^2 (\Delta_j^n W)^2) \\ &= \sum_{i=0}^{n-1} \frac{3T^2}{n^2} + 2 \sum_{i < j} \frac{T^2}{n^2} \\ &= \frac{3T^2}{n} + 2 \binom{n}{2} \frac{T^2}{n^2} \\ &= T^2 + \frac{2T^2}{n} \rightarrow T^2,\end{aligned}$$

which completes the proof. □

Note 16. A Brownian stochastic sum

Statement. Let W be standard Brownian motion, $T \geq 0$, and $t_j^n = jT/n$. In L^2 ,

$$\sum_{j=0}^{n-1} W(t_j^n)(W(t_{j+1}^n) - W(t_j^n)) \longrightarrow \frac{1}{2}(W(T)^2 - T).$$

Proof. First, we note that

$$W(t_j^n)(W(t_{j+1}^n) - W(t_j^n)) = \frac{1}{2}(W(t_{j+1}^n)^2 - W(t_j^n)^2) - \frac{1}{2}(W(t_j^n) - W(t_{j+1}^n))^2$$

Now we know that

$$\lim_{n \rightarrow \infty} \frac{1}{2} \sum_{j=0}^{n-1} (W(t_j^n) - W(t_{j+1}^n))^2 = \frac{T}{2}$$

and since

$$\sum_{j=0}^{n-1} (W(t_{j+1}^n)^2 - W(t_j^n)^2) = W(T)^2$$

we conclude that the limit is $\frac{1}{2}(W(T)^2 - T)$.

□

Note 44. Scaling Brownian hitting times

Statement. For standard Brownian motion, define $\tau(a) = \inf\{t \geq 0 : W(t) = a\}$, where $a \geq 0$. For every $b > 0$,

$$\tau(a) \stackrel{\mathcal{D}}{=} b\tau(a/\sqrt{b}).$$

Proof.

$$\begin{aligned} b\tau(a/\sqrt{b}) &= b \min\{t : W(t) = a/\sqrt{b}\} \\ &= b \min\{t : \sqrt{b}W(t) = a\} \\ &= \min\{t : \sqrt{b}W(t/b) = a\} \end{aligned}$$

Since $\sqrt{b}W(t/b) \stackrel{\mathcal{D}}{=} W(t)$, we deduce that $\tau(a) \stackrel{\mathcal{D}}{=} b\tau(a/\sqrt{b})$.

□

2.2 Conditional expectation and stochastic integration

Note 45. Conditional expectation as a projection

Statement. Let $(\Omega, \mathcal{F}, \mu)$ be a probability space and $\mathcal{G} \subseteq \mathcal{F}$ a σ -algebra. For $X \in L^2(\mathcal{F})$, the conditional expectation $\mathbb{E}[X | \mathcal{G}]$ is the orthogonal projection of X onto $L^2(\mathcal{G})$, with inner product $\langle X, Y \rangle = \mathbb{E}[XY]$.

Proof. Let $P : L^2(\mathcal{F}) \rightarrow L^2(\mathcal{G})$ denote the orthogonal projection from $L^2(\mathcal{F})$ to $L^2(\mathcal{G})$. Thus

$$\langle X - PX, Y \rangle = 0,$$

for every $Y \in L^2(\mathcal{G})$. In particular

$$\langle X - PX, 1_A \rangle = 0,$$

for every $A \in \mathcal{G}$. In other words

$$\int_A X \, d\mu = \int_A PX \, d\mu,$$

for every $A \in \mathcal{G}$. Since $\mathbb{E}(X|\mathcal{G})$ is simply the unique $\mathcal{G}$ -measurable function that satisfies the above equality, we conclude that $PX = \mathbb{E}(X|\mathcal{G})$.

□

Note 17. Cross-isometry for stochastic step integrals

Statement. Let f, g be adapted square-integrable step processes, and write $I(f) = \int_0^\infty f(t) dW(t)$. Then

$$\mathbb{E}[I(f)I(g)] = \mathbb{E} \left[\int_0^\infty f(t)g(t) dt \right].$$

Proof. Refining their partitions, write

$$f = \sum_{i=0}^{n-1} \eta_i 1_{[t_i, t_{i+1})}, \quad g = \sum_{i=0}^{n-1} \zeta_i 1_{[t_i, t_{i+1})},$$

where $0 = t_0 < \dots < t_n$ and η_i, ζ_i are $\mathcal{F}_{t_i}$ -measurable. Put $\Delta_i W = W(t_{i+1}) - W(t_i)$ and $\Delta_i t = t_{i+1} - t_i$. We show that

$$\mathbb{E}[I(f)I(g)] = \mathbb{E} \left[\int_0^\infty f(t)g(t) dt \right].$$

Expanding the two stochastic integrals gives

$$\mathbb{E}[I(f)I(g)] = \sum_{i,j=0}^{n-1} \mathbb{E}[\eta_i \zeta_j \Delta_i W \Delta_j W].$$

On the diagonal, independence of the increment from $\mathcal{F}_{t_i}$ yields

$$\mathbb{E}[\eta_i \zeta_i (\Delta_i W)^2] = \mathbb{E}[\eta_i \zeta_i] \Delta_i t.$$

If $i < j$, then $\eta_i \zeta_j \Delta_i W$ is $\mathcal{F}_{t_j}$ -measurable. Conditioning on $\mathcal{F}_{t_j}$, the zero mean of $\Delta_j W$ makes the corresponding expectation zero. The case $i > j$ is symmetric. Hence

$$\mathbb{E}[I(f)I(g)] = \sum_{i=0}^{n-1} \mathbb{E}[\eta_i \zeta_i] \Delta_i t = \mathbb{E} \left[\int_0^\infty f(t)g(t) dt \right].$$

□

Note 62. Stochastic step integrals are martingales

Statement. If f is an adapted square-integrable step process, then

$$I_t(f) = \int_0^t f(s) dW(s)$$

is a square-integrable martingale. Consequently, $\mathbb{E}[I_t(f) \mid \mathcal{F}_s] = I_s(f)$ almost surely for $0 \leq s \leq t$.

Proof. Let $f \in M_{\text{step}}^2$ be an adapted square-integrable step process,

$$f(u) = \sum_{j=0}^{n-1} \eta_j 1_{[t_j, t_{j+1})}(u), \quad 0 = t_0 < t_1 < \dots < t_n,$$

where η_j is $\mathcal{F}_{t_j}$ -measurable. Truncation to any interval $[0, t)$ gives another square-integrable step process, so $f \in M_t^2$ for every $t > 0$.

We show that $I_t(f) = \int_0^t f(u) dW(u)$ is a martingale. Fix $0 \leq s < t$. Refine the partition by inserting s and t . On every subinterval $[u_j, u_{j+1}) \subseteq [s, t)$, the corresponding coefficient ξ_j is $\mathcal{F}_{u_j}$ -measurable. Independence and the zero mean of Brownian increments give

$$\mathbb{E}[\xi_j(W(u_{j+1}) - W(u_j)) \mid \mathcal{F}_{u_j}] = 0.$$

Since $\mathcal{F}_s \subseteq \mathcal{F}_{u_j}$, the tower property implies

$$\mathbb{E}[I_t(f) - I_s(f) \mid \mathcal{F}_s] = 0.$$

The process is adapted and square-integrable by the Itô isometry. Consequently $\mathbb{E}[I_t(f) \mid \mathcal{F}_s] = I_s(f)$ almost surely. □

Note 63. Square-integrability of Brownian motion

Statement. For every $T > 0$, standard Brownian motion is an admissible square-integrable stochastic integrand on $[0, T]$. More precisely,

$$\mathbb{E} \left[\int_0^T |W(t)|^2 dt \right] = \frac{T^2}{2} < \infty.$$

Proof. It suffices to show that

$$\mathbb{E} \left[\int_0^T |W(t)|^2 dt \right] < \infty$$

since $W(t)$ has a.s. continuous paths. Using Doob's maximal L^2 inequality we deduce that

$$\begin{aligned} \mathbb{E} \left[\int_0^T |W(t)|^2 dt \right] &\leq \mathbb{E} \left[T \max_{s \leq T} |W(s)|^2 \right] \\ &\leq 4T \mathbb{E} |W(T)|^2 \\ &= 4T^2 < \infty \end{aligned}$$

Alternatively, we can proceed as follows

$$\begin{aligned} \mathbb{E} \left[\int_0^T |W(t)|^2 dt \right] &= \int_0^T \mathbb{E}[|W(t)|^2] dt \\ &= \int_0^T t dt \\ &= \frac{1}{2} T^2 < \infty \end{aligned}$$

□

Note 64. The stochastic integral of Brownian motion

Statement. For standard Brownian motion and $T \geq 0$, almost surely

$$\int_0^T W(t) dW(t) = \frac{1}{2}(W(T)^2 - T).$$

Proof. For each $n \in \mathbb{Z}_+$, we let $0 = t_0^n < t_1^n < \dots < t_n^n = T$ denote a partition of $[0, T]$ into parts that have equal length, namely T/n . By Note 16, we know that

$$\lim_{n \rightarrow \infty} \sum_{j=0}^{n-1} W(t_j^n)(W(t_{j+1}^n) - W(t_j^n)) = \frac{1}{2}W(T)^2 - \frac{1}{2}T$$

where the limit is taken in the L^2 -sense. For every $n \in \mathbb{Z}_+$ we consider the following step function

$$f_n = \sum_{j=0}^{n-1} W(t_j^n) 1_{[t_j^n, t_{j+1}^n)}$$

We first show that $\{f_n\}$ approximates $1_{[0,T)}W(t)$. This can be established as follows.

$$\begin{aligned} \mathbb{E} \left[\int_0^\infty |1_{[0,T)}W(t) - f_n(t)|^2 dt \right] &= \mathbb{E} \left[\int_0^T |W(t) - f_n(t)|^2 dt \right] \\ &= \mathbb{E} \left[\sum_{j=0}^{n-1} \int_{t_j^n}^{t_{j+1}^n} |W(t) - W(t_j^n)|^2 dt \right] \\ &= \sum_{j=0}^{n-1} \int_{t_j^n}^{t_{j+1}^n} \mathbb{E}[|W(t) - W(t_j^n)|^2] dt \\ &= \sum_{j=0}^{n-1} \int_{t_j^n}^{t_{j+1}^n} (t - t_j^n) dt \\ &= \frac{1}{2} \sum_{j=0}^{n-1} (t_{j+1}^n - t_j^n)^2 \\ &= \frac{1}{2} \frac{T^2}{n} \rightarrow 0 \end{aligned}$$

We also have that

$$\lim_{n \rightarrow \infty} \mathbb{E}[|I_T(f) - I(f_n)|^2] = 0,$$

where $I_T(f) := \frac{1}{2}W(T)^2 - \frac{1}{2}T$, since we have already established that $I(f_n) \xrightarrow{L^2} I_T(f)$. $\square$

Note 65. Integration by parts against Brownian motion

Statement. For standard Brownian motion and $T \geq 0$, almost surely

$$\int_0^T t dW(t) = TW(T) - \int_0^T W(t) dt.$$

Proof. For each $n \in \mathbb{Z}_+$, we let $0 = t_0^n < t_1^n < \dots < t_n^n = T$ denote a partition of $[0, T]$ into parts that have equal length, namely T/n . For every $n \in \mathbb{Z}_+$ we consider the following step function

$$f_n = \sum_{j=0}^{n-1} t_j^n 1_{[t_j^n, t_{j+1}^n)}$$

Clearly $\{f_n\}$ approximates $1_{[0, T]}t$. Now, we have that

$$I(f_n) = \sum_{j=0}^{n-1} t_j^n (W(t_{j+1}^n) - W(t_j^n))$$

and we want to determine the limit of this sum in the L^2 -sense. Applying the identity

$$c(b - a) = (db - ca) - b(d - c)$$

we can rewrite the sum as follows

$$\begin{aligned} \sum_{j=0}^{n-1} t_j^n (W(t_{j+1}^n) - W(t_j^n)) &= \sum_{j=0}^{n-1} \left((t_{j+1}^n W(t_{j+1}^n) - t_j^n W(t_j^n)) - W(t_{j+1}^n)(t_{j+1}^n - t_j^n) \right) \\ &= \underbrace{\sum_{j=0}^{n-1} (t_{j+1}^n W(t_{j+1}^n) - t_j^n W(t_j^n))}_{= \text{telescopic series}} - \sum_{j=0}^{n-1} W(t_{j+1}^n)(t_{j+1}^n - t_j^n) \\ &= TW(T) - \sum_{j=0}^{n-1} W(t_{j+1}^n)(t_{j+1}^n - t_j^n) \end{aligned}$$

Now we claim that $\sum_{j=0}^{n-1} W(t_{j+1}^n)(t_{j+1}^n - t_j^n) \xrightarrow{L^2} \int_0^T W(t) dt$. Indeed, we have the following:

$$\begin{aligned} &\mathbb{E} \left[\left| \int_0^T W(t) dt - \sum_{j=0}^{n-1} W(t_{j+1}^n)(t_{j+1}^n - t_j^n) \right|^2 \right] \\ &= \mathbb{E} \left[\left| \sum_{j=0}^{n-1} \int_{t_j^n}^{t_{j+1}^n} W(t_{j+1}^n) - W(t) dt \right|^2 \right] \\ &\leq n \sum_{j=0}^{n-1} \mathbb{E} \left[\left(\int_{t_j^n}^{t_{j+1}^n} W(t_{j+1}^n) - W(t) dt \right)^2 \right] \\ &\leq n \sum_{j=0}^{n-1} (t_{j+1}^n - t_j^n) \mathbb{E} \left[\int_{t_j^n}^{t_{j+1}^n} (W(t_{j+1}^n) - W(t))^2 dt \right] \\ &= n \sum_{j=0}^{n-1} (t_{j+1}^n - t_j^n) \int_{t_j^n}^{t_{j+1}^n} (t_{j+1}^n - t) dt \\ &= \frac{T^3}{2n} \rightarrow 0 \end{aligned}$$

Hence we are done. □

Note 66. Square-integrability of a stochastic integral

Statement. Let f be an admissible adapted integrand satisfying $\mathbb{E} \int_0^T |f(s)|^2 ds < \infty$. The stochastic integral $\xi(t) = \int_0^t f(s) dW(s)$ is itself an admissible square-integrable integrand on $[0, T]$, and

$$\mathbb{E} \int_0^T |\xi(t)|^2 dt \leq T \mathbb{E} \int_0^T |f(s)|^2 ds.$$

Proof. Without loss of generality, we can assume that ξ has a.s. continuous paths and thus it suffices to show that

$$\mathbb{E} \left(\int_0^T |\xi(t)|^2 dt \right) < \infty$$

Applying the Itô isometry, we see that

$$\begin{aligned} \mathbb{E} \left(\int_0^T |\xi(t)|^2 dt \right) &= \mathbb{E} \left(\int_0^T \left| \int_0^t f(s) dW(s) \right|^2 dt \right) \\ &= \int_0^T \mathbb{E} \left(\left| \int_0^t f(s) dW(s) \right|^2 \right) dt \\ &= \int_0^T \mathbb{E} \left(\int_0^t |f(s)|^2 ds \right) dt \\ &\leq T \mathbb{E} \left(\int_0^T |f(s)|^2 ds \right) < \infty \end{aligned}$$

Since $f \in M_t^2$, we know that

$$\mathbb{E} \left(\int_0^t |f(s)|^2 ds \right) < \infty$$

and hence we also have that

$$\mathbb{E} \left(\int_0^T |\xi(t)|^2 dt \right) < \infty$$

□

Note 67. The exponential Brownian martingale

Statement. The process $X(t) = \exp(W(t) - t/2)$ satisfies $X(0) = 1$ and the stochastic differential equation

$$dX(t) = X(t) dW(t).$$

Proof. Let $F(t, x) := e^x e^{-t/2}$. Applying the Itô formula we see that

$$\begin{aligned} &F(T, W(T)) - F(0, W(0)) \\ &= \int_0^T (F'_t(t, W(t)) + \frac{1}{2} F''_{xx}(t, W(t))) dt + \int_0^T F'_x(t, W(t)) dW(t) \end{aligned}$$

Since

$$F'_t(t, x) = -\frac{1}{2} e^x e^{-t/2}$$

and

$$F''_{xx}(t, x) = e^{x-t/2}$$

we have that

$$\int_0^T (F'_t(t, W(t)) + \frac{1}{2} F''_{xx}(t, W(t))) dt = 0$$

Hence, using the differential notation, we have that

$$de^{W(t)-t/2} = e^{W(t)-t/2} dW(t),$$

which is what we wanted to show. □

2.3 Linear stochastic differential equations

Note 68. An affine Ornstein–Uhlenbeck equation

Statement. For $\alpha > 0$ and real constants β, σ, x_0 , the solution of

$$dX_t = (-\alpha X_t + \beta) dt + \sigma dW(t), \quad X_0 = x_0,$$

is

$$X_t = x_0 e^{-\alpha t} + \frac{\beta}{\alpha} (1 - e^{-\alpha t}) + \sigma \int_0^t e^{-\alpha(t-s)} dW(s).$$

Its distribution converges to $\mathcal{N}(\beta/\alpha, \sigma^2/(2\alpha))$ as $t \rightarrow \infty$.

Proof. We look for a solution in the class of processes that can be written as

$$X_t = a(t) \left\{ x_0 + \int_0^t c(s) ds + \int_0^t b(s) dW(s) \right\},$$

where $a(\cdot), b(\cdot)$ and $c(\cdot)$ are differentiable functions. Applying the product rule we see that

$$\begin{aligned} dX_t &= a'(t) \left\{ x_0 + \int_0^t c(s) ds + \int_0^t b(s) dW(s) \right\} dt + a(t)c(t) dt + a(t)b(t) dW(t) \\ &= \frac{a'(t)}{a(t)} X_t dt + a(t)c(t) dt + a(t)b(t) dW(t) \\ &= \left(\frac{a'(t)}{a(t)} X_t + a(t)c(t) \right) dt + a(t)b(t) dW(t) \end{aligned}$$

As in the solution of the Ornstein–Uhlenbeck SDE, we see that $a(t) = e^{-\alpha t}$ and $b(t) = \sigma e^{\alpha t}$. Furthermore, given the constraint $a(t)c(t) = \beta$, we see that $c(t) = \beta e^{\alpha t}$. Since

$$\beta \int_0^t e^{\alpha s} ds = \frac{\beta}{\alpha} (e^{\alpha t} - 1),$$

we have the following solution

$$\begin{aligned} X_t &= e^{-\alpha t} \left\{ x_0 + \frac{\beta}{\alpha} (e^{\alpha t} - 1) + \sigma \int_0^t e^{\alpha s} dW(s) \right\} \\ &= x_0 e^{-\alpha t} + \frac{\beta}{\alpha} (1 - e^{-\alpha t}) + \sigma \int_0^t e^{-\alpha(t-s)} dW(s) \end{aligned}$$

Letting $t \rightarrow \infty$, we see that $\mathbb{E}(X_t) \rightarrow \frac{\beta}{\alpha}$. Similarly to the case of Ornstein-Uhlenbeck SDE we also have that $\text{Var}(X_t) \rightarrow \frac{\sigma^2}{2\alpha}$. Thus the distribution of X_t converges to a Gaussian distribution with mean $\frac{\beta}{\alpha}$ and variance $\frac{\sigma^2}{2\alpha}$. $\square$

Note 69. A linear SDE with a time-dependent drift

Statement. *The stochastic differential equation*

$$dX_t = tX_t dt + e^{t^2/2} dW(t), \quad X_0 = 0,$$

has solution $X_t = e^{t^2/2}W(t)$.

Proof. We look for a solution in the class of processes that can be written as

$$X_t = a(t) \int_0^t b(s) dW(s),$$

where $a(\cdot)$ and $b(\cdot)$ are differentiable functions and $a(0) = 1$. Applying the product rule we see that

$$\begin{aligned} dX_t &= a'(t) \int_0^t b(s) dW(s) dt + a(t)b(t) dW(t) \\ &= \frac{a'(t)}{a(t)} X_t dt + a(t)b(t) dW(t) \end{aligned}$$

From this we get the constraints $a'(t)/a(t) = t$ and $a(t)b(t) = e^{t^2/2}$. Solving the first differential equation we get that $a(t) = e^{t^2/2}$ (recall that $a(0) = 1$) and hence we also get that $b(t) = 1$. Thus

$$X_t = e^{t^2/2} \cdot \int_0^t 1 dW(s) = e^{t^2/2} \cdot W(t)$$

$\square$

Note 70. A linear SDE with a vanishing diffusion

Statement. *For $0 \leq t < 1$, the equation*

$$dX_t = -\frac{2X_t}{1-t} dt + \sqrt{2t(1-t)} dW(t), \quad X_0 = 0,$$

has the centred solution

$$X_t = (1-t)^2 \int_0^t \frac{\sqrt{2s(1-s)}}{(1-s)^2} dW(s).$$

For $0 \leq s \leq t < 1$, its covariance is $\text{Cov}(X_t, X_s) = s^2(1-t)^2$.

Proof. We look for a solution in the class of processes that can be written as

$$X_t = a(t) \int_0^t b(s) dW(s)$$

where a is differentiable and b is locally square-integrable and $a(0) = 1$. As in Note 69, we have that

$$dX_t = \frac{a'(t)}{a(t)} X_t dt + a(t)b(t) dW(t)$$

Matching coefficients we see that $a'(t)/a(t) = -2/(1-t)$, which implies that $a(t) = (1-t)^2$. Furthermore, given the constraint $a(t)b(t) = \sqrt{2t(1-t)}$, we see that $b(t) = \sqrt{2t(1-t)}/(1-t)^2$. Thus we have that

$$X_t = (1-t)^2 \int_0^t \frac{\sqrt{2s(1-s)}}{(1-s)^2} dW(s)$$

Next we calculate $\text{Cov}(X_t, X_s)$, for $0 \leq s \leq t < 1$. First, for every $t \in [0, 1)$ we have that $\mathbb{E}(X_t) = 0$. Thus we only need to calculate $\mathbb{E}(X_t X_s)$. This can be done as follows.

$$\begin{aligned} \mathbb{E}(X_t X_s) &= (1-t)^2 (1-s)^2 \mathbb{E} \left(\int_0^t \frac{\sqrt{2u(1-u)}}{(1-u)^2} dW(u) \int_0^s \frac{\sqrt{2v(1-v)}}{(1-v)^2} dW(v) \right) \\ &= (1-t)^2 (1-s)^2 \mathbb{E} \left(\left| \int_0^s \frac{\sqrt{2v(1-v)}}{(1-v)^2} dW(v) \right|^2 \right) \\ &= (1-t)^2 (1-s)^2 \int_0^s \frac{2v(1-v)}{(1-v)^4} dv \\ &= (1-t)^2 (1-s)^2 \frac{s^2}{(1-s)^2} \\ &= s^2 (1-t)^2 \end{aligned}$$

Hence $\text{Cov}(X_t, X_s) = s^2(1-t)^2$, for $0 \leq s \leq t < 1$.

□

Note 71. Moments of geometric Brownian motion

Statement. Let $X_0 = x_0$ be deterministic and let $dX_t = \mu X_t dt + \sigma X_t dW(t)$. Then

$$X_t = x_0 e^{(\mu - \sigma^2/2)t + \sigma W(t)}, \quad \mathbb{E}[X_t] = x_0 e^{\mu t},$$

and $\text{Var}(X_t) = x_0^2 e^{2\mu t} (e^{\sigma^2 t} - 1)$.

Proof. Since the above SDE has a unique solution (which we already know), we have that

$$X_t = x_0 \exp \left(\left(\mu - \frac{1}{2} \sigma^2 \right) t + \sigma W(t) \right),$$

where $X_0 = x_0$. Now, we know that

$$\exp \left(\sigma W(t) - \frac{1}{2} \sigma^2 t \right)$$

is a martingale, which means that

$$\begin{aligned}\mathbb{E}\left(\exp\left(\sigma W(t) - \frac{1}{2}\sigma^2 t\right)\right) &= \mathbb{E}\left(\mathbb{E}\left(\exp\left(\sigma W(t) - \frac{1}{2}\sigma^2 t\right)\middle|\mathcal{F}_0\right)\right) \\ &= \mathbb{E}\left(\exp\left(\sigma W(0) - \frac{1}{2}\sigma^2 \cdot 0\right)\right) \\ &= \mathbb{E}(1) = 1\end{aligned}$$

Hence we have that

$$\mathbb{E}(X_t) = x_0 \exp(\mu t)$$

To calculate the variance of X_t we note that

$$\begin{aligned}\mathbb{E}\left(\exp\left(2\sigma W(t) - \sigma^2 t\right)\right) &= \exp(\sigma^2 t) \mathbb{E}\left(\exp\left(2\sigma W(t) - 2\sigma^2 t\right)\right) \\ &= \exp(\sigma^2 t)\end{aligned}$$

and hence

$$\text{Var}(X_t) = x_0^2 \exp(2\mu t) \exp(\sigma^2 t) - x_0^2 \exp(2\mu t) = x_0^2 \exp(2\mu t) (\exp(\sigma^2 t) - 1)$$

□

2.4 Exponential waiting times and Poisson processes

Note 46. Moments of the exponential distribution

Statement. If X has the exponential distribution with rate $\lambda > 0$, then

$$\mathbb{E}[X] = \frac{1}{\lambda}, \quad \text{Var}(X) = \frac{1}{\lambda^2}.$$

Proof. First, we have that

$$\begin{aligned}\mathbb{E}(X) &= \int_0^\infty \lambda t e^{-\lambda t} dt \\ &= \lambda \int_0^\infty t e^{-\lambda t} dt \\ &= \lambda \left(\underbrace{-\frac{t}{\lambda} e^{-\lambda t}}_{=0} \bigg|_0^\infty + \frac{1}{\lambda} \underbrace{\int_0^\infty e^{-\lambda t} dt}_{=1/\lambda} \right) \\ &= \frac{1}{\lambda}\end{aligned}$$

Secondly, we have that

$$\begin{aligned}
 \mathbb{E}(X^2) &= \int_0^\infty \lambda t^2 e^{-\lambda t} dt \\
 &= \lambda \int_0^\infty t^2 e^{-\lambda t} dt \\
 &= \lambda \underbrace{-\frac{t^2}{\lambda} e^{-\lambda t}}_{=0} \Big|_0^\infty + 2 \int_0^\infty t e^{-\lambda t} dt \\
 &= \frac{2}{\lambda} \underbrace{\int_0^\infty \lambda t e^{-\lambda t} dt}_{=1/\lambda} \\
 &= \frac{2}{\lambda^2}
 \end{aligned}$$

Hence

$$\text{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{1}{\lambda^2}$$

□

Note 47. The mean of a Poisson process

Statement. If $(N(t))_{t \geq 0}$ is a Poisson process with rate $\lambda > 0$, then $\mathbb{E}[N(t)] = \lambda t$ for every $t \geq 0$.

Proof.

$$\begin{aligned}
 \mathbb{E}(N(t)) &= \sum_{n=0}^{\infty} n \Pr(N(t) = n) \\
 &= \sum_{n=1}^{\infty} n \frac{(\lambda t)^n}{n!} e^{-\lambda t} \\
 &= \lambda t e^{-\lambda t} \sum_{n=1}^{\infty} \frac{(\lambda t)^{n-1}}{(n-1)!} \\
 &= \lambda t \cdot e^{-\lambda t} \cdot e^{\lambda t} \\
 &= \lambda t
 \end{aligned}$$

□

Note 48. Strictly increasing Poisson arrival times

Statement. Let $\eta_1, \eta_2, \dots$ be independent exponential waiting times with rate $\lambda > 0$, and put $\xi_0 = 0$ and $\xi_n = \sum_{j=1}^n \eta_j$. Almost surely,

$$0 = \xi_0 < \xi_1 < \xi_2 < \dots$$

Proof. We first note that for each $n \geq 1$ we have that

$$\Pr\{\xi_{n-1} < \xi_n\} = \Pr\{\eta_n > 0\} = 1$$

In other words, a.s. we have that $\xi_n > \xi_{n-1}$. Thus a.s. we have that for every $n > 0$, $\xi_n > \xi_{n-1}$. □

Note 49. Divergence of Poisson arrival times

Statement. *For the arrival times $\xi_n = \sum_{j=1}^n \eta_j$ associated with independent exponential waiting times of rate $\lambda > 0$, almost surely*

$$\frac{\xi_n}{n} \longrightarrow \frac{1}{\lambda}, \quad \text{and hence} \quad \xi_n \longrightarrow \infty.$$

Proof. Recall that ξ_n is a sum of i.i.d. random variables each having $1/\lambda$ as the expected value, for some $\lambda > 0$. Thus the strong law of large numbers entails that a.s.

$$\frac{\xi_n}{n} \rightarrow \frac{1}{\lambda}$$

as $n \rightarrow \infty$. Clearly this also implies that a.s. $\xi_n \rightarrow \infty$. □

3. Information, randomization and cryptography

3.1 Entropy and coding

Note 16a. Shannon entropy as a Rényi limit

Statement. Let $p_1, \dots, p_n \geq 0$ sum to one. With $0 \log 0 = 0$, the Rényi entropy converges to Shannon entropy as its order tends to one:

$$\lim_{\alpha \rightarrow 1} \frac{\log(\sum_{i=1}^n p_i^\alpha)}{1 - \alpha} = - \sum_{i=1}^n p_i \log p_i.$$

Proof. Letting

$$f(\alpha) := \log\left(\sum_{i=1}^n p_i^\alpha\right)$$

we observe that by definition

$$\lim_{\alpha \rightarrow 1} H_R^\alpha(p_1, \dots, p_n) = - \left(\frac{d}{d\alpha} f(\alpha) \right)_{\alpha=1},$$

since $f(1) = 0$. Now

$$\frac{d}{d\alpha} f(\alpha) = \frac{1}{\sum_{i=1}^n p_i^\alpha} \cdot \sum_{i=1}^n p_i^\alpha \log(p_i)$$

Taking the negative of this derivative at $\alpha = 1$ gives $H_S(p_1, \dots, p_n)$. Zero-probability terms are omitted, with the convention $0 \log 0 = 0$. □

Note 26. Information per symbol and the strong law

Statement. Let $X_1, X_2, \dots$ be independent samples from the finite alphabet $\{a_1, \dots, a_r\}$, with $\Pr(X_j = a_s) = p_s$. Writing $P_N = \prod_{j=1}^N p_{X_j}$, almost surely

$$-\frac{1}{N} \log P_N \longrightarrow - \sum_{s=1}^r p_s \log p_s,$$

where zero-probability symbols are omitted and p_{X_j} denotes the mass of the observed symbol.

Proof. First, we note that we can rewrite the left-hand side as follows:

$$\frac{1}{N} \log \frac{1}{\Pr(D_N)} = \frac{1}{N} \sum_{j=1}^N -\log \Pr(X_j),$$

where $\Pr(X_j)$ is just the random variable

$$\Pr(X_j)((\ell_1, \dots, \ell_N)) := \Pr(X_j = \ell_j)$$

Consider now the following sequence of random variables $(Y_j)_{j=1}^\infty$, where

$$Y_j := -\log \Pr(X_j).$$

This is clearly a sequence of i.i.d. random variables. Furthermore,

$$\mathbb{E}(Y_j) = -\sum_{s=1}^r \Pr(X_j = a_s) \log \Pr(X_j = a_s) = -\sum_{s=1}^r p_s \log(p_s)$$

The claim follows now from the strong law of large numbers. □

Note 29. An entropy identity for finite partitions

Statement. Let $(C_i)_{i \in I}$ be a partition of $[n]$ into nonempty sets, and let X be uniform on $[n]$. If $p_i = \Pr(X \in C_i) = |C_i|/n$, then

$$\log n = -\sum_{i \in I} p_i \log p_i + \sum_{i \in I} p_i \log |C_i|.$$

Proof. For every $i \in I$ and $m \in C_i$ we have that

$$\underbrace{\Pr(X = m)}_{=1/n} = \Pr(X \in C_i) \cdot \underbrace{\Pr(X = m \mid X \in C_i)}_{=1/|C_i|}.$$

Taking logs and multiplying both sides with -1 we get

$$\log(n) = -\log \Pr(X \in C_i) + \log(|C_i|)$$

This identity holds for every $i \in I$. Hence

$$\begin{aligned} \log(n) &= \sum_{i \in I} \Pr(X \in C_i) \log(n) \\ &= -\sum_{i \in I} \Pr(X \in C_i) \log \Pr(X \in C_i) + \sum_{i \in I} \Pr(X \in C_i) \log(|C_i|) \end{aligned}$$

□

Note 52. Kraft's inequality

Statement. If $c_1, \dots, c_N$ are distinct words in a prefix-free binary code and ℓ_i is the length of c_i , then

$$\sum_{i=1}^N 2^{-\ell_i} \leq 1.$$

Proof. Let $m = \max\{\ell_i \mid i = 1, \dots, N\}$ and consider the uniform probability distribution μ over $\{0, 1\}^m$. For each $i = 1, \dots, N$ we define the event

$$A_i := \{x \in \{0, 1\}^m \mid c_i \text{ occurs as a prefix in } x\}$$

Since $\{c_1, \dots, c_N\}$ is a prefix-free encoding, we have that $A_i \cap A_j = \emptyset$, for every $i \neq j$. Clearly $\mu(A_i) = 2^{-\ell_i}$, and hence

$$\sum_{i=1}^N 2^{-\ell_i} = \sum_{i=1}^N \mu(A_i) = \mu\left(\bigcup_{i=1}^N A_i\right) \leq 1,$$

which is what we wanted to show. □

3.2 Randomized computation

Note 30. Zero-error randomized computation

Statement. *The class of languages decidable by zero-error randomized algorithms in expected polynomial time satisfies*

$$\text{ZPP} = \text{RP} \cap \text{coRP}.$$

Proof. We start by showing that $\text{ZPP} \subseteq \text{RP} \cap \text{coRP}$. Let $L \in \text{ZPP}$. Then there exists a probabilistic algorithm A and a polynomial p such that A decides L correctly in expected time $p(|x|)$, i.e., $\mathbb{E}(A(x)) \leq p(|x|)$, where $A(x)$ is the random variable that counts the number of steps that A takes on input x . Markov's inequality gives us that

$$\Pr(A(x) > 2p(|x|)) \leq \frac{\mathbb{E}(A(x))}{2p(|x|)} \leq \frac{1}{2}$$

Consider now the following randomized polynomial time algorithm.

Input: $x \in \{0, 1\}^*$

1. Simulate A on input x for $2p(|x|)$ steps.
2. **if** A stops during this simulation **then** return $A(x)$
3. return 0

This algorithm will never output 1 if $x \notin L$, since A decides every input correctly. Furthermore, since $A(x)$ takes more than $2p(|x|)$ steps with probability at most $1/2$, this algorithm is correct on instances $x \in L$ with probability at least $1/2$. Hence $L \in \text{RP}$. If we modify it so that it returns by default 1, then we get a coRP algorithm for recognizing the language L .

Next we show that $\text{RP} \cap \text{coRP} \subseteq \text{ZPP}$. Let $L \in \text{RP} \cap \text{coRP}$ and let A_1, A_2 be the corresponding algorithms. Consider the following probabilistic algorithm.

Input: $x \in \{0, 1\}^*$

1. **while true do**
2. $b_1 \leftarrow A_1(x)$
3. $b_2 \leftarrow A_2(x)$
4. **if** $b_1 = b_2$ **then** return b_1

Note that this algorithm always outputs a correct answer. Indeed, if $x \in L$, then always $A_2(x) = 1$ and furthermore $A_1(x) = 1$ with positive probability. Likewise, if $x \notin L$, then always $A_1(x) = 0$ and furthermore $A_2(x) = 0$ with positive probability. Finally, if X is the random variable which counts the number of times the while loop was executed, then we have that

$$\mathbb{E}(X) = \sum_{k=0}^{\infty} \Pr(X > k) \leq \sum_{k=0}^{\infty} 2^{-k} = 2$$

and hence the expected number of steps this algorithm takes is clearly polynomial with respect to $|x|$. □

Note 33. Covering the Boolean cube by translates

Statement. Let $\ell \geq 1$, let $S \subseteq \{0, 1\}^\ell$, and write $S \oplus z = \{s \oplus z : s \in S\}$. If $|S| \leq 2^\ell/(4\ell)$, the union of any ℓ translates has size at most $2^\ell/4$. If $|S| \geq (1 - 1/(4\ell))2^\ell$, there exist $z_1, \dots, z_\ell$ such that

$$\bigcup_{i=1}^{\ell} (S \oplus z_i) = \{0, 1\}^\ell.$$

Proof. Suppose that $S \subseteq \{0, 1\}^\ell$ is a set of size at most $2^\ell/(4\ell)$. Then for any $z_1, \dots, z_\ell \in \{0, 1\}^\ell$ we have that

$$\left| \bigcup_{i=1}^{\ell} (S \oplus z_i) \right| \leq \sum_{i=1}^{\ell} |S \oplus z_i| = \sum_{i=1}^{\ell} |S| \leq \ell \cdot 2^\ell/(4\ell) = 2^\ell/4.$$

Suppose then that $|S| \geq (1 - 1/(4\ell))2^\ell$. We claim that there exists $z_1, \dots, z_\ell \in \{0, 1\}^\ell$ such that $\bigcup_{i=1}^{\ell} (S \oplus z_i) = \{0, 1\}^\ell$. To prove this, we use a probabilistic argument. Pick $z_1, \dots, z_\ell \in \{0, 1\}^\ell$ independently at random. Now

$$\begin{aligned} \Pr \left[\exists x \in \{0, 1\}^\ell : x \notin \bigcup_{i=1}^{\ell} (S \oplus z_i) \right] &\leq \sum_{x \in \{0, 1\}^\ell} \Pr \left[x \notin \bigcup_{i=1}^{\ell} (S \oplus z_i) \right] \\ &= \sum_{x \in \{0, 1\}^\ell} \Pr \left[\forall 1 \leq i \leq \ell : x \notin (S \oplus z_i) \right] \\ &= \sum_{x \in \{0, 1\}^\ell} \prod_{i=1}^{\ell} \Pr[x \notin (S \oplus z_i)] \\ &\leq \sum_{x \in \{0, 1\}^\ell} \prod_{i=1}^{\ell} \frac{1}{4\ell} \\ &= 2^\ell \left(\frac{1}{4\ell} \right)^\ell \\ &< 1 \end{aligned}$$

Thus there must exist $z_1, \dots, z_\ell \in \{0, 1\}^\ell$ such that $\bigcup_{i=1}^{\ell} (S \oplus z_i) = \{0, 1\}^\ell$. □

Note 34. Randomized NP protocols with polynomial advice

Statement. Suppose a probabilistic polynomial-time mapping M produces a Boolean formula such that $\Pr(3\text{SAT}(M(x)) = 1_L(x)) \geq 2/3$ for every input x . Then $L \in \text{NP/poly}$. Consequently, $\text{BP} \cdot \text{NP} \subseteq \text{NP/poly}$.

Proof. Let $L \in \text{BP} \cdot \text{NP}$. Thus there exists a probabilistic polynomial time TM M such that

$$\forall x \in \{0, 1\}^* : \Pr[L(x) = 3\text{SAT}(M(x))] \geq \frac{2}{3}$$

Note that the composition of 3SAT and M can be seen as a Arthur-Merlin protocol. Verifier first computes $M(x)$ obtaining (an encoding of) a 3CNF-formula ϕ and then sends it as a message to the prover. Prover returns an assignment and the verifier accepts iff this assignment satisfies ϕ . Let N denote this protocol. Using standard error reduction, we can assume that

$$\forall x \in \{0, 1\}^* : \Pr[L(x) = N(x, r)] \geq 1 - 2^{-(|x|+1)},$$

where r denotes the coins used by the verifier. Just as in the proof of $\text{BPP} \subseteq \text{P/poly}$, for each input length n there exists r_n such that for every $x \in \{0, 1\}^n$ we have that $L(x) = N(x, r_n)$. But $N(x, r_n)$ is just a non-deterministic polynomial time TM with advice and hence it can be clearly implemented as a family of NP/poly-circuits. □

Note 56. A PSPACE-complete oracle collapses P and NP

Statement. If B is PSPACE-complete under polynomial-time many-one reductions, then $\text{P}^B = \text{NP}^B$.

Proof. Clearly $\text{P}^B \subseteq \text{NP}^B$. Suppose then that $L \in \text{NP}^B$. Thus there exists a polynomial p and a polynomial time Turing machine $M(x, y)$ which has access to an oracle for B such that

$$\forall x \in \{0, 1\}^* : x \in L \Leftrightarrow \exists y \in \{0, 1\}^{p(|x|)} : M(x, y) = 1$$

Since $B \in \text{PSPACE}$, the language $\{x\#y \mid M(x, y) = 1\}$ is clearly in PSPACE. Hence, also the language L is clearly in PSPACE, since one can loop through the members of $\{0, 1\}^{p(|x|)}$ using only polynomial space. Since B is PSPACE-complete, there exists a polynomial time computable mapping $f : \{0, 1\}^* \rightarrow \{0, 1\}^*$ such that $x \in L$ iff $f(x) \in B$.

A polynomial Turing machine which has access to the oracle B can then clearly compute L : given x , compute $f(x)$ and ask the oracle whether or not $f(x) \in B$. □

3.3 Perfect secrecy

Note 54. A key-length obstruction to perfect secrecy

Statement. Consider a deterministic encryption scheme with correct decryption for every key, m -bit plaintexts, and uniformly chosen n -bit keys, where $n < m$. There exist plaintexts x, x' and a ciphertext y for which

$$\Pr_k(E_k(x) = y) \neq \Pr_k(E_k(x') = y).$$

Thus such a scheme cannot have perfect secrecy.

Proof. We first establish the following stronger property: there exist messages x, x' and a key k such that

$$E_k(x) \neq E_{k'}(x'),$$

for every $k' \in \{0, 1\}^n$. First, since E_k is an injection for every $k \in \{0, 1\}^n$, we have that the range of each E_k has size exactly 2^m . However, for a fixed plaintext x the set $\text{supp}(x) := \{E_k(x) \mid k \in \{0, 1\}^n\}$ is of size at most $2^n < 2^m$. Hence, if we fix any plaintext x and a key k , then there must exist $x' \in \{0, 1\}^m$ such that $E_k(x') \notin \text{supp}(x)$. In other words, for any $k' \in \{0, 1\}^n$ we have that $E_k(x') \neq E_{k'}(x)$, which is what we wanted to show. $\square$

Note 55. Predicting one bit through a one-time pad

Statement. Let X, K be independent and uniform on $\{0, 1\}^n$, and set $C = X \oplus K$. For any fixed index $i \in [n]$ and function $A_i : \{0, 1\}^n \rightarrow \{0, 1\}$,

$$\Pr(A_i(C) = X_i) = \frac{1}{2}.$$

Proof. Indeed, for every plaintext x and ciphertext c ,

$$\Pr(X = x, C = c) = \Pr(X = x, K = x \oplus c) = 2^{-2n}.$$

Thus X and C are independent and uniform. Conditional on $C = c$, the prediction $A_i(c)$ is fixed while X_i remains a fair bit. Averaging over c proves the identity.

The success probability is exactly one half, rather than merely at most one half. If a Boolean predictor did worse, reversing its output would produce a predictor that did better. The adaptive choice of which bit to predict is treated in Note 131. $\square$

Note 131. Adaptive bit prediction through a one-time pad

Statement. Let X, K be independent uniform n -bit strings, set $C = X \oplus K$, and let $A : \{0, 1\}^n \rightarrow [n] \times \{0, 1\}$. If $A(C) = (I, B)$, then

$$\Pr(B = X_I) = \frac{1}{2},$$

even though the selected index I may depend on the ciphertext.

Proof. By Note 55, X and C are independent. Conditional on $C = c$, the output $A(c) = (i, b)$ is fixed and X_i remains a fair bit. Hence

$$\Pr(B = X_I) = \mathbb{E}[\Pr(B = X_I \mid C)] = \mathbb{E}[1/2] = \frac{1}{2}.$$

□

4. Learning and optimization

4.1 Estimation and moments

Note 41. The expectation and variance of a sample mean

Statement. Let $X_1, \dots, X_n$ be independent samples with mean μ and finite variance σ^2 , and define $\boldsymbol{\mu}_S = n^{-1} \sum_{i=1}^n X_i$. Then

$$\mathbb{E}[\boldsymbol{\mu}_S] = \mu, \quad \text{Var}(\boldsymbol{\mu}_S) = \frac{\sigma^2}{n}.$$

Proof. Clearly $\mathbb{E}(\boldsymbol{\mu}_S) = \mathbb{E}(X)$. Furthermore

$$\begin{aligned} \text{Var}(\boldsymbol{\mu}_S) &= \frac{1}{n^2} \text{Var} \left(\sum_{i=1}^n X_i \right) \\ &= \frac{1}{n^2} \left(\sum_{i=1}^n \text{Var}(X_i) + \sum_{i \neq j} \underbrace{\text{Cov}(X_i, X_j)}_{=0} \right) \\ &= \frac{1}{n} \text{Var}(X) \end{aligned}$$

Here $\text{Cov}(X_i, X_j) = 0$ follows from the fact that X_i and X_j are independent. □

Note 42. The bias of the sample variance

Statement. For independent samples $X_1, \dots, X_n$ with mean μ and variance $\sigma^2 < \infty$, let $\boldsymbol{\mu}_S = n^{-1} \sum_i X_i$ and $\boldsymbol{\sigma}_S^2 = n^{-1} \sum_i (X_i - \boldsymbol{\mu}_S)^2$. Then

$$\mathbb{E}[\boldsymbol{\sigma}_S^2] = \frac{n-1}{n} \sigma^2.$$

For $n \geq 2$, replacing the denominator n by $n-1$ gives an unbiased estimator of σ^2 .

Proof. Now

$$\mathbb{E}(\boldsymbol{\sigma}_S^2) = \frac{1}{n} \sum_{i=1}^n \mathbb{E}(X_i - \boldsymbol{\mu}_S)^2$$

To calculate $\mathbb{E}(X_i - \boldsymbol{\mu}_S)^2$, we first do the following computation

$$\begin{aligned} (X_i - \boldsymbol{\mu}_S)^2 &= ((X_i - \mathbb{E}(X)) - (\boldsymbol{\mu}_S - \mathbb{E}(X)))^2 \\ &= (X_i - \mathbb{E}(X))^2 - 2(X_i - \mathbb{E}(X))(\boldsymbol{\mu}_S - \mathbb{E}(X)) + (\boldsymbol{\mu}_S - \mathbb{E}(X))^2 \end{aligned}$$

This gives us

$$\begin{aligned} \mathbb{E}(X_i - \boldsymbol{\mu}_S)^2 &= \text{Var}(X) - 2 \text{Cov}(X_i, \boldsymbol{\mu}_S) + \text{Var}(\boldsymbol{\mu}_S) \\ &= \text{Var}(X) - 2 \text{Cov}(X_i, \boldsymbol{\mu}_S) + \frac{1}{n} \text{Var}(X) \end{aligned}$$

Observe that

$$\begin{aligned}\text{Cov}(X_i, \boldsymbol{\mu}_S) &= \frac{1}{n} \left(\text{Var}(X_i) + \sum_{i \neq j} \underbrace{\text{Cov}(X_i, X_j)}_{=0} \right) \\ &= \frac{1}{n} \text{Var}(X)\end{aligned}$$

Hence

$$\mathbb{E}(X_i - \boldsymbol{\mu}_S)^2 = \frac{(n-1)}{n} \text{Var}(X),$$

which gives us that

$$\mathbb{E}(\boldsymbol{\sigma}_S^2) = \frac{(n-1)}{n} \text{Var}(X)$$

On the other hand, if we were to define sample variance as

$$\boldsymbol{\sigma}_S^2 := \frac{1}{n-1} \sum_{i=1}^n (X_i - \boldsymbol{\mu}_S)^2,$$

then we would get that $\mathbb{E}(\boldsymbol{\sigma}_S^2) = \text{Var}(X)$.

□

Note 43. A lower bound involving absolute moments

Statement. If $0 < \mathbb{E}|X| < \infty$ and $\mathbb{E}[X^2] < \infty$, then

$$\frac{\mathbb{E}[X^2]}{\mathbb{E}|X|} \geq \frac{|\mathbb{E}[X]| + \sqrt{\text{Var}(X)}}{2}.$$

Proof. First, by applying Jensen's inequality, we obtain that $\mathbb{E}[|X|] \leq \sqrt{\mathbb{E}[X^2]}$. Hence it suffices to show that

$$2\sqrt{\mathbb{E}[X^2]} \geq |\mathbb{E}[X]| + \sqrt{\text{Var}(X)}$$

First we note that clearly $\sqrt{\mathbb{E}[X^2]} \geq \sqrt{\mathbb{E}[X^2] - \mathbb{E}[X]^2} = \sqrt{\text{Var}(X)}$. Secondly, by applying Jensen again, we see that

$$\sqrt{\mathbb{E}[X^2]} \geq \sqrt{\mathbb{E}[X]^2} = |\mathbb{E}[X]|$$

The desired inequality follows from these two observations.

□

Note 76. The median-of-means estimate

Statement. Let independent samples have mean μ and variance $\sigma^2 < \infty$. Fix $\varepsilon > 0$ and $0 < \delta < 1$. Divide KN samples into K blocks of size N , where

$$N \geq \max\{1, \lceil 4\sigma^2/\varepsilon^2 \rceil\}, \quad K \geq 32 \log(1/\delta)$$

and K is odd. The median of the block means lies within ε of μ with probability at least $1 - \delta$.

Proof. Let $X_1, \dots, X_N$ be i.i.d. samples from a fixed random variable X with mean μ and variance σ^2 . Consider the empirical mean $\hat{\mu} := \frac{1}{N} \sum_{i=1}^N X_i$. Now $\mathbb{E}[\hat{\mu}] = \mu$ and $\text{Var}[\hat{\mu}] = \sigma^2/N$. Hence, by applying Chebyshev's inequality we obtain that

$$\Pr[|\hat{\mu} - \mu| \geq \varepsilon] \leq \frac{\sigma^2}{N\varepsilon^2}$$

By setting $N := \max\{1, \lceil 4\sigma^2/\varepsilon^2 \rceil\}$ we obtain that $\Pr[|\hat{\mu} - \mu| \leq \varepsilon] \geq 3/4$.

Suppose now that we take KN i.i.d. samples from the same random variable X , where $N := \max\{1, \lceil 4\sigma^2/\varepsilon^2 \rceil\}$ and K will be optimized later. For each $1 \leq k \leq K$ we thus have N samples $X_{k,1}, \dots, X_{k,N}$. For each k we set $\hat{\mu}_k := \frac{1}{N} \sum_{i=1}^N X_{k,i}$. Now, with probability $p \geq 3/4$ we have that $\hat{\mu}_k \in (\mu - \varepsilon, \mu + \varepsilon)$. For each k we let Y_k denote a Bernoulli random variable for the event “ $\hat{\mu}_k \notin (\mu - \varepsilon, \mu + \varepsilon)$ ”. Set $Y := \sum_{k=1}^K Y_k$. Now $\mathbb{E}[Y] \leq \frac{1}{4}K$ and hence we have that if $Y < \mathbb{E}[Y] + \frac{1}{8}K$, then $Y < \frac{3}{8}K$.

But if the latter inequality holds, then the median of $\hat{\mu}_1, \dots, \hat{\mu}_K$ will be in the interval $(\mu - \varepsilon, \mu + \varepsilon)$. So, it is enough for us to guarantee that with probability at least $1 - \delta$ we have that $Y < \mathbb{E}[Y] + \frac{1}{8}K$. Towards this we use Hoeffding's inequality, which gives us that

$$\Pr\left[Y \geq \mathbb{E}[Y] + \frac{1}{8}K\right] \leq \exp\left(-\frac{1}{32}K\right)$$

Thus, if $K \geq 32 \log(\frac{1}{\delta})$, then the above probability is at most δ . Hence, if we take a sample of size $\mathcal{O}((1 + \sigma^2/\varepsilon^2) \log(1/\delta))$, then with probability at least $1 - \delta$ our estimate for μ — the median of the empirical means — is in the interval $(\mu - \varepsilon, \mu + \varepsilon)$. □

Note 107. Mean distance from the centre of a disk

Statement. Choose a point uniformly from a disk of radius $R > 0$, and let X be its distance from the centre. Then X has density $f_X(x) = 2x/R^2$ on $[0, R]$, and $\mathbb{E}[X] = 2R/3$.

Proof. By definition, we have that

$$\mathbb{E}(X) = \int_0^R x f_X(x) dx.$$

To calculate $f_X(x)$ we proceed as follows. First, note that

$$F_X(x) = \frac{\pi x^2}{\pi R^2} = \frac{x^2}{R^2}.$$

Hence

$$f_X(x) = \frac{d}{dx} F_X(x) = \frac{2x}{R^2}$$

and thus

$$\int_0^R x f_X(x) dx = \int_0^R \frac{2x^2}{R^2} dx = \frac{2x^3}{3R^2} \Big|_0^R = \frac{2R}{3}.$$

□

Note 108. Mean-square error implies consistency

Statement. Let $(\hat{\theta}_n)$ be estimators of θ . If $\mathbb{E}[(\hat{\theta}_n - \theta)^2] \rightarrow 0$, then $\hat{\theta}_n \rightarrow \theta$ in probability. Thus convergence of mean-square error to zero implies consistency.

Proof. This can be proved in the same way that Chebyshev's inequality can be proved using Markov's inequality. Namely, we have that

$$\Pr(|\hat{\theta}_n - \theta| > \varepsilon) = \Pr((\hat{\theta}_n - \theta)^2 > \varepsilon^2) \leq \frac{\mathbb{E}((\hat{\theta}_n - \theta)^2)}{\varepsilon^2} = \frac{\text{MSE}(\hat{\theta}_n)}{\varepsilon^2} \rightarrow 0.$$

That is, $\hat{\theta}_n$ is a consistent estimator for θ . □

Note 109. Conditional expectation and mean-square error

Statement. For a square-integrable random variable X and $\hat{X}_M = \mathbb{E}[X | Y]$,

$$\mathbb{E}[(X - \hat{X}_M)^2] = \mathbb{E}[\text{Var}(X | Y)].$$

Proof. This follows simply from the law of iterated expectation, since

$$\mathbb{E}[\text{Var}(X|Y)] = \mathbb{E}[\mathbb{E}[(X - \hat{X}_M)^2|Y]] = \mathbb{E}[(X - \hat{X}_M)^2].$$

□

Note 110. Variance decomposition in linear regression

Statement. Let $Y = \beta_0 + \beta_1 X + \varepsilon$, where X has finite variance and $\varepsilon \sim \mathcal{N}(0, \sigma^2)$ is independent of X . For $\hat{Y} = \beta_0 + \beta_1 X$,

$$\mathbb{E}[(Y - \mathbb{E}Y)^2] = \mathbb{E}[(\hat{Y} - \mathbb{E}Y)^2] + \mathbb{E}[(Y - \hat{Y})^2].$$

Equivalently, $\text{Var}(Y) = \text{Var}(\hat{Y}) + \sigma^2$.

Proof. Since $\mathbb{E}(Y) = \mathbb{E}(\hat{Y})$, we have that

$$\mathbb{E}[(\hat{Y} - \mathbb{E}(\hat{Y}))^2] = \text{Var}(\hat{Y}).$$

Furthermore

$$\mathbb{E}[(Y - \hat{Y})^2] = \mathbb{E}[\varepsilon^2] = \text{Var}(\varepsilon),$$

as $\mathbb{E}[\varepsilon] = 0$. So our claim is that

$$\text{Var}(Y) = \text{Var}(\hat{Y}) + \text{Var}(\varepsilon).$$

This follows, because $\hat{Y}$ is independent of ε . □

Note 111. Maximum likelihood for a Poisson parameter

Statement. *Given independent observations $X_1, \dots, X_n$ from $\text{Poisson}(\lambda)$, with unknown $\lambda \geq 0$, a maximum-likelihood estimator is*

$$\hat{\lambda} = \frac{1}{n} \sum_{i=1}^n X_i.$$

When all observations are zero, the maximum occurs at $\lambda = 0$.

Proof. Now

$$\mathcal{L}(X_1, \dots, X_n; \lambda) = \prod_{i=1}^n \frac{\lambda^{X_i} e^{-\lambda}}{X_i!} = e^{-\lambda n} \prod_{i=1}^n \frac{\lambda^{X_i}}{X_i!}$$

and hence

$$\log \mathcal{L}(X_1, \dots, X_n; \lambda) = -\lambda n + \sum_{i=1}^n X_i \log(\lambda) - \sum_{i=1}^n \log(X_i!).$$

We want to find λ in terms of $X_1, \dots, X_n$ for which this term is maximized. First, the partial derivative with respect to λ is

$$-n + \sum_{i=1}^n \frac{X_i}{\lambda}.$$

This is 0 when

$$\lambda = \frac{1}{n} \sum_{i=1}^n X_i,$$

which is the MLE for λ .

If every observation is zero, the maximum is attained at the boundary $\lambda = 0$.

□

Note 139. An affine Gaussian observation model

Statement. *Let $x \sim \mathcal{N}(\mu_x, \Sigma_x)$ and $\varepsilon \sim \mathcal{N}(0, Q)$ be independent, and put $y = Ax + b + \varepsilon$ for fixed A, b of compatible dimensions. The pair (y, x) is jointly Gaussian, with*

$$\begin{aligned} \mathbb{E}[y] &= A\mu_x + b, & \text{Var}(y) &= A\Sigma_x A^\top + Q, \\ \text{Cov}(y, x) &= A\Sigma_x, & y \mid x &\sim \mathcal{N}(Ax + b, Q). \end{aligned}$$

Proof. We know that

$$Ax + b \sim \mathcal{N}(A\mu_x + b, A\Sigma_x A^\top)$$

and hence

$$y \sim \mathcal{N}(A\mu_x + b, A\Sigma_x A^\top + Q).$$

Furthermore, we know that

$$(y, x) \sim \mathcal{N}\left(\begin{bmatrix} A\mu_x + b \\ \mu_x \end{bmatrix}, \begin{bmatrix} A\Sigma_x A^\top + Q & A\Sigma_x \\ \Sigma_x A^\top & \Sigma_x \end{bmatrix}\right)$$

and therefore

$$y|x \sim \mathcal{N}(Ax + b, Q).$$

□

4.2 Probability distributions on Boolean types

Note 59. Expected Bayes error under a uniform type prior

Statement. Let τ be a finite set of Boolean features and $q \notin \tau$ a Boolean target. Write T_τ for the feature types and $T_{\tau,q}$ for the complete $\tau \cup \{q\}$ -types. Choose a distribution μ on $T_{\tau,q}$ from the symmetric Dirichlet distribution with all parameters equal to one. Its expected optimal type-based classification error is

$$\mathbb{E}\left[\sum_{t \in T_\tau} \min\{\mu(t \wedge q), \mu(t \wedge \neg q)\}\right] = \frac{1}{4}.$$

Proof. Applying linearity of expectation, the left-hand side is

$$\sum_{t \in T_\tau} \mathbb{E}(\min\{\mu(t \wedge q), \mu(t \wedge \neg q)\}) \quad (59.1)$$

For each $t \in T_\tau$ we define three new random variables:

$$\begin{aligned} \mu(t) &:= \mu(t \wedge q) + \mu(t \wedge \neg q) \\ \mu(q|t) &:= \frac{\mu(q \wedge t)}{\mu(t)} \\ \mu(\neg q|t) &:= \frac{\mu(\neg q \wedge t)}{\mu(t)} \end{aligned}$$

Then (59.1) can be re-written as

$$\sum_{t \in T_\tau} \mathbb{E}(\mu(t) \min\{\mu(q|t), 1 - \mu(q|t)\}) \quad (59.2)$$

Assume first that $K > 2$. Next we will show that $\mu(t)$ is independent from $\mu(q|t)$. To see this, consider the transformation $T(y_1, y_2) = (y_1 y_2, y_1(1 - y_2))$. The absolute value of the Jacobian determinant is y_1 . Since the joint probability distribution of $(\mu(t \wedge q), \mu(t \wedge \neg q))$ is

$$f_{\mu(t \wedge q), \mu(t \wedge \neg q)}(x_1, x_2) = \frac{\Gamma(K)}{\Gamma(K-2)} (1 - x_1 - x_2)^{K-3},$$

where $K = 2^{|\tau \cup \{q\}|}$, the joint probability distribution of $(\mu(t), \mu(q|t))$ is

$$\begin{aligned} g_{\mu(t), \mu(q|t)}(y_1, y_2) &= \frac{\Gamma(K)}{\Gamma(K-2)} y_1 (1 - y_1 y_2 - y_1 + y_1 y_2)^{K-3} \\ &= \frac{\Gamma(K)}{\Gamma(K-2)} y_1 (1 - y_1)^{K-3} \\ &= g_{\mu(t)}(y_1) \times g_{\mu(q|t)}(y_2) \end{aligned}$$

Here the last equality follows from the following first two observations. First, we note that

$$\frac{\Gamma(K)}{\Gamma(K-2)} y_1 (1-y_1)^{K-3}$$

is the probability density function for the distribution $\text{Beta}(2, K-2)$, which is the distribution that $\boldsymbol{\mu}(t)$ follows. Secondly, $g_{\boldsymbol{\mu}(q|t)}(y_2)$ is

$$\int_0^1 g_{\boldsymbol{\mu}(t), \boldsymbol{\mu}(q|t)}(y_1, y_2) dy_1 = \int_0^1 g_{\boldsymbol{\mu}(t)}(y_1) dy_1,$$

which is one, because $g_{\boldsymbol{\mu}(t)}(y_1)$ is a probability density function.

Hence $\boldsymbol{\mu}(t)$ and $\boldsymbol{\mu}(q|t)$ are independent. Thus (59.2) can be simplified to

$$\sum_{t \in T_\tau} \mathbb{E}(\boldsymbol{\mu}(t)) \mathbb{E}(\min\{\boldsymbol{\mu}(q|t), 1 - \boldsymbol{\mu}(q|t)\}) \quad (59.3)$$

Since $g_{\boldsymbol{\mu}(q|t)}(y_2) = 1$ in its support, which is $[0, 1]$, it is clear that

$$\mathbb{E}(\min\{\boldsymbol{\mu}(q|t), 1 - \boldsymbol{\mu}(q|t)\}) = \frac{1}{4}$$

Using this together with the fact that $\sum_{t \in T_\tau} \boldsymbol{\mu}(t) = 1$, we see that (59.3) simplifies to $1/4$, which is what we wanted to show.

If $K = 2$, then $\boldsymbol{\mu}(t) = 1$ almost surely and $\boldsymbol{\mu}(q | t)$ is uniform on $[0, 1]$, so the same conclusion follows directly. □

Note 60. Mean absolute deviation of a beta variable

Statement. If $X \sim \text{Beta}(\alpha, \beta)$, where $\alpha, \beta > 0$, and $\mu = \mathbb{E}[X] = \alpha/(\alpha + \beta)$, then

$$\mathbb{E}|X - \mu| = \frac{2\mu^{\alpha+1}(1-\mu)^\beta}{\alpha B(\alpha, \beta)},$$

where B is the beta function.

Proof. Write $B_z(a, b) = \int_0^z u^{a-1}(1-u)^{b-1} du$ for the incomplete beta function. We begin our calculations as follows:

$$\begin{aligned} & \int_0^1 |x - \mu| x^{\alpha-1} (1-x)^{\beta-1} dx \\ &= \int_0^\mu (\mu - x) x^{\alpha-1} (1-x)^{\beta-1} dx + \int_\mu^1 (x - \mu) x^{\alpha-1} (1-x)^{\beta-1} dx \\ &= \int_0^\mu (\mu - x) x^{\alpha-1} (1-x)^{\beta-1} dx + \int_0^1 (x - \mu) x^{\alpha-1} (1-x)^{\beta-1} dx \\ &\quad - \int_0^\mu (x - \mu) x^{\alpha-1} (1-x)^{\beta-1} dx \\ &= \mu B_\mu(\alpha, \beta) - B_\mu(\alpha + 1, \beta) + \underbrace{B(\alpha + 1, \beta) - \mu B(\alpha, \beta)}_{=0} - B_\mu(\alpha + 1, \beta) + \mu B_\mu(\alpha, \beta) \\ &= 2(\mu B_\mu(\alpha, \beta) - B_\mu(\alpha + 1, \beta)) \end{aligned}$$

Thus

$$\mathbb{E}[|X - \mu|] = \frac{2(\mu B_\mu(\alpha, \beta) - B_\mu(\alpha + 1, \beta))}{B(\alpha, \beta)}$$

and hence it suffices to show that

$$\mu B_\mu(\alpha, \beta) - B_\mu(\alpha + 1, \beta) = \frac{\mu^{\alpha+1}(1 - \mu)^\beta}{\alpha}$$

For this we use the identity

$$\frac{B_\mu(\alpha + 1, \beta)}{B(\alpha + 1, \beta)} = \frac{B_\mu(\alpha, \beta)}{B(\alpha, \beta)} - \frac{\mu^\alpha(1 - \mu)^\beta}{\alpha B(\alpha, \beta)}$$

which is equivalent to

$$B_\mu(\alpha + 1, \beta) = \mu B_\mu(\alpha, \beta) - \frac{\mu^{\alpha+1}(1 - \mu)^\beta}{\alpha}$$

Thus

$$\mu B_\mu(\alpha, \beta) - B_\mu(\alpha + 1, \beta) = \frac{\mu^{\alpha+1}(1 - \mu)^\beta}{\alpha}$$

and we are done. □

Note 61. Expected Bayes error under a Dirichlet prior

Statement. Use the feature and target types of Note 59, with a symmetric Dirichlet parameter $\alpha > 0$ on every complete type. Put

$$\boldsymbol{\mu}(t) = \boldsymbol{\mu}(t \wedge q) + \boldsymbol{\mu}(t \wedge \neg q), \quad \boldsymbol{\mu}(q \mid t) = \frac{\boldsymbol{\mu}(t \wedge q)}{\boldsymbol{\mu}(t)}.$$

For each t , the variables $\boldsymbol{\mu}(t)$ and $\boldsymbol{\mu}(q \mid t)$ are independent, and $\boldsymbol{\mu}(q \mid t) \sim \text{Beta}(\alpha, \alpha)$. Moreover,

$$\mathbb{E} \left[\sum_{t \in T_\tau} \min\{\boldsymbol{\mu}(t \wedge q), \boldsymbol{\mu}(t \wedge \neg q)\} \right] = \frac{1}{2} - \frac{1}{\alpha B(\alpha, \alpha) 2^{2\alpha}}.$$

Proof. Again, we consider the transformation

$$T(y_1, y_2) = (y_1 y_2, y_1(1 - y_2)),$$

for which the Jacobian was y_1 . Write

$$\alpha_{-t} = (|T_{\tau, q}| - 2)\alpha$$

for the total parameter mass on the types incompatible with t . Since the joint probability distribution of

$$(\boldsymbol{\mu}(t \wedge q), \boldsymbol{\mu}(t \wedge \neg q))$$

is

$$\begin{aligned} f_{\boldsymbol{\mu}(t \wedge q), \boldsymbol{\mu}(t \wedge \neg q)}(x_1, x_2) \\ = \frac{\Gamma(\alpha_{-t} + 2\alpha)}{\Gamma(\alpha)\Gamma(\alpha)\Gamma(\alpha_{-t})} x_1^{\alpha-1} x_2^{\alpha-1} (1 - x_1 - x_2)^{\alpha_{-t}-1}, \end{aligned}$$

the joint probability distribution of $(\boldsymbol{\mu}(t), \boldsymbol{\mu}(q|t))$ is

$$g_{\boldsymbol{\mu}(t), \boldsymbol{\mu}(q|t)}(y_1, y_2) = \frac{\Gamma(\alpha_{-t} + 2\alpha)}{\Gamma(\alpha)\Gamma(\alpha)\Gamma(\alpha_{-t})} h(y_1, y_2),$$

where

$$\begin{aligned} h(y_1, y_2) &= (y_1 y_2)^{\alpha-1} (y_1(1-y_2))^{\alpha-1} (1 - y_1 y_2 - y_1(1-y_2))^{\alpha_{-t}-1} y_1 \\ &= y_1^{2\alpha-1} (1-y_1)^{\alpha_{-t}-1} y_2^{\alpha-1} (1-y_2)^{\alpha-1} \end{aligned}$$

We now factor this joint density. Writing the normalizing constant as

$$\frac{\Gamma(\alpha_{-t} + 2\alpha)}{\Gamma(\alpha)\Gamma(\alpha)\Gamma(\alpha_{-t})} = \frac{\Gamma(\alpha_{-t} + 2\alpha)}{\Gamma(2\alpha)\Gamma(\alpha_{-t})} \cdot \frac{\Gamma(2\alpha)}{\Gamma(\alpha)\Gamma(\alpha)},$$

we obtain

$$\begin{aligned} g_{\boldsymbol{\mu}(t), \boldsymbol{\mu}(q|t)}(y_1, y_2) &= \underbrace{\frac{\Gamma(\alpha_{-t} + 2\alpha)}{\Gamma(2\alpha)\Gamma(\alpha_{-t})} y_1^{2\alpha-1} (1-y_1)^{\alpha_{-t}-1}}_{\text{density of Beta}(2\alpha, \alpha_{-t})} \\ &\quad \cdot \underbrace{\frac{\Gamma(2\alpha)}{\Gamma(\alpha)\Gamma(\alpha)} y_2^{\alpha-1} (1-y_2)^{\alpha-1}}_{\text{density of Beta}(\alpha, \alpha)}. \end{aligned}$$

The joint density thus factors as a function of y_1 alone times a function of y_2 alone, and each factor is itself a probability density. A joint density that factors in this way is exactly the product of its marginals. Hence $\boldsymbol{\mu}(t)$ and $\boldsymbol{\mu}(q|t)$ are independent. Furthermore, by integrating out y_1 , we see that the probability density function of $\boldsymbol{\mu}(q|t)$ is $\text{Beta}(\alpha, \alpha)$.

As in Note 59, we can use the independence of $\boldsymbol{\mu}(t)$ and $\boldsymbol{\mu}(q|t)$ to reduce the calculation of the expectation

$$\mathbb{E} \left(\sum_{t \in T_\tau} \min\{\boldsymbol{\mu}(t \wedge q), \boldsymbol{\mu}(t \wedge \neg q)\} \right)$$

to simply calculating

$$\sum_{t \in T_\tau} \mathbb{E} \boldsymbol{\mu}(t) \mathbb{E} \min\{\boldsymbol{\mu}(q|t), 1 - \boldsymbol{\mu}(q|t)\}$$

Using the probability density function for $\boldsymbol{\mu}(q|t)$ we can compute the expectation of the minimum as follows. First, using the identity

$$\min\{a, b\} = \frac{1}{2}(a + b) - \frac{1}{2}|a - b|$$

we see that

$$\mathbb{E} \min\{\boldsymbol{\mu}(q|t), 1 - \boldsymbol{\mu}(q|t)\} = \frac{1}{2} - \mathbb{E} \left(\left| \boldsymbol{\mu}(q|t) - \frac{1}{2} \right| \right)$$

Since $\mu(q|t) \sim \text{Beta}(\alpha, \alpha)$ has mean $\frac{1}{2}$, Note 60 gives

$$\mathbb{E} \min\{\mu(q|t), 1 - \mu(q|t)\} = \frac{1}{2} - \frac{1}{\alpha B(\alpha, \alpha) 2^{2\alpha}}$$

Thus we obtain that

$$\begin{aligned} & \mathbb{E} \left(\sum_{t \in T_\tau} \min\{\mu(t \wedge q), \mu(t \wedge \neg q)\} \right) \\ &= \left(\frac{1}{2} - \frac{1}{\alpha B(\alpha, \alpha) 2^{2\alpha}} \right) \sum_{t \in T_\tau} \mathbb{E} \mu(t) \\ &= \left(\frac{1}{2} - \frac{1}{\alpha B(\alpha, \alpha) 2^{2\alpha}} \right) \mathbb{E} \underbrace{\sum_{t \in T_\tau} \mu(t)}_{=1} \\ &= \frac{1}{2} - \frac{1}{\alpha B(\alpha, \alpha) 2^{2\alpha}} \end{aligned}$$

Some examples:

1. When $\alpha = 1$, we have that

$$\frac{1}{2} - \frac{1}{B(1, 1) 2^2} = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

2. When $\alpha = 0.1$, we have that

$$\frac{1}{2} - \frac{1}{0.1 B(0.1, 0.1) 2^{0.2}} \approx 0.06$$

3. When $\alpha = 10$, we have that

$$\frac{1}{2} - \frac{1}{10 B(10, 10) 2^{20}} = \frac{1}{2} - \frac{46189}{524288} \approx 0.41$$

For a sanity check on these results, note that the parameter α in the symmetric Dirichlet controls how uniform the randomly picked probability distribution is. If α is very small, then we should expect that the probability distribution we pick is very far from being uniform, while if α is very large, then we should expect the probability distribution to be close to uniform. In the first case there is a type $t \in T_\tau$ such that either $t \wedge q$ or $t \wedge \neg q$ has most of the probability mass, and this type alone will achieve a good accuracy. In the second case we have that for every $t \in T_\tau$ both $t \wedge q$ and $t \wedge \neg q$ have roughly the same amount of probability mass, which means that every formula must have an accuracy that is roughly 50%.

If $\tau = \emptyset$, there is a single feature type t and $\mu(t) = 1$ almost surely. The conditional probability $\mu(q | t)$ still has the $\text{Beta}(\alpha, \alpha)$ distribution, so the final expectation has the same value in this case.

□

Note 77. Expected disagreement under a Dirichlet prior

Statement. Let $t_1, \dots, t_k$ be the Boolean types over a finite vocabulary. Suppose formulas φ, ψ disagree exactly on $t_1, \dots, t_s$, and let $\mu \sim \text{Dir}(\alpha_1, \dots, \alpha_k)$, with all $\alpha_i > 0$. Writing $\alpha_0 = \sum_i \alpha_i$,

$$\mathbb{E}_\mu \left[\Pr_{t \sim \mu}(\varphi(t) \neq \psi(t)) \right] = \frac{\alpha_1 + \dots + \alpha_s}{\alpha_0}.$$

Proof. Clearly it can be represented as

$$\sum_{\ell=1}^s \mu(t_\ell).$$

Thus

$$\begin{aligned} & \mathbb{E}_{\mu \sim \text{Dir}(\alpha_1, \dots, \alpha_k)} \left[\Pr_{t \sim \mu} [\varphi \text{ and } \psi \text{ disagree on } t] \right] \\ &= \sum_{\ell=1}^s \mathbb{E}_{\mu \sim \text{Dir}(\alpha_1, \dots, \alpha_k)} [\mu(t_\ell)] \\ &= \sum_{\ell=1}^s \frac{\alpha_\ell}{\alpha_0} \\ &= \frac{\alpha_1 + \dots + \alpha_s}{\alpha_0} \end{aligned}$$

Here $\alpha_0 := \sum_{\ell=1}^k \alpha_\ell$. In particular, when $\alpha_1 = \dots = \alpha_k = 1$, the expected value is s/k . $\square$

4.3 Representation and convex optimization

Note 79. Margin perceptron and hinge-loss descent

Statement. Consider a nonempty labelled sample $(\mathbf{x}_i, y_i)$, with $y_i \in \{-1, 1\}$ and $\|\mathbf{x}_i\| \leq 1$. Starting at $\mathbf{w} = 0$, margin perceptron updates by $\mathbf{w} \leftarrow \mathbf{w} + y_i \mathbf{x}_i$ when $y_i \langle \mathbf{w}, \mathbf{x}_i \rangle < 1$. This is stochastic subgradient descent with learning rate one for hinge loss. If some $\mathbf{w}^*$ satisfies $y_i \langle \mathbf{w}^*, \mathbf{x}_i \rangle \geq 1$ for all i and $\|\mathbf{w}^*\| \leq 1/\gamma$, the algorithm makes at most $3/\gamma^2$ updates and terminates with normalized margin at least $\gamma/3$.

Proof. (1) We start by arguing why margin perceptron is equivalent to running SGD with learning rate 1 on the class of linear predictors with hinge-loss as the loss function. Recall the definition of hinge-loss:

$$\ell(y, y') = \max\{0, 1 - y \cdot y'\}$$

Suppose that the linear predictor is given by $\mathbf{w}$. If $1 > \mathbf{w} \cdot \mathbf{x}_i \cdot y_i$, then

$$\frac{\partial \ell(\mathbf{w} \cdot \mathbf{x}_i, y_i)}{\partial w_j} = -\frac{\partial (\mathbf{w} \cdot \mathbf{x}_i \cdot y_i)}{\partial w_j} = -x_i^j y_i$$

and hence

$$-\nabla \ell(\mathbf{w} \cdot \mathbf{x}_i, y_i) = \mathbf{x}_i y_i.$$

This establishes our claim.

(2) Suppose that $\|\mathbf{x}_i\| \leq 1$, for every i . Suppose that there exists $\mathbf{w}^*$ such that $(\mathbf{w}^* \cdot \mathbf{x}_i) y_i \geq 1$ for all i and $1/\|\mathbf{w}^*\| \geq \gamma$. Each update increases $\mathbf{w} \cdot \mathbf{w}^*$ by at least one. Furthermore, each update increases $\|\mathbf{w}\|^2$ by at most 3. Hence, if the margin perceptron makes m updates, then $\|\mathbf{w}\|^2 \leq 3m$ and $\mathbf{w} \cdot \mathbf{w}^* \geq m$. A simple sequence of equivalences gives us that $m \leq 3\|\mathbf{w}^*\|^2 \leq 3/\gamma^2$.

(3) Since $\|\mathbf{w}\|^2 \leq 9/\gamma^2$, we get that $1/\|\mathbf{w}\| \geq \gamma/3$. In other words, the margin of the linear predictor is at least $\gamma/3$. $\square$

Note 81. Majority forests versus DNF formulas

Statement. *The strict-majority function on n Boolean variables is represented by a forest of n single-variable decision trees. Every equivalent DNF has at least $\binom{n}{\lfloor n/2 \rfloor + 1}$ terms, and therefore has size $\Omega(2^{n/2})$.*

Proof. Consider the following simple random forest:

$$\mathfrak{F} := \{x_1, \dots, x_n\}.$$

$\mathfrak{F}$ clearly accepts those inputs which contain at least $\lfloor n/2 \rfloor + 1$ ones. Suppose that ϕ is a DNF formula that is equivalent to $\mathfrak{F}$. We claim that ϕ is of size $\Omega(2^{n/2})$. By assumption we have that for every input with precisely $\lfloor n/2 \rfloor + 1$ ones there must exist at least one term in ϕ which accepts it. Thus it suffices to show that there is no term which accepts more than one input with precisely $\lfloor n/2 \rfloor + 1$. Aiming for a contradiction, suppose that t is such a term.

Now, t can contain at most $\lfloor n/2 \rfloor$ positive instances of propositional symbols. But this means that t will accept an input with at most $\lfloor n/2 \rfloor$ ones, which is a contradiction. $\square$

Note 104. An exponential decision-tree lower bound

Statement. *Let all $p_1, q_1, \dots, p_k, q_k$ be distinct variables. Every binary decision tree representing*

$$\bigvee_{i=1}^k (p_i \wedge q_i)$$

has size at least 2^k . Thus this family has linear-size DNF representations but requires exponential-size decision trees.

Proof. For any $I \subseteq \{1, \dots, k\}$ we use $f(I)$ to denote the size of the smallest decision tree which expresses the formula

$$\phi_I := \bigvee_{\ell \in I} (p_\ell \wedge q_\ell).$$

Clearly $f(I) = f(I')$, for every I, I' with $|I| = |I'|$. Hence we can define $f(\ell) := f(\{1, \dots, \ell\})$. We claim that $f(\ell) \geq 2^\ell$.

Trivially we have that $f(1) \geq 2$. Let $\mathcal{T}$ be any decision tree that expresses the formula ϕ_I , where $I = \{1, \dots, \ell\}$. Let $i \in I$ be the index of the proposition symbol that occurs as the label of the topmost node in $\mathcal{T}$ which we assume w.l.o.g. to be p_i . (The case of q_i is completely analogous.) Now the right subtree of $\mathcal{T}$ needs to compute the formula $\phi_{I-\{i\}}$ while the left subtree of $\mathcal{T}$ needs to compute the formula

$$q_i \vee \phi_{I-\{i\}}.$$

The former subtree has by definition size at least $f(I - \{i\})$. We claim that also the latter subtree — which we now denote by $\mathcal{T}'$ — has size at least $f(I - \{i\})$. Indeed, on inputs where q_i is zero it needs to compute the formula $\phi_{I-\{i\}}$.

Hence, if $\mathcal{T}''$ denotes the tree obtained from $\mathcal{T}'$ by removing all nodes labeled with q_i and connecting their right children with their parents, then $\mathcal{T}''$ computes the formula $\phi_{I-\{i\}}$

and hence has size at least $f(I - \{i\})$. Since $\mathcal{T}'$ is larger than $\mathcal{T}''$, also $\mathcal{T}'$ has size at least $f(I - \{i\})$.

Thus $\mathcal{T}$ has size at least

$$2f(I - \{i\}) = 2f(\ell - 1).$$

Since $\mathcal{T}$ was arbitrary, we can deduce that

$$f(\ell) \geq 2f(\ell - 1) \geq 2^\ell.$$

Hence $f(k) \geq 2^k$.

For a complementary comparison between majority forests and DNF formulas, see Note 81. $\square$

Note 105. Local minima of a convex function

Statement. *Every local minimum of a convex function $f : \mathbb{R} \rightarrow \mathbb{R}$ is a global minimum.*

Proof. Let x_0 be a local minimum, and choose $\varepsilon > 0$ such that $f(x_0) \leq f(x)$ whenever $|x - x_0| < \varepsilon$. Aiming for a contradiction, suppose that there exists $x \in \mathbb{R}$ such that $f(x) < f(x_0)$. If

$$0 < \lambda < \min\left\{1, \frac{\varepsilon}{|x - x_0|}\right\},$$

then

$$|(\lambda x + (1 - \lambda)x_0) - x_0| < \varepsilon.$$

Since f was convex and x_0 was a local minimum of f , we get that

$$\begin{aligned} f(x_0) &\leq f(\lambda x + (1 - \lambda)x_0) \\ &\leq \lambda f(x) + (1 - \lambda)f(x_0) \\ &< \lambda f(x_0) + (1 - \lambda)f(x_0) \\ &= f(x_0) \end{aligned}$$

which is a clear contradiction. $\square$

Note 112. Comparing the one-norm and the two-norm

Statement. *For every $x \in \mathbb{R}^n$, the vector norms satisfy $\|x\|_2 \leq \|x\|_1$.*

Proof. It suffices to show that

$$\|x\|_2^2 \leq \|x\|_1^2.$$

Suppose that

$$|x_1| \geq \dots \geq |x_n|.$$

The claim follows now from the following chain of inequalities.

$$\begin{aligned}
\|x\|_2^2 &= \sum_{i=1}^n x_i^2 \\
&= \sum_{i=1}^n |x_i|^2 \\
&\leq |x_1| \sum_{i=1}^n |x_i| \\
&\leq \left(\sum_{i=1}^n |x_i| \right) \left(\sum_{i=1}^n |x_i| \right) \\
&= \|x\|_1^2.
\end{aligned}$$

□

Note 113. The error of retaining the largest coordinates

Statement. Let $x \in \mathbb{R}^n$, and let $H_k(x)$ retain its k largest coordinates in absolute value and replace the others by zero. For $1 \leq k \leq n$,

$$\|x - H_k(x)\|_2 \leq \frac{\|x\|_1}{\sqrt{k}}.$$

Proof. After reindexing, assume $|x_1| \geq \dots \geq |x_n|$, so $H_k(x)$ retains the first k coordinates. It suffices to show that

$$\|x - H_k(x)\|_2^2 \leq \frac{1}{k} \|x\|_1^2.$$

We start as in Note 112 with the following chain of inequalities.

$$\begin{aligned}
\|x - H_k(x)\|_2^2 &= \sum_{i=k+1}^n x_i^2 \\
&= \sum_{i=k+1}^n |x_i|^2 \\
&\leq |x_{k+1}| \sum_{i=k+1}^n |x_i| \\
&\leq |x_{k+1}| \sum_{i=1}^n |x_i|
\end{aligned}$$

It suffices to show that

$$|x_{k+1}| \leq \frac{1}{k} \sum_{i=1}^n |x_i|,$$

which follows from the following chain of inequalities.

$$\begin{aligned}
\sum_{i=1}^n |x_i| &\geq \sum_{i=1}^k |x_i| \\
&\geq k|x_{k+1}|.
\end{aligned}$$

For $k = n$, the claimed bound is immediate because $H_n(x) = x$.

□

Note 114. Quadratic regularization and strong convexity

Statement. Suppose $\ell(\bar{w}, \bar{z})$ is differentiable and convex in $\bar{w}$ on a convex domain, and let $\alpha > 0$. Then

$$r(\bar{w}, \bar{z}) = \ell(\bar{w}, \bar{z}) + \frac{\alpha}{2} \|\bar{w}\|_2^2$$

is α -strongly convex in $\bar{w}$.

Proof. First, because $\ell(\bar{w}, \bar{z})$ is convex, we have that

$$\ell(\bar{w}', \bar{z}) \geq \ell(\bar{w}, \bar{z}) + \langle \nabla \ell(\bar{w}, \bar{z}), \bar{w}' - \bar{w} \rangle$$

and hence it suffices to show that

$$\|\bar{w}'\|_2^2 \geq \|\bar{w}\|_2^2 + \langle \nabla \|\bar{w}\|_2^2, \bar{w}' - \bar{w} \rangle + \|\bar{w} - \bar{w}'\|_2^2.$$

This follows from the fact that $\|\cdot\|_2^2$ is 2-strongly convex.

□

Note 141. The hardness of finding a sparse solution

Statement. For rational A, b , minimizing the number $\|x\|_0$ of nonzero coordinates of a solution to $Ax = b$ is NP-hard. Equivalently, deciding whether a rational linear system has a solution with at most a prescribed number of nonzero coordinates is NP-hard.

Proof. We will give a reduction from 1-in-3-SAT. Let φ be a 3CNF formula over n variables $x_1, \dots, x_n$. Let $\bar{x}_1, \dots, \bar{x}_n$ be fresh variables and consider the following system of linear equations.

$$\begin{aligned} x_1 + \bar{x}_1 &= 1 \\ &\vdots \\ x_n + \bar{x}_n &= 1 \end{aligned}$$

If $y = (x_1, \dots, x_n, \bar{x}_1, \dots, \bar{x}_n) \in \mathbb{Q}^{2n}$ solves it, then for each $1 \leq i \leq n$ either $x_i \neq 0$ or $\bar{x}_i \neq 0$. Therefore $\|y\|_0 \geq n$ for any such solution. Furthermore, if $\|y\|_0 \leq n$, then either $x_i = 1$ or $\bar{x}_i = 1$. Thus any solution y with $\|y\|_0 = n$ encodes a Boolean assignment over the variables $x_1, \dots, x_n$.

Let us return to φ . Write it out as

$$\varphi := \bigwedge_{k=1}^m (\ell_1^k \vee \ell_2^k \vee \ell_3^k).$$

Consider then the following system of linear equations.

$$\begin{aligned}
x_1 + \bar{x}_1 &= 1 \\
&\vdots \\
x_n + \bar{x}_n &= 1 \\
\ell_1^1 + \ell_2^1 + \ell_3^1 &= 1 \\
&\vdots \\
\ell_1^m + \ell_2^m + \ell_3^m &= 1
\end{aligned}$$

We claim that the above system of linear equations has a solution y with $\|y\|_0 \leq n$ iff there exists a Boolean assignment which satisfies precisely one literal in each clause of φ . The right to left direction is trivial. For the converse direction, let y be a solution of this system of linear equations such that $\|y\|_0 \leq n$. Then, as discussed above, y encodes a Boolean assignment over the variables $x_1, \dots, x_n$. It is clear that for this Boolean assignment we have that for each k exactly one of the literals $\ell_1^k, \ell_2^k, \ell_3^k$ is true. □

4.4 Attention and classification objectives

Note 138. Scaled dot-product attention, entry by entry

Statement. Let $Q \in \mathbb{R}^{m \times d_k}$, $K \in \mathbb{R}^{n \times d_k}$, and $V \in \mathbb{R}^{n \times d_v}$, with $d_k > 0$. Let q_j, k_i be rows of Q, K , and apply softmax rowwise. For $\text{Attention}(Q, K, V) = \text{softmax}(QK^\top / \sqrt{d_k})V$,

$$\text{Attention}(Q, K, V)_{jr} = \frac{\sum_{i=1}^n e^{q_j k_i^\top / \sqrt{d_k}} V_{ir}}{\sum_{i=1}^n e^{q_j k_i^\top / \sqrt{d_k}}}.$$

Proof. Let $q_1, \dots, q_m$ be the rows of Q and $k_1, \dots, k_n$ be the rows of K . The j th row in $QK^\top$ is then $[q_j k_1^\top, \dots, q_j k_n^\top]$. Dividing with $\sqrt{d_k}$ and applying softmax, we get that the j th row of

$$\text{softmax}\left(\frac{QK^\top}{\sqrt{d_k}}\right)$$

is

$$\left[\frac{\exp(\text{score}(q_j, k_1^T))}{\sum_i \exp(\text{score}(q_j, k_i^T))}, \dots, \frac{\exp(\text{score}(q_j, k_n^T))}{\sum_i \exp(\text{score}(q_j, k_i^T))} \right]$$

where

$$\text{score}(v, u) := \frac{vu}{\sqrt{d_k}}.$$

Finally, let $v_1, \dots, v_{d_v}$ be the columns of V . Then the (j, k) -cell of $\text{Attention}(Q, K, V)$ is

$$\left[\frac{\exp(\text{score}(q_j, k_1^T))}{\sum_i \exp(\text{score}(q_j, k_i^T))}, \dots, \frac{\exp(\text{score}(q_j, k_n^T))}{\sum_i \exp(\text{score}(q_j, k_i^T))} \right] v_k = \frac{\sum_i \exp(\text{score}(q_j, k_i^T)) v_{i,k}}{\sum_i \exp(\text{score}(q_j, k_i^T))}$$

□

Note 140. Permutation equivariance of self-attention

Statement. Suppose $V[X], K[X], Q[X]$ apply the same transformation to every column of X , and softmax acts independently on each column. Define

$$\text{Sa}[X] = V[X] \text{Softmax}(K[X]^\top Q[X]).$$

For every compatible permutation matrix P , one has $\text{Sa}[XP] = \text{Sa}[X]P$.

Proof. First, we notice that

$$V[XP] = V[X]P, \quad K[XP] = K[X]P, \quad Q[XP] = Q[X]P$$

because in $V[\cdot], K[\cdot]$ and $Q[\cdot]$ the same transformation is applied to each column. Thus

$$\text{Sa}[XP] = V[X]P \cdot \text{Softmax}[P^T K[X]^T Q[X]P].$$

Because **Softmax** is applied column-wise and independently for each column, we get that

$$\text{Softmax}[P^T K[X]^T Q[X]P] = \text{Softmax}[P^T K[X]^T Q[X]]P.$$

Furthermore, since **Softmax** is equivariant under row permutations, we get that

$$\text{Softmax}[P^T K[X]^T Q[X]]P = P^T \text{Softmax}[K[X]^T Q[X]]P$$

To conclude, we note that for a permutation matrix P its transpose P^T is the inverse of P . $\square$

Note 127. The Bayes rule for asymmetric binary loss

Statement. Let (X, Y) have distribution $\mathcal{D}$ on a finite feature space $\mathcal{X}$ and labels $\{0, 1\}$. Let $\ell(\hat{y}, y)$ charge $\alpha > 0$ for a false positive, $\beta > 0$ for a false negative, and zero for a correct prediction. An expected-loss-minimizing classifier is

$$h(x) = 1 \iff \Pr(Y = 1 \mid X = x) > \frac{\alpha}{\alpha + \beta}.$$

Either prediction is optimal at equality, and values at zero-probability feature points are arbitrary.

Proof. Fix a hypothesis $h : \mathcal{X} \rightarrow \{0, 1\}$. We can compute the expected loss as follows.

$$\begin{aligned} & \mathbb{E}_{(X,Y) \sim \mathcal{D}}[\ell(h(X), Y)] \\ &= \sum_{(x,y) \in \mathcal{X} \times \{0,1\}} \ell(h(x), y) \cdot \Pr[X = x, Y = y] \\ &= \sum_{x \in \mathcal{X}} \ell(h(x), 0) \cdot \Pr[X = x, Y = 0] + \sum_{x \in \mathcal{X}} \ell(h(x), 1) \cdot \Pr[X = x, Y = 1] \\ &= \alpha \sum_{x \in h^{-1}(1)} \Pr[X = x, Y = 0] + \beta \sum_{x \in h^{-1}(0)} \Pr[X = x, Y = 1] \\ &= \alpha \Pr[h(X) = 1, Y = 0] + \beta \Pr[h(X) = 0, Y = 1] \end{aligned}$$

A hypothesis h which minimizes this loss minimizes it pointwise. For $x \in \mathcal{X}$ we want to minimize

$$\alpha \Pr[h(x) = 1, Y = 0 | X = x] + \beta \Pr[h(x) = 0, Y = 1 | X = x].$$

If $h(x) = 1$, then the loss is

$$\alpha \Pr[Y = 0 | X = x]$$

and otherwise it is

$$\beta \Pr[Y = 1 | X = x].$$

So for the ideal classifier we have that $h(x) = 1$ iff

$$\begin{aligned} \beta \Pr[Y = 1 | X = x] &> \alpha \Pr[Y = 0 | X = x] \\ \Leftrightarrow \beta \Pr[Y = 1 | X = x] &> \alpha(1 - \Pr[Y = 1 | X = x]) \\ \Leftrightarrow \Pr[Y = 1 | X = x] &> \frac{\alpha}{\alpha + \beta} \end{aligned}$$

□

Note 168. Optimizing confusion-matrix objectives

Statement. Let $\mathfrak{M}$ be a finite unary $\tau \cup \{q\}$ -structure of size n , with τ the Boolean features and q the target. Any polynomial-time-computable objective depending only on the confusion matrix of a deterministic type-based classifier can be maximized in polynomial time. If k feature types are realized and P, N are the positive and negative counts, a dynamic program computes all attainable true-positive and false-positive pairs using $O(k(P + 1)(N + 1))$ elementary updates.

Proof. Let $t_1, \dots, t_k$ enumerate all the (quantifier-free) 1-types over τ realized in $\mathfrak{M}$. Now each $S \subseteq [k]$ determines a classifier f_S defined by

$$f_S(t) = \begin{cases} 1 & \text{if } t \text{ is } t_\ell \text{ for } \ell \in S \\ 0 & \text{otherwise} \end{cases}$$

Let Φ be a polynomial-time computable objective function that only depends on the confusion matrix of f_S . Our goal is to show that Φ can be maximized in polynomial time with respect to n .

Given $S \subseteq [k]$ we define

$$P(S) := \sum_{\ell \in S} |t_\ell \wedge q|_{\mathfrak{M}} \text{ and } N(S) := \sum_{\ell \in S} |t_\ell \wedge \neg q|_{\mathfrak{M}}.$$

Now the confusion matrix of f_S depends only on $P(S)$ and $N(S)$. Thus it suffices to show that we can calculate in polynomial time all pairs

$$(n, m) \in \{0, \dots, |q|_{\mathfrak{M}}\} \times \{0, \dots, |\neg q|_{\mathfrak{M}}\}$$

such that (n, m) is $(P(S), N(S))$, for some $S \subseteq [k]$. Such a pair is called attainable. Note that there are at most $(|q|_{\mathfrak{M}} + 1)(|\neg q|_{\mathfrak{M}} + 1)$ many attainable pairs.

To do this, we employ dynamic programming. Given $0 \leq \ell \leq k$ we write

$$A[\ell, r, s] := \begin{cases} 1 & \text{if } (r, s) \text{ is attainable using } t_1, \dots, t_\ell \\ 0 & \text{otherwise} \end{cases}$$

Initialize $A[0, 0, 0] = 1$, set all other entries with $\ell = 0$ to 0, and interpret entries with out-of-range indices as 0. We have

$$A[\ell, r, s] = \max\{A[\ell - 1, r, s], A[\ell - 1, r - |t_\ell \wedge q|_{\mathfrak{M}}, s - |t_\ell \wedge \neg q|_{\mathfrak{M}}]\}.$$

Storing intermediate results in memory, for each $1 \leq \ell \leq k$ calculating $A[\ell, \cdot, \cdot]$ takes time $\mathcal{O}((|q|_{\mathfrak{M}} + 1)(|\neg q|_{\mathfrak{M}} + 1))$ and so calculating $A[\cdot, \cdot, \cdot]$ takes time $\mathcal{O}(k(|q|_{\mathfrak{M}} + 1)(|\neg q|_{\mathfrak{M}} + 1))$. $\square$

Note 169. Optimizing the minimum of precision and recall

Statement. *In the setting of Note 168, suppose $P > 0$ is the number of positive examples and k is the number of realized feature types. A deterministic type-based classifier maximizing the minimum of precision and recall, with value zero for the empty classifier, can be found in $\mathcal{O}(kn)$ elementary updates after the types have been grouped. Its objective is*

$$\min \left\{ \frac{TP}{TP + FP}, \frac{TP}{P} \right\} = \frac{TP}{\max\{TP + FP, P\}}.$$

Proof. For $0 \leq T \leq n$ and $1 \leq \ell \leq k$ we define

$$A[\ell, T] := \max_{S \subseteq [\ell]} \left\{ \sum_{j \in S} |t_j \wedge q|_{\mathfrak{M}} : \sum_{j \in S} |t_j|_{\mathfrak{M}} = T \right\}$$

Assume $P = |q|_{\mathfrak{M}} > 0$. Initialize $A[0, 0] = 0$, and assign $-\infty$ to impossible or out-of-range states. We have

$$A[\ell, T] := \max\{A[\ell - 1, T], A[\ell - 1, T - |t_\ell|_{\mathfrak{M}}] + |t_\ell \wedge q|_{\mathfrak{M}}\}$$

Storing intermediate results in memory, for each $1 \leq \ell \leq k$ calculating $A[\ell, \cdot]$ takes time $\mathcal{O}(n)$ and so calculating $A[\cdot, \cdot]$ takes time $\mathcal{O}(kn)$. Once we have calculated $A[\cdot, \cdot]$ we can trivially calculate

$$\max_{\substack{0 \leq T \leq n \\ A[k, T] > -\infty}} \frac{A[k, T]}{\max\{T, P\}}$$

and so the total running time is $\mathcal{O}(kn) + \mathcal{O}(n) = \mathcal{O}(kn)$. $\square$

5. Model theory and algebra

5.1 Cardinality, order and infinite games

Note 87. Well-orders inside the real line

Statement. *No uncountable well-ordered set admits an order-preserving embedding into $(\mathbb{R}, <)$.*

Proof. Aiming for a contradiction, suppose that there exists an embedding $i : \alpha \rightarrow \mathbb{R}$, where $\alpha \geq \omega_1$ is an uncountable ordinal. For each $\beta < \omega_1$ we define

$$I_\beta := (i(\beta), i(\beta + 1)) \subseteq \mathbb{R}.$$

Clearly we have that if $\gamma < \beta$, then $I_\beta \cap I_\gamma = \emptyset$. Thus $\{I_\beta \mid \beta < \omega_1\}$ is an uncountable family of open pairwise disjoint intervals. To get a contradiction, we show that every family of open pairwise disjoint intervals must be countable. Let $\mathcal{I}$ be a family of open pairwise disjoint intervals. For each $I \in \mathcal{I}$ there exists a rational number $q_I \in I$.

Since the intervals in $\mathcal{I}$ are pairwise disjoint, the mapping $I \mapsto q_I$ gives an injection from $\mathcal{I}$ to $\mathbb{Q}$, i.e., $\mathcal{I}$ is countable. □

Note 88. Countability of finite Cartesian powers

Statement. *For every positive integer k , the Cartesian power $\mathbb{N}^k$ is countably infinite.*

Proof. Clearly the mapping $n \mapsto (n, 0)$ is an injection from $\mathbb{N}$ to $\mathbb{N}^2$. For the converse direction, we note that the mapping $(n, m) \mapsto 2^n \cdot 3^m$ is an injection from $\mathbb{N}^2$ to $\mathbb{N}$. Hence, the Schröder-Bernstein theorem implies that there is a bijection between $\mathbb{N}$ and $\mathbb{N}^2$. Using induction, we can use this to show that $|\mathbb{N}^k| = |\mathbb{N}|$ for each positive integer k . □

Note 99. A supremum that is a successor ordinal

Statement. *If X is a set of ordinals and $\sup X$ is a successor ordinal, then $\sup X \in X$. Thus X has a largest element.*

Proof. If X has a largest member, then $\sup X$ would be it. Furthermore, X needs to have a largest member, since otherwise $\sup X$ would be a limit ordinal. □

Note 102. Counting structures of an infinite cardinality

Statement. For a countable first-order language L and an infinite cardinal κ , there are at most 2^κ isomorphism types of L -structures of cardinality κ .

Proof. Since we are counting structures only up to isomorphism, we can assume that their domain is κ . Now there are at most

$$\prod_{R \in L} 2^{\kappa^{ar(R)}} = \prod_{R \in L} 2^\kappa = 2^{\kappa \cdot |L|} = 2^\kappa$$

many L -structures over κ , which is what we wanted to show. □

Note 106. Cancellation of ordinal successors

Statement. For ordinals α, β , the inequality $\alpha + 1 \leq \beta + 1$ implies $\alpha \leq \beta$.

Proof. Aiming for a contradiction, suppose that $\alpha > \beta$. Then $\alpha \geq \beta + 1$ and hence $\alpha + 1 > \beta + 1$, a contradiction. □

Note 153. Pigeonhole principles for infinite cardinals

Statement. The cardinal $\aleph_\omega$ can be partitioned into countably many sets, each of cardinality less than $\aleph_\omega$. In contrast, for an infinite cardinal κ , a union of at most κ sets of size at most κ has size at most κ . Consequently, a set of size κ^+ cannot be partitioned into fewer than κ^+ sets all of size less than κ^+ .

Proof. For example, we have

$$\aleph_\omega := \bigcup_{n < \omega} \left(\aleph_n \setminus \bigcup_{m < n} \aleph_m \right).$$

Here

$$\left| \aleph_n \setminus \bigcup_{m < n} \aleph_m \right| \leq |\aleph_n| = \aleph_n < \aleph_\omega,$$

and so $\aleph_\omega$ can be partitioned into $< \aleph_\omega$ many pieces each having size $< \aleph_\omega$. On the other hand, pigeonhole principle holds for successor cardinals. Indeed, suppose that

$$X = \bigcup_{i \in I} X_i,$$

where $|X| = \kappa^+$, $|X_i| < \kappa^+$ and $|I| < \kappa^+$. Thus $|X_i| \leq \kappa$ and $|I| \leq \kappa$ and hence

$$\left| \bigcup_{i \in I} X_i \right| \leq \kappa < |X|,$$

a contradiction. □

Note 159. Winning decimal games and perfect subsets

Statement. Let $A \subseteq (0, 1)$. In an infinite game, player $\forall$ chooses the decimal digits in positions $0, 2, 4, \dots$, and player $\exists$ chooses those in positions $1, 3, 5, \dots$. Player $\exists$ wins when the resulting real $0.d_0d_1d_2\dots$ belongs to A . If $\exists$ has a winning strategy, then A contains a nonempty perfect subset and hence has cardinality $2^{\aleph_0}$.

Proof. Fix a winning strategy σ for player $\exists$. Consider the set

$$B := \{0.d_0d_1d_2\dots \mid \text{the play } (d_0, d_1, d_2, \dots) \text{ is consistent with } \sigma\}.$$

When we say that $(d_0, d_1, d_2, \dots)$ is consistent with σ , we mean that for each i we have that $d_{2i+1} = \sigma((d_0, d_1, \dots, d_{2i}))$. Since σ is a winning strategy for $\exists$, $B \subseteq A$. Also clearly B is non-empty. To show that it is perfect, let

$$0.d_0d_1d_2\dots \in B$$

and $n \in \mathbb{N}$. We have two cases. If $n = 2i$, then let $c_n \in \{0, \dots, 9\} \setminus \{d_n\}$ such that $|c_n - d_n| \geq 2$.¹ Now, since σ is a winning strategy, there is a play of the form

$$(d_0, \dots, d_{n-1}, c_n, c_{n+1}, \dots),$$

which is consistent with σ . Furthermore,

$$0.d_0d_1d_2\dots \neq 0.d_0d_1d_2\dots d_{n-1}c_nc_{n+1}\dots$$

and $0.d_0d_1d_2\dots d_{n-1}c_nc_{n+1}\dots \in B$. If $n = 2i + 1$, then we repeat the above argument for index $2(i + 1)$.

Finally, the set of infinite digit sequences consistent with σ is closed in the compact product space $\{0, \dots, 9\}^{\mathbb{N}}$. Its image under the continuous decimal-expansion map is B , so B is compact and hence closed. Together with the absence of isolated points proved above, this establishes that B is perfect in the usual topological sense. □

5.2 Compactness and elementary equivalence**Note 83. Model completeness and universal formulas**

Statement. For a first-order theory T , the following conditions are equivalent:

- (i) T is model complete.
- (ii) Whenever $M \subseteq N$ are models of T , one has $M \preceq N$.
- (iii) Every formula is equivalent modulo T to a universal formula, with the same free variables.

Proof. 1. We start by showing that T is model complete iff for every two models $\mathcal{M}_1, \mathcal{M}_2$ of T we have that if $\mathcal{M}_1 \subseteq \mathcal{M}_2$ then $\mathcal{M}_1 \preceq \mathcal{M}_2$. Suppose first that T is model complete. Let $\mathcal{M}_1 \subseteq \mathcal{M}_2$ be models of T . We want to show that $\mathcal{M}_1 \preceq \mathcal{M}_2$. We use the Tarski-Vaught test. Let $\varphi(\bar{x}, y)$ be a formula and $\bar{a}$ a tuple of elements of $\mathcal{M}_1$ such that $\mathcal{M}_2 \models \exists y \varphi(\bar{a}, y)$.

¹Why $|c_n - d_n| \geq 2$? Consider $0.100\dots$ and $0.099\dots$. These are the same real number, even though they have different digit representations. By setting $|c_n - d_n| \geq 2$, we can guarantee that the decimals below c_n will not force c_n to round up.

Aiming for a contradiction, suppose that $\mathcal{M}_1 \models \forall y \neg \varphi(\bar{a}, y)$. Since $T \cup \text{Diag}(\mathcal{M}_1)$ is complete, we have that $T \cup \text{Diag}(\mathcal{M}_1) \vdash \forall y \neg \varphi(\bar{a}, y)$. Since

$$T \cup \text{Diag}(\mathcal{M}_1) \subseteq T \cup \text{Diag}(\mathcal{M}_2),$$

we have that $T \cup \text{Diag}(\mathcal{M}_2) \vdash \forall y \neg \varphi(\bar{a}, y)$. Hence $\mathcal{M}_2 \models \forall y \neg \varphi(\bar{a}, y)$, a contradiction.

Next we consider the right to left direction. Let $\mathcal{M}$ be a model of T . We want to show that $T \cup \text{Diag}(\mathcal{M})$ is complete. Let φ be a sentence in the vocabulary L_M . Let $\mathcal{M}^+$ be the extension of $\mathcal{M}$ to the vocabulary L_M . Now, either $\mathcal{M}^+ \models \varphi$ or $\mathcal{M}^+ \models \neg \varphi$. Without loss of generality, we assume that the first case is true. We claim that $T \cup \text{Diag}(\mathcal{M}) \vdash \varphi$. Let $\mathcal{N}^+$ be a model of $T \cup \text{Diag}(\mathcal{M})$ in the vocabulary L_M .

Let $\mathcal{N}$ be the reduct of $\mathcal{N}^+$ to the vocabulary L . Now $\mathcal{M} \subseteq \mathcal{N}$ and hence by assumption $\mathcal{M} \preceq \mathcal{N}$. Let $\varphi(\bar{x})$ be the formula obtained from φ by replacing each constant symbol corresponding to an element of M with a variable. Since $\mathcal{M}^+ \models \varphi$, there exists a tuple $\bar{a}$ such that $\mathcal{M} \models \varphi(\bar{a})$. Since $\mathcal{M} \preceq \mathcal{N}$, we have that $\mathcal{N} \models \varphi(\bar{a})$. Now $\mathcal{N}^+ \models \varphi$. Thus $T \cup \text{Diag}(\mathcal{M}) \vdash \varphi$.

2. We show that if T eliminates quantifiers down to universal quantifiers, then it is model complete. We use the first part here. Let $\mathcal{M}$ and $\mathcal{N}$ be models of T such that $\mathcal{M} \subseteq \mathcal{N}$. We show that $\mathcal{M} \preceq \mathcal{N}$. For this we use the Tarski-Vaught test. Let $\varphi(\bar{x}, y)$ be an L -formula such that $\mathcal{N} \models \exists y \varphi(\bar{a}, y)$. Now, there exists an universal formula $\theta(\bar{x})$ such that $T \vdash \forall \bar{x} (\exists y \varphi(\bar{x}, y) \leftrightarrow \theta(\bar{x}))$. Since universal formulas are preserved under substructures, we have that $\mathcal{M} \models \theta(\bar{a})$. Thus $\mathcal{M} \models \exists y \varphi(\bar{a}, y)$, which is what we wanted to show.

3. We show that if T is model complete, then it eliminates quantifiers down to universal quantifiers. Let $\varphi(\bar{x})$ be an L -formula, $\bar{c}$ be new constant symbols and $L' = L \cup \{\bar{c}\}$. Let Σ be the set of all universal L' -sentences ψ such that $T \vdash \varphi(\bar{c}) \rightarrow \psi$.

We claim that $T \cup \Sigma \vdash \varphi(\bar{c})$. Suppose not. Let $\mathcal{M}$ be a model of $T \cup \Sigma \cup \{\neg \varphi(\bar{c})\}$. Clearly we can assume that $L' \subseteq L_M$. Since $T \cup \text{Diag}(\mathcal{M})$ is a complete theory, we have that $T \cup \text{Diag}(\mathcal{M}) \vdash \neg \varphi(\bar{c})$. By compactness, there exists quantifier-free sentences $\psi_1, \dots, \psi_r \in \text{Diag}(\mathcal{M})$ such that $T \cup \{\psi_1, \dots, \psi_r\} \vdash \neg \varphi(\bar{c})$. Let ψ be the conjunction of the sentences $\psi_1, \dots, \psi_r$. Now ψ is of the form $\theta(\bar{d})$, where $\theta(\bar{x})$ is an L' -formula and $\bar{d} \subseteq L_M \setminus L'$. Since T is an L -theory, we have that $T \vdash \exists \bar{x} \theta(\bar{x}) \rightarrow \neg \varphi(\bar{c})$.

Thus $T \vdash \varphi(\bar{c}) \rightarrow \forall \bar{x} \neg \theta(\bar{x})$. Hence $\forall \bar{x} \neg \theta(\bar{x}) \in \Sigma$ and we get that $\mathcal{M} \models \forall \bar{x} \neg \theta(\bar{x})$, which is a contradiction, since $\mathcal{M} \models \theta(\bar{d})$.

Thus $T \cup \Sigma \vdash \varphi(\bar{c})$. By compactness, there exists universal sentences $\psi_1, \dots, \psi_r \in \Sigma$ such that $T \cup \{\psi_1, \dots, \psi_r\} \vdash \varphi(\bar{c})$. Let ψ be the conjunction of the sentences $\psi_1, \dots, \psi_r$. Now ψ is of the form $\theta(\bar{c})$, where $\theta(\bar{x})$ is an L -formula. Since T is an L -theory, we have that $T \vdash \forall \bar{x} (\theta(\bar{x}) \leftrightarrow \varphi(\bar{x}))$. Since $\theta(\bar{x})$ is an universal formula, we are done.

4. Finally, we show that if T eliminates quantifiers down to universal quantifiers, then it also eliminates quantifiers down to existential quantifiers. We note that the converse direction is also true and equally easy to show.

Let $\varphi(\bar{x})$ be an L -formula. Now, there exists an existential formula $\theta(\bar{x})$ such that $T \vdash \forall \bar{x} (\neg \varphi(\bar{x}) \leftrightarrow \theta(\bar{x}))$. Thus $T \vdash \forall \bar{x} (\varphi(\bar{x}) \leftrightarrow \neg \theta(\bar{x}))$. The claim follows, since $\neg \theta(\bar{x})$ is equivalent to a universal formula.

□

Note 89. Obstructions to universal second-order definitions

Statement. *Neither of the following infinite structures is definable up to isomorphism by a universal second-order sentence:*

- (i) *the unique model of cardinality κ of a countable κ -categorical first-order theory, where κ is infinite;*
- (ii) *a countably infinite recursively saturated structure in a recursive vocabulary.*

Proof. For (i), let $\mathfrak{A}$ be the unique model of T of cardinality κ . Aiming for a contradiction, suppose that φ defines $\mathfrak{A}$ up to isomorphism. Now $\neg\varphi$ is equivalent to an ESO-sentence which is true in a structure $\mathfrak{B}$ iff $\mathfrak{A} \not\cong \mathfrak{B}$. Since T is an FO-theory with an infinite model, it has a model of cardinality $\lambda > \kappa$.

This is also an infinite model of $T \cup \{\neg\varphi\}$. Hence there exists a model $\mathfrak{B}$ of size κ such that $\mathfrak{B} \models T \cup \{\neg\varphi\}$. Since T was κ -categorical, $\mathfrak{B} \cong \mathfrak{A}$. This is a contradiction, because $\mathfrak{B} \models \neg\varphi$.

For (ii), let $\mathfrak{A}$ be a countably infinite recursively saturated structure. We claim that $\mathfrak{A}$ is not USO-definable. Aiming for a contradiction, suppose that φ defines $\mathfrak{A}$ up to isomorphism. We use the following fact: there exists a countable set $\{\varphi_n \mid n \in \mathbb{N}\}$ of FO-sentences such that

1. for every structure $\mathfrak{B}$ we have that

$$\text{if } \mathfrak{B} \models \bigvee_{n \in \mathbb{N}} \varphi_n, \text{ then } \mathfrak{B} \models \varphi;$$

2. and for every countable recursively saturated structure $\mathfrak{B}$ we have that

$$\text{if } \mathfrak{B} \models \varphi, \text{ then } \mathfrak{B} \models \bigvee_{n \in \mathbb{N}} \varphi_n.$$

Now $\mathfrak{A} \models \bigvee_{n \in \mathbb{N}} \varphi_n$. Let $\mathfrak{B}$ be an uncountable elementary extension of $\mathfrak{A}$. Then $\mathfrak{B} \models \bigvee_{n \in \mathbb{N}} \varphi_n$ and hence $\mathfrak{B} \models \varphi$, a contradiction. □

Note 90. First-order Craig interpolation

Statement. *First-order logic has Craig interpolation: if $\varphi \in \text{FO}[\tau]$ and $\psi \in \text{FO}[\sigma]$ are sentences with $\varphi \models \psi$, then some sentence $\theta \in \text{FO}[\tau \cap \sigma]$ satisfies*

$$\varphi \models \theta \models \psi.$$

Proof. Let $R_1, \dots, R_n$ be the relation symbols occurring in φ but not in ψ , and put

$$\Phi := \exists R_1 \dots \exists R_n \varphi.$$

Then $\Phi \models \psi$. Using the first-order approximation property underlying Note 89, choose sentences $\varphi_i \in \text{FO}[\tau \cap \sigma]$ such that $\Phi \models \varphi_i$ for every i , and such that, on countable recursively saturated structures, $\bigwedge_i \varphi_i$ implies Φ .

We claim that $\{\varphi_i : i \in \mathbb{N}\} \models \psi$. Otherwise this first-order theory together with $\neg\psi$ has a countable recursively saturated model. That model satisfies $\Phi \wedge \neg\psi$, a contradiction.

By compactness there are $i_1, \dots, i_m$ with $\varphi_{i_1} \wedge \dots \wedge \varphi_{i_m} \models \psi$. Their conjunction is in the common vocabulary and is implied by Φ , hence by φ . It is the required interpolant. □

Note 91. Oligomorphic groups and quantifier elimination

Statement. *Let G be an oligomorphic permutation group acting on a countably infinite set A . Form a relational structure $\mathfrak{A}$ by naming every G -orbit on A^n , for each $n \geq 1$. Then $\text{Th}(\mathfrak{A})$ eliminates quantifiers and is ω -categorical.*

Proof. We first show that $\mathfrak{A}$ has the quantifier-elimination property. First, observe that every relation over $\mathfrak{A}$ definable via a quantifier-free formula is a finite union of orbits. Indeed, this follows from the fact that

$$\neg R_O(\bar{x}) = \bigvee_{O' \neq O^*} R_{O'}(\bar{x}),$$

where the disjunction is finite, because G was oligomorphic. Since every finite union of orbits is trivially definable via a quantifier-free formula, finite unions of orbits are precisely the relations definable over $\mathfrak{A}$ via a quantifier-free formula.

Now, if we have a finite union of orbits, then their projection is still a finite union of orbits. Indeed, consider a single orbit $O \subseteq A^n$. Let $p(O) \subseteq A^{n-1}$ denote its projection where we remove, say, the last element from each tuple. We claim that

$$p(O) = \bigcup \{O' \mid \exists (\bar{a}, c) \in A^n : \bar{a} \in O' \text{ and } (\bar{a}, c) \in O\}.$$

The direction " $\Rightarrow$ " is clear. Suppose then that $\bar{a} \in O'$. Thus there exists $(\bar{b}, c) \in A^n$ for which $\bar{b} \in O'$ and $(\bar{b}, c) \in O$. Hence there exists $g \in G$ such that $g(\bar{b}) = \bar{a}$. Thus $g((\bar{b}, c)) = (\bar{a}, g(c))$, i.e., $\bar{a} \in p(O)$. Since projections of finite unions of orbits are still definable via a quantifier-free formula, the structure $\mathfrak{A}$ has quantifier-elimination.

Since $\mathfrak{A}$ has quantifier-elimination, so does $\text{Th}(\mathfrak{A})$. This implies that every type in $S_n(\text{Th}(\mathfrak{A}))$ is principal for each $n \in \mathbb{Z}_+$. From Ryll-Nardzewski we get that $\text{Th}(\mathfrak{A})$ is ω -categorical, i.e., $\mathfrak{A}$ is ω -categorical. □

Note 94. Local-to-global existence by compactness

Statement. *Let $\mathfrak{A}$ be a structure and let*

$$\varphi = \exists X_1 \cdots \exists X_k \forall x_1 \cdots \forall x_n \psi,$$

where ψ is quantifier-free and the X_i are relation variables. If every finitely generated substructure of $\mathfrak{A}$ satisfies φ , then $\mathfrak{A} \models \varphi$.

Proof. Consider the atomic diagram $\text{Diag}(\mathfrak{A})$ of $\mathfrak{A}$ and the ESO theory $\text{Diag}(\mathfrak{A}) \cup \{\varphi\}$. We first show that the above theory is consistent. Since ESO is compact, it suffices to show that it is finitely consistent. Let Σ be a finite subset of $\text{Diag}(\mathfrak{A})$. Now only finitely many elements from A occur in Σ , say $a_1, \dots, a_s \in A$. Since every finitely generated substructure of $\mathfrak{A}$ satisfies φ , the τ -structure $\langle a_1, \dots, a_s \rangle_{\mathfrak{A}}$ is a model of $\Sigma \cup \{\varphi\}$.

Let $\mathfrak{B}$ be a model of $\text{Diag}(\mathfrak{A}) \cup \{\varphi\}$. Since $\mathfrak{B} \models \text{Diag}(\mathfrak{A})$ w.l.o.g. we can assume that $\mathfrak{A}$ is a substructure of $\mathfrak{B}$. Let $\mathfrak{B}^+$ be an expansion of $\mathfrak{B}$ such that $\mathfrak{B}^+ \models \forall x_1 \cdots \forall x_n \psi$. Now $\mathfrak{B}^+ \upharpoonright \text{dom}(\mathfrak{A}) \models \forall x_1 \cdots \forall x_n \psi$, because universal sentences are preserved under substructures. The reduct of $\mathfrak{B}^+ \upharpoonright \text{dom}(\mathfrak{A})$ to the vocabulary of φ is $\mathfrak{A}$ and therefore $\mathfrak{A}$ is a model of φ .

Some consequences of the above result.

- Let G be a graph. If every finite subgraph of G is k -colorable, then G is k -colorable.
- Let G be a group. If every finitely generated subgroup of G has a total ordering, then so does G .

Indeed, the above result is quite general.

□

Note 151. Elementary equivalence without partial isomorphism

Statement. In the vocabulary $\{P_n : n \in \mathbb{N}\}$, define $\mathfrak{A}$ on $\mathbb{N}$ by $P_n^{\mathfrak{A}} = \{m : m \geq n\}$, and define $\mathfrak{B}$ on $\mathbb{N} \cup \{\infty\}$ by $P_n^{\mathfrak{B}} = P_n^{\mathfrak{A}} \cup \{\infty\}$. Then $\mathfrak{A} \equiv \mathfrak{B}$, but $\mathfrak{A}$ and $\mathfrak{B}$ are not partially isomorphic.

Proof. Notice that

$$\mathfrak{A} \not\equiv_p \mathfrak{B},$$

because Spoiler can pick ∞ from $\mathfrak{B}$ and Duplicator has no non-losing choices because there is no element in $\mathfrak{A}$ that belongs to the interpretation of all the relations $P_n^{\mathfrak{A}}$. On the other hand, we have

$$\mathfrak{A} \equiv \mathfrak{B}.$$

To prove this, it suffices to show that

$$\forall \text{ finite } \tau_0 \subseteq \tau : \mathfrak{A} \upharpoonright \tau_0 \cong \mathfrak{B} \upharpoonright \tau_0.$$

So, pick a finite $\tau_0 \subseteq \tau$. It suffices to show that

$$\mathfrak{A} \upharpoonright \tau_0 \cong \mathfrak{B} \upharpoonright \tau_0.$$

This follows immediately from the simple observation that both structures realize each quantifier-free 1-type the same number of times.

□

Note 152. Continuum many countable models of true arithmetic

Statement. The complete first-order theory of $(\mathbb{N}, 0, 1, +, \cdot)$ has exactly $2^{\aleph_0}$ non-isomorphic countable models.

Proof. We start by observing that clearly $\text{Th}(\mathbb{N})$ has at most 2^ω many countable models, because the underlying vocabulary is finite. Next, as an auxiliary result, we prove that for each $Q \subseteq \mathbb{P} := \{\text{prime numbers}\}$ there exists a countable model of $\text{Th}(\mathbb{N})$ which has an element whose prime divisors are exactly the prime numbers in Q . We call a set of the form

$$\pi_Q(x) := \{ \underbrace{\exists y(y \cdot (1 + \cdots + 1) = x)}_{p\text{-times}} \mid p \in Q \} \cup \{ \underbrace{\neg \exists y(y \cdot (1 + \cdots + 1) = x)}_{p\text{-times}} \mid p \notin Q \},$$

where $Q \subseteq \mathbb{P}$, a divisor-type. Note that $\pi_Q(x)$ simply states that the set of prime divisors of x is exactly Q . Let $\tau := \{0, 1, +, \cdot\} \cup \{c\}$ and consider the following τ -theory:

$$\text{Th}(\mathbb{N}) \cup \pi_Q(c).$$

By the Compactness Theorem and the Löwenheim-Skolem Theorem, this theory has a countable model and the reduct of this model to the vocabulary $\{0, 1, +, \cdot\}$ is the desired model. To show that $\text{Th}(\mathbb{N})$ must have at least 2^ω many non-isomorphic countable models, we assume — aiming for a contradiction — that it has at most $\kappa < 2^\omega$ many non-isomorphic countable models.

Since a countable model of $\text{Th}(\mathbb{N})$ realizes at most countably many divisor-types and different divisor-types are mutually exclusive, the models of $\text{Th}(\mathbb{N})$ realize at most $\max\{\kappa, \omega\} < 2^\omega$ many divisor-types. Since each divisor-type is realized by a countable model of $\text{Th}(\mathbb{N})$, there are $< 2^\omega$ many divisor types. But this is a contradiction, because there are 2^ω many divisor-types. □

Note 165. A saturated model at an inaccessible cardinal

Statement. *Let κ be strongly inaccessible and let T be a theory in a language L with $|L| < \kappa$. If T has a model of cardinality κ , then it has a saturated model of cardinality κ , realizing every type over each parameter set of size less than κ .*

Proof. Let T be an L -theory such that T has an infinite model of cardinality κ , where $\kappa > |L|$ and it is strongly inaccessible. Given a structure $\mathfrak{A}$ we write $\mathfrak{A}^+$ for the expansion of $\mathfrak{A}$ obtained by naming each element $a \in A$ with a constant c_a . Furthermore, given $S \subseteq A$ we write $\mathfrak{A}^S$ for the expansion of $\mathfrak{A}$ obtained by naming each element $a \in S$ with a constant c_a .

Let $\mathfrak{A}$ be any model of T of cardinality κ . Now, for each $S \subseteq A$ such that $|S| < \kappa$ we write P^S for the set of all n -types consistent with $\text{Th}(\mathfrak{A}^S)$. Each n -type is a subset of the set of all formulas over the vocabulary $L \cup \{c_a \mid a \in S\}$. There are at most $\max\{\aleph_0, |L| + |S|\}$ many such formulas, and therefore there are at most

$$2^{\max\{\aleph_0, |L| + |S|\}}$$

many such n -types. Since $|L|, |S| < \kappa$ and κ is strongly inaccessible, we have

$$2^{\max\{\aleph_0, |L| + |S|\}} < \kappa.$$

The number of subsets of A of size $< \kappa$ is $\kappa^{<\kappa} = \kappa$, because κ is strongly inaccessible. Thus

$$P := \bigcup \{P^S \mid S \subseteq A \text{ and } |S| < \kappa\}$$

has size at most κ . Now there exists an elementary extension $\mathfrak{B}^+$ of $\mathfrak{A}^+$ of size κ which realizes every type in P . Let $\mathfrak{B}$ be the L -reduct of $\mathfrak{B}^+$. Now $\mathfrak{B}$ is an elementary extension of $\mathfrak{A}$.

By repeating the above procedure and taking unions at limit ordinals, we get an increasing sequence $(\mathfrak{A}_\lambda)_{\lambda < \kappa}$ of L -structures such that for each $\delta < \lambda < \kappa$ we have that (i) $\mathfrak{A}_\lambda$ has cardinality κ (ii) $\mathfrak{A}_\delta$ is an elementary substructure of $\mathfrak{A}_\lambda$ (iii) for every $S \subseteq A_\delta$ such that $|S| < \kappa$ we have that every n -type consistent with $\text{Th}(\mathfrak{A}_\delta^S)$ is realized in $\mathfrak{A}_\lambda$.

Let $\mathfrak{M}$ denote $\bigcup_{\lambda < \kappa} \mathfrak{A}_\lambda$. Now $\mathfrak{M}$ contains each $\mathfrak{A}_\lambda$ as an elementary substructure and it has cardinality κ . We claim that $\mathfrak{M}$ is saturated. Let $S \subseteq M$ such that $|S| < \kappa$ and let p be

an n -type consistent with $\text{Th}(\mathfrak{M}^S)$. For each $a \in S$ we pick λ_a such that a is an element of A_{λ_a} . Now

$$\lambda := \sup_{a \in S} \lambda_a < \kappa.$$

Since $(\mathfrak{A}_\lambda)_{\lambda < \kappa}$ is an increasing sequence, we have that $\mathfrak{A}_{\lambda_a} \subseteq \mathfrak{A}_\lambda$, for every $a \in S$, and in particular $S \subseteq A_\lambda$. Now p is consistent with $\text{Th}(\mathfrak{A}_\lambda^S)$, because $\mathfrak{A}_\lambda^S$ is an elementary substructure of $\mathfrak{M}^S$. Hence p is realized in $\mathfrak{A}_{\lambda+1}^S$ and therefore in $\mathfrak{M}^S$. $\square$

5.3 Fields and definability

Note 28. Infinitesimals in elementary extensions

Statement. *Every proper elementary extension of the ordered field $\mathbb{R}$ contains a nonzero infinitesimal. Here $\varepsilon \neq 0$ is infinitesimal when $|\varepsilon| < 1/n$ for every positive integer n .*

Proof. Let $a \in \text{dom}(\mathcal{R}) \setminus \mathbb{R}$. Without loss of generality, we assume that $a > 0$, since otherwise can replace a with $-a$. If a is infinite, then $1/a$ is a nonzero infinitesimal. Thus, assume that a is finite. Consider the sets

$$A := \{r \in \mathbb{R} \mid r < a\}, B := \{r \in \mathbb{R} \mid a < r\} \subseteq \mathbb{R}$$

Since $\mathbb{R}$ is complete, both $\sup A = b_1$ and $\inf B = b_2$ exist. We claim that $b_1 = b_2$. First, note that every member of B is larger than every member of A , which means that b_2 is an upper bound on A , i.e., $b_1 \leq b_2$. Aiming for a contradiction, suppose that there exists $n \in \mathbb{N}$ such that $b_2 - b_1 > \frac{1}{n}$, i.e., $b_2 - 1/n > b_1$. Then either $b_2 - 1/n < a$ or $b_2 - 1/n > a$. In the first case we have that $b_2 - 1/n \in A$, which would entail that $b_2 - 1/n \leq b_1$, a contradiction.

In the second case we have that $b_2 - 1/n \in B$, which would entail that $b_2 \leq b_2 - 1/n$, a contradiction. Thus $b_1 = b_2$.

Let $b = \sup A = \inf B$. We have now two cases.

1. Suppose that $b < a$. We claim that $a - b > 0$ is infinitesimal. Let $n \in \mathbb{N}$. We claim that $a - b < 1/n$. Aiming for a contradiction, suppose that $a \geq b + 1/n$. Since $a \notin \mathbb{R}$, we must have that $a > b + 1/n$, which means that $b + 1/n \in A$. But since $b = \sup A$, we have that $b + 1/n \leq b$, a contradiction. Hence $a - b < 1/n$, and the claim follows since n was arbitrary.
2. Suppose that $b > a$. We claim that $b - a$ is infinitesimal. Let $n \in \mathbb{N}$. We claim that $b - a < 1/n$. Aiming for a contradiction, suppose that $a \leq b - 1/n$. Since $a \notin \mathbb{R}$, we must have that $a < b - 1/n$, which means that $b - 1/n \in B$. But since $b = \inf B$, we have that $b \leq b - 1/n$, a contradiction. Hence $b - a < 1/n$, and the claim follows since n was arbitrary.

Hence $\mathcal{R}$ must contain an infinitesimal. $\square$

Note 82. Injective polynomial maps over closed fields

Statement. *For an algebraically closed field K and $n \geq 1$, every injective polynomial map $f : K^n \rightarrow K^n$ is surjective. The converse fails in general, and the implication can fail when K is not algebraically closed.*

Proof. Fix n, d , and let $\Psi(K, n, d)$ assert that every injective polynomial self-map of K^n whose coordinate polynomials have degree at most d is surjective. First, observe that $\Psi(K, n, d)$ is true when K is finite. Indeed, in this case K^n is also finite, which means that every injection from K^n to itself is surjective. Now, consider $\mathbb{F}_p^{\text{alg}}$, for prime p . We claim that $\Psi(K, n, d)$ is true for $\mathbb{F}_p^{\text{alg}}$. Let $f : (\mathbb{F}_p^{\text{alg}})^n \rightarrow (\mathbb{F}_p^{\text{alg}})^n$ be an injective polynomial mapping. Let $\mathbb{F}_q$ be a finite subfield of $\mathbb{F}_p^{\text{alg}}$ such that all the coefficients of f are from $\mathbb{F}_q$.

If $K \supseteq \mathbb{F}_q$ is a finite subfield of $\mathbb{F}_p^{\text{alg}}$, we know that $f \upharpoonright K^n$ is an injection from K^n to itself, because fields are closed under polynomials. But since $\Psi(K, n, d)$ holds for every finite field K which contains $\mathbb{F}_q$, we know that $f \upharpoonright K^n$ is a surjection. Since $\mathbb{F}_p^{\text{alg}}$ is a union of its finite subfields, we can conclude that f is also a surjection.

Since for each prime p there exists an algebraically closed field of characteristic p for which the statement $\Psi(K, n, d)$ is true, by completeness of ACF_p we know that $\Psi(K, n, d)$ follows from ACF_p . Furthermore, by compactness we get that $\Psi(K, n, d)$ follows also from ACF_0 . Hence $\Psi(K, n, d)$ is true for all algebraically closed fields.

Note that the mapping $f : \mathbb{C} \rightarrow \mathbb{C}$, where $f(x) = x^2$ is surjective but not injective. Indeed, it is surjective because $\mathbb{C}$ is algebraically closed, but it is not injective because e.g. $1^2 = (-1)^2$. More generally, if $f : \mathbb{C} \rightarrow \mathbb{C}$ is any non-constant polynomial, then it is a surjection, but of course not necessarily an injection.

We also note that $\Psi(K, n, d)$ can fail when K is not algebraically closed. Consider for instance the case $K = \mathbb{Q}$, $n = 1$ and $d = 3$. The polynomial X^3 is clearly an injection $\mathbb{Q} \rightarrow \mathbb{Q}$. However, it is not a surjection, since e.g. the equation $X^3 = 2$ has no solutions over $\mathbb{Q}$. □

Note 85. Characteristic zero in existential monadic logic

Statement. *Among commutative rings with identity, characteristic zero is definable by an existential monadic second-order sentence. One such sentence is*

$$\exists X (X(1) \wedge \neg X(0) \wedge \forall x (X(x) \rightarrow X(x+1))).$$

Proof. Consider the following sentence of EMSO.

$$\varphi := \exists X (X(1) \wedge \neg X(0) \wedge \forall x (X(x) \rightarrow X(x+1)))$$

If $(R, +, \cdot, 0, 1)$ has characteristic zero, then we can clearly take

$$X = \{1, 1+1, 1+1+1, \dots\},$$

whence $(R, +, \cdot, 0, 1) \models \varphi$. If $(R, +, \cdot, 0, 1) \models \varphi$, then the ring must have characteristic zero, because otherwise 0 would be included in every set X satisfying the formula

$$X(1) \wedge \forall x (X(x) \rightarrow X(x+1)).$$

□

Note 92. Radical ideals and algebraic sets

Statement. Let K be algebraically closed and $I \subseteq K[X_1, \dots, X_n]$ an ideal. Then

$$I = I(V(I)) \iff I \text{ is radical.}$$

The forward implication holds over any field, while the reverse implication can fail over a field that is not algebraically closed.

Proof. The direction " $\Leftarrow$ " follows from Hilbert's Nullstellensatz. Suppose then that $I = I(V(I))$, or equivalently $I = I(V(\sqrt{I}))$. Since $V(\sqrt{I})$ contains only points where all the polynomials from $\sqrt{I}$ vanish, we know that every polynomial in $\sqrt{I}$ vanishes on $V(\sqrt{I})$. Since I contains all the polynomials which vanish on $V(\sqrt{I})$, we get that $\sqrt{I} \subseteq I$.

Hence I is a radical ideal. Note that we did not require K to be algebraically closed for " $\Rightarrow$ ".

The assumption that K is algebraically closed is crucial for " $\Leftarrow$ ". Consider for instance $I_1 = \mathbb{R}[X]$ and $I_2 = \langle X^2 + 1 \rangle$. Both of these ideals are radical. However, we have that $V(I_1) = V(I_2) = \emptyset$ and hence

$$I(V(I_1)) = I(V(I_2)) = \mathbb{R}[X].$$

Thus $I_2 \neq I(V(I_2))$.

□

Note 95. Types and prime ideals over a field

Statement. Let F be a field and $n \geq 1$. For a tuple in an algebraically closed extension, put $I_F(\bar{a}) = \{f \in F[\bar{X}] : f(\bar{a}) = 0\}$. Then:

- (i) $I_F(\bar{a})$ is prime, and $\text{tp}(\bar{a}/F) = \text{tp}(\bar{b}/F)$ exactly when the corresponding ideals agree.
- (ii) Complete n -types over F in algebraically closed extensions correspond bijectively to prime ideals of $F[\bar{X}]$.
- (iii) $V(I_F(\bar{a}))$ is the smallest algebraic set defined over F that contains $\bar{a}$.
- (iv) The type-to-ideal bijection is continuous from the Stone topology to the Zariski topology, but its inverse is not continuous.

Proof. 1. Let $K \supseteq F$ be an algebraically closed field. Given $\bar{a} \in K^n$ we define

$$I_F(\bar{a}) := \{f(\bar{x}) \in F[\bar{x}] \mid f(\bar{a}) = 0\}.$$

Clearly $I_F(\bar{a})$ is a proper ideal. Since K is a domain, $I_F(\bar{a})$ is also a prime ideal.

Let $\bar{a}, \bar{b} \in K^n$. We claim that

$$\text{tp}(\bar{a}/F) = \text{tp}(\bar{b}/F)$$

iff $I_F(\bar{a}) = I_F(\bar{b})$. The direction from left to right follows from the fact that for every $f(\bar{x}) \in I_F(\bar{a})$ we have that $f(\bar{x}) = 0 \in \text{tp}(\bar{a}/F)$. Let's consider then the converse direction. Since ACF has quantifier-elimination, every formula with parameters from F is equivalent to a Boolean combination of formulas of the form $f(\bar{x}) = 0$, with $f \in F[\bar{x}]$. Whether or not a tuple $\bar{a}$ satisfies such a Boolean combination is of course completely determined by $I_F(\bar{a})$.

2. We define

$$f : S_n(\text{ACF}, F) \rightarrow \text{Spec}(F[x_1, \dots, x_n])$$

by setting that

$$f(p) = \{f(\bar{x}) \in F[\bar{x}] \mid f(\bar{x}) = 0 \in p\}.$$

The first part essentially shows that this is an injection. To prove that it is surjection, note that if I is any prime ideal, then we can define

$$p := \{f(\bar{x}) = 0 \mid f \in I\} \cup \{\neg f(\bar{x}) = 0 \mid f \notin I\}.$$

To justify consistency, embed the integral domain $F[\bar{x}]/I$ into its fraction field and then into an algebraically closed field. The images of $x_1, \dots, x_n$ realize exactly the displayed polynomial equalities and inequalities. Their complete type p satisfies $f(p) = I$. Quantifier elimination makes this complete extension unique.

3. It is easy to see that $V(I_F(\bar{a}))$ is the smallest variety which contains $\bar{a}$ and which is defined by polynomials over F . Indeed, let $P \subseteq F[\bar{x}]$ be a set of polynomials such that $\bar{a} \in V(P)$. Thus $f(\bar{a}) = 0$, for every $f(\bar{x}) \in P$. Hence $P \subseteq I_F(\bar{a})$, which implies that $V(I_F(\bar{a})) \subseteq V(P)$.
4. We prove that f is continuous. For a basis of $\text{Spec}(F[\bar{x}])$, we take sets of the form

$$\{I \mid f \in I\} \subseteq \text{Spec}(F[\bar{x}])$$

to be closed. The pre-image of such a set is

$$\{p \mid f(\bar{x}) = 0 \in p\},$$

which is a closed set in the standard topology of $S_n(\text{ACF}, F)$.

For $n \geq 1$, f^{-1} is not continuous, whether F is finite or infinite. If f^{-1} were continuous, then f would be a homeomorphism from $S_n(\text{ACF}, F)$ to $\text{Spec}(F[\bar{x}])$. However, this cannot be the case, because the former space is Hausdorff while the latter is not.

Here we show that $S_n(\text{ACF}, F)$ is Hausdorff. Let p_1, p_2 be complete n -types over F . Now there exists $\varphi(\bar{x})$ such that $\varphi(\bar{x}) \in p_1$ and $\neg\varphi(\bar{x}) \in p_2$. The sets

$$U_1 := \{p \mid \varphi(\bar{x}) \in p\} \text{ and } U_2 := \{p \mid \neg\varphi(\bar{x}) \in p\}$$

are open and disjoint. Furthermore $p_1 \in U_1$ and $p_2 \in U_2$.

□

Note 98. Parameter-free subsets of a vector space

Statement. *In the usual vector-space language containing scalar multiplication by every element of K , the parameter-free definable subsets of a K -vector space V in one variable are precisely*

$$\emptyset, \quad \{0\}, \quad V \setminus \{0\}, \quad V,$$

with repetitions possible in the trivial vector space.

Proof. Clearly $\emptyset$ and V are definable subsets of V . Furthermore, the formula $x = 0$ defines the set $\{0\}$ and hence its negation defines $V \setminus \{0\}$. (Note that even if 0 did not have a corresponding constant in our vocabulary, we could have defined it with the formula $1x = 0x$.) We claim that these are all the parameter-free definable subsets of V .

Aiming for a contradiction, suppose that there exists a definable set $U \subseteq V$ which is not one of the previously mentioned sets. Let $\varphi(x)$ be a formula which defines U . Pick any $v \in U \setminus \{0\}$. Since for every $0 \neq \lambda \in K$ the mapping $x \mapsto \lambda x$ is an automorphism of V , $\langle v \rangle \setminus \{0\} \subseteq U$. Since U is not V nor $V \setminus \{0\}$, there must exist $u \in V \setminus (U \cup \{0\})$. Since $\langle v \rangle \setminus \{0\} \subseteq U$, the vectors v and u are linearly independent.

Hence there exists an automorphism of V which maps v to u . Since $V \models \varphi(v)$ we also have that $V \models \varphi(u)$, a contradiction. $\square$

Note 133. Primitive powers in a finite field

Statement. Let p be prime and g a generator of $\mathbb{F}_p^*$. For an integer m , the element g^m is primitive exactly when $\gcd(m, p-1) = 1$. More generally, the order of m in the additive group $\mathbb{Z}_n$ is $n/\gcd(m, n)$.

Proof. Indeed, in the additive group $\mathbb{Z}_n$, the element m has order $n/\gcd(m, n)$. To see this, write $m = dm'$ and $n = dn'$, where $d = \gcd(m, n)$ and $\gcd(m', n') = 1$. Then

$$km \equiv 0 \pmod{n} \iff n' \mid k.$$

Thus m generates $\mathbb{Z}_n$ exactly when $\gcd(m, n) = 1$. Apply this with $n = p-1$ under the isomorphism $m \mapsto g^m$. $\square$

Note 163. An infinitesimal from a Dedekind cut

Statement. If $\mathcal{R}$ is a proper elementary extension of $\mathbb{R}$ as an ordered field, then $\mathcal{R}$ contains a positive infinitesimal. Such an element can be obtained from the difference between a nonstandard element and the real supremum of its lower Dedekind cut, or from the reciprocal of an infinite element.

Proof. Aiming for a contradiction, suppose that this is not the case. Pick $r \in \mathcal{R} \setminus \mathbb{R}$ and consider the set $\{q \in \mathbb{R} \mid q <^{\mathcal{R}} r\} \subseteq \mathbb{R}$. Since $\mathcal{R}$ does not contain infinitesimals, it does not contain infinities. In particular $-n <^{\mathcal{R}} r <^{\mathcal{R}} n$, for some $n \in \mathbb{N}$, and thus the set $\{q \in \mathbb{R} \mid q <^{\mathcal{R}} r\}$ is non-empty and bounded in $\mathbb{R}$. Let

$$r^* := \sup\{q \in \mathbb{R} \mid q <^{\mathcal{R}} r\} \in \mathbb{R}.$$

Since $\mathcal{R}$ is an extension of $\mathbb{R}$, we have $r^* \in \mathcal{R}$. Now, by assumption, $|r^* - r| \in \mathcal{R}$ is not infinitesimal. Hence $|r^* - r| >^{\mathcal{R}} \varepsilon$, for some positive real ε . We have two cases.

1. Suppose first that $r <^{\mathcal{R}} r^* - \varepsilon$. Then $r^* - \varepsilon$ would be an upper bound for the set $\{q \in \mathbb{R} \mid q <^{\mathcal{R}} r\}$ which is strictly smaller than r^* , a contradiction.
2. Suppose then that $r >^{\mathcal{R}} r^* + \varepsilon$. Then $(r^* + \varepsilon) \in \{q \in \mathbb{R} \mid q <^{\mathcal{R}} r\}$ and thus $r^* + \varepsilon \leq r^*$, a contradiction.

Thus there must exist an infinitesimal in $\mathcal{R}$.

□

Note 164. Polynomial derivatives through infinitesimals

Statement. Let $\mathcal{R}$ be an elementary extension of $\mathbb{R}$, let $p \in \mathbb{R}[X]$, and let $x \in \mathbb{R}$. For every nonzero infinitesimal $\varepsilon \in \mathcal{R}$,

$$\frac{p(x + \varepsilon) - p(x)}{\varepsilon} = p'(x) + \eta,$$

where η is zero or infinitesimal.

Proof. Let

$$I := \{0\} \cup \{\varepsilon \in \mathcal{R} \mid \varepsilon \text{ is infinitesimal}\}.$$

Note that I is closed under addition and multiplication. Furthermore, if $r \in \mathbb{R}$ and $\alpha \in I$, then $r \cdot \alpha \in I$.

We claim that for every $p(X) \in \mathbb{R}[X]$ there exist $q(X) \in \mathbb{R}[X]$ such that for every $x \in \mathbb{R}$ and infinitesimal ε , there exist $\alpha \in I$ such that

$$\frac{p(x + \varepsilon) - p(x)}{\varepsilon} = q(x) + \alpha.$$

Fix x and a nonzero infinitesimal ε . The constant-polynomial case is immediate. We first prove this in the case where $p(X)$ is of the form X^n , for $n \geq 1$. We have that

$$\begin{aligned} & \frac{(x + \varepsilon)^n - x^n}{\varepsilon} \\ &= \frac{\sum_{i=0}^n \binom{n}{i} \varepsilon^i x^{n-i} - x^n}{\varepsilon} \\ &= \frac{\sum_{i=1}^n \binom{n}{i} \varepsilon^i x^{n-i}}{\varepsilon} \\ &= nx^{n-1} + \underbrace{\sum_{i=2}^n \binom{n}{i} \varepsilon^{i-1} x^{n-i}}_{\in I} \end{aligned}$$

Thus for $p(X) = X^n$ we can take $q(X) = nX^{n-1}$. Now we show that if the claim holds for $p(X)$ and $p'(X)$, then it holds for

$$ap(X) \text{ and } p(X) + p'(X)$$

and thus for arbitrary polynomials. First, we have

$$\frac{ap(x + \varepsilon) - ap(x)}{\varepsilon} = a \frac{p(x + \varepsilon) - p(x)}{\varepsilon} = a(q(x) + \alpha) = aq(x) + \underbrace{a\alpha}_{\in I}.$$

Secondly, we have

$$\begin{aligned}
 & \frac{p(x + \varepsilon) + p'(x + \varepsilon) - (p(x) + p'(x))}{\varepsilon} \\
 &= \frac{p(x + \varepsilon) - p(x)}{\varepsilon} + \frac{p'(x + \varepsilon) - p'(x)}{\varepsilon} \\
 &= q(x) + \beta + q'(x) + \beta' \\
 &= (q(x) + q'(x)) + \underbrace{(\beta + \beta')}_{\in I}.
 \end{aligned}$$

We are done. □

Note 166. Definability of the inverse Frobenius map

Statement. *If K is algebraically closed of characteristic $p > 0$, then every $a \in K$ has a unique p th root. The function $a \mapsto a^{1/p}$ is definable by the formula $y^p = x$.*

Proof. First, since K is algebraically closed, for every $a \in K$ we have that $a^{1/p}$ exists as a root of $X^p - a$. This also shows that the function is definable, since it is defined by $\varphi(X, Y) := Y^p - X = 0$. To show that it is a well-defined function, pick $a \in K$ and suppose that $b^p - a = c^p - a$ for some $b, c \in K$.

Now $b^p - c^p = 0$ and since K has characteristic $p > 0$, Freshman's dream implies that $0 = b^p - c^p = (b - c)^p$ and hence that $b = c$. Thus $X^p - a$ has a unique root for every a and therefore $a \mapsto a^{1/p}$ is indeed a well-defined function. □

Note 167. The algebraic closure of a finite field

Statement. *Every element of $\mathbb{F}_p^{\text{alg}}$ lies in a finite subfield. Equivalently,*

$$\mathbb{F}_p^{\text{alg}} = \bigcup \{K : K \leq \mathbb{F}_p^{\text{alg}}, K \text{ finite}\}.$$

Proof. The direction " $\supseteq$ " is immediate, so we focus on " $\subseteq$ ". Let a be an element of $\mathbb{F}_p^{\text{alg}}$. Every element of $\mathbb{F}_p(a) \leq \mathbb{F}_p^{\text{alg}}$ is of the form $p(a)/q(a)$, where $p, q \in \mathbb{F}_p[X]$ and $q(a) \neq 0$. Furthermore, since $\mathbb{F}_p^{\text{alg}} = \text{acl}^K(\mathbb{F}_p)$ for an algebraically closed field K , there is a polynomial $f(X)$ with coefficients in $\mathbb{F}_p$ such that $f(a) = 0$. Let d be the degree of f .

Now a^d can be expressed as a polynomial of degree at most $d - 1$ with coefficients from $\mathbb{F}_p$. Hence every element of $\mathbb{F}_p(a)$ is of the form $p(a)/q(a)$, where $p, q \in \mathbb{F}_p[X]$, $q(a) \neq 0$ and p, q have degree at most $d - 1$. Since $\mathbb{F}_p$ is finite, there are only finitely many such elements and hence $\mathbb{F}_p(a)$ is finite. Since it is a subfield of $\mathbb{F}_p^{\text{alg}}$, we are done. □

Note 171. Maximal ideals over an algebraically closed field

Statement. *For an algebraically closed field K , an ideal $I \subseteq K[X_1, \dots, X_n]$ is maximal if and only if*

$$I = (X_1 - a_1, \dots, X_n - a_n)$$

for some $(a_1, \dots, a_n) \in K^n$. This characterization can fail when K is not algebraically closed.

Proof. First, if I is generated by polynomials of the form $X_1 - a_1, \dots, X_n - a_n$, then

$$K[X_1, \dots, X_n] / \langle X_1 - a_1, \dots, X_n - a_n \rangle \cong K$$

by applying the first isomorphism theorem to the evaluation map $f(X_1, \dots, X_n) \mapsto f(a_1, \dots, a_n)$. Since K is a field, this is only possible if $\langle X_1 - a_1, \dots, X_n - a_n \rangle$ is maximal.

Conversely, suppose that I is maximal. Since I is proper, by weak Nullstellensatz we have that $V(I) \neq \emptyset$. Let $(a_1, \dots, a_n) \in V(I)$. We claim that $X_i - a_i \in I$, for each $1 \leq i \leq n$. Aiming for a contradiction, suppose that $X_i - a_i \notin I$, for some $1 \leq i \leq n$. By Hilbert's basis theorem, there is $f_1, \dots, f_m \in I$ which generate I . We claim that $\langle I \cup \{X_i - a_i\} \rangle = \langle f_1, \dots, f_m, X_i - a_i \rangle$ is proper. If it were not, it would contain 1, and therefore there would exist $g_1, \dots, g_{m+1} \in K[X_1, \dots, X_n]$ such that

$$1 = \sum_{j=1}^m g_j f_j + g_{m+1}(X_i - a_i),$$

which is not possible because the polynomial on the right vanishes on $(a_1, \dots, a_n)$. Thus $\langle I \cup \{X_i - a_i\} \rangle$ is a proper ideal extending I , a contradiction with the fact that I was maximal. Since $X_1 - a_1, \dots, X_n - a_n \in I$, we have $\langle X_1 - a_1, \dots, X_n - a_n \rangle \subseteq I$. As we have just shown that the former is a maximal ideal, we get that $I = \langle X_1 - a_1, \dots, X_n - a_n \rangle$, so I is generated by $X_1 - a_1, \dots, X_n - a_n$.

The assumption that K is algebraically closed is necessary. Indeed, consider the case $K = \mathbb{R}$ and $I = \langle X^2 + 1 \rangle$. Now $\mathbb{R}[X] / \langle X^2 + 1 \rangle \cong \mathbb{C}$, so $\langle X^2 + 1 \rangle$ is maximal. On the other hand, clearly $\langle X^2 + 1 \rangle$ is not generated by any polynomial of the form $X - a$, where $a \in \mathbb{R}$, since every polynomial in $\langle X - a \rangle$ vanishes on a . On the other hand, in the setting $K = \mathbb{C}$ we have that $\langle X^2 + 1 \rangle \subset \langle X - i \rangle$ and so it is not maximal. □

Note 172. One-variable definable sets in closed fields

Statement. *Every subset of an algebraically closed field K that is first-order definable in one field variable, with parameters allowed, is finite or cofinite.*

Proof. First, we know that every set definable over K is a finite union of sets of the form $V \setminus W$, where $V, W \subseteq K$ are varieties. But a variety is either $\emptyset$, K or a finite set of points. So $V \setminus W$ is clearly either finite or co-finite and so is any finite union of such sets. □

5.4 Groups and Boolean algebras

Note 84. Ordering torsion-free Abelian groups

Statement. *An Abelian group admits a translation-invariant total order if and only if it is torsion-free.*

Proof. Suppose first that $(G, \cdot, e)$ is a group with torsion. We claim that it does not have a total ordering. Aiming for a contradiction, suppose that $\leq$ is a translation-invariant total ordering of the group. Since the group has torsion, there exists $g \in G$ such that $g^n = e$ and $g^m \neq e$, for every $1 \leq m < n$. We have two cases. If $g < e$, then $e = g \cdot g^{n-1} < g^{n-1}$. On the other hand, from $g < e$ we get that

$$e > g > g^2 > g^3 > g^4 > \cdots > g^{n-1}$$

and thus $e < g$, a contradiction. The case $g > e$ is completely analogous.

Suppose then that G is a torsion-free Abelian group. We first consider the case where G is finitely generated. Thus

$$G \cong \mathbb{Z}^n,$$

for some $n \in \mathbb{N}$. The natural ordering $\leq$ of $\mathbb{Z}$ is a translation-invariant total ordering of it. Let $\leq_{\text{lex}}$ be the canonical lexicographical ordering of $\mathbb{Z}^n$. We show that this is a translation-invariant total ordering of $\mathbb{Z}^n$. Let $\bar{a}, \bar{b}, \bar{c} \in \mathbb{Z}^n$ such that $\bar{a} <_{\text{lex}} \bar{b}$. Thus there exists $1 \leq i \leq n$ such that

$$a_j = b_j \quad (j < i), \quad a_i < b_i.$$

Since $\leq$ is a translation-invariant total ordering of $\mathbb{Z}$, we get that

$$a_j + c_j = b_j + c_j \quad (j < i), \quad a_i + c_i < b_i + c_i.$$

Hence $\bar{a} + \bar{c} \leq_{\text{lex}} \bar{b} + \bar{c}$.

Now, let G be an arbitrary torsion-free Abelian group. We know that each of its finitely generated subgroups has a total ordering. We will now use a compactness argument to show that also G must have a total ordering. Let $\text{Diag}(G)$ denote the diagram of G and let $\varphi(\leq)$ be a sentence which states the binary relation $\leq$ is a translation-invariant total ordering of G . Clearly it suffices to show that $\text{Diag}(G) \cup \{\varphi(\leq)\}$ has a model.

By compactness, it suffices to show that each of its finite subsets has a model. Let $\Gamma \subseteq \text{Diag}(G)$ be a finite subset. Now only finitely many constants occur in Γ , say $g_1, \dots, g_n \in G$. Since $\langle g_1, \dots, g_n \rangle_G$ has a total ordering, Γ has a model.

□

Note 179. Groups with finitely many subgroups

Statement. *Every group with only finitely many subgroups is finite.*

Proof. First, since G has only finitely many subgroups, there exists $g_1, \dots, g_N \in G$ such that for every $g \in G$ we have that $\langle g \rangle = \langle g_i \rangle$, for some $1 \leq i \leq N$. Next, we note that

$$G \subseteq \langle g_1 \rangle \cup \cdots \cup \langle g_N \rangle.$$

Indeed, if $g \in G$, then for some $1 \leq i \leq N$ we have $\langle g \rangle = \langle g_i \rangle$ and thus in particular $g \in \langle g_i \rangle$. So all we need to show is that each $\langle g_i \rangle$ is finite. Aiming for a contradiction, suppose that some $\langle g_i \rangle$ is infinite. Thus g_i has infinite order. Consider the subgroups

$$\langle g_i^2 \rangle, \langle g_i^4 \rangle, \langle g_i^8 \rangle, \dots$$

Clearly $\langle g_i^{2^n} \rangle \supseteq \langle g_i^{2^{n+1}} \rangle$. On the other hand, the containment is proper, as $g_i^{2^n} \notin \langle g_i^{2^{n+1}} \rangle$. Indeed, otherwise

$$g_i^{2^n} = g_i^{2^{n+1}m},$$

for some $m \in \mathbb{Z}$, which contradicts the fact that g_i has infinite order. So $\langle g_i \rangle$ and therefore also G contains infinitely many subgroups, a contradiction. So each $\langle g_i \rangle$ is finite. □

Note 158. The two functionally complete binary connectives

Statement. *Exactly two binary Boolean connectives are functionally complete as the sole connective, without constants: NAND, $p \mid q = \neg(p \wedge q)$, and NOR, $p \uparrow q = \neg(p \vee q)$.*

Proof. Each defines negation by applying the connective to two copies of the same variable. NAND then defines conjunction, and NOR defines disjunction, so both are functionally complete.

Conversely, let f be a functionally complete binary connective. If $f(0,0) = 0$, every term built from f preserves zero. If $f(1,1) = 1$, every such term preserves one. Functional completeness therefore requires

$$f(0,0) = 1, \quad f(1,1) = 0.$$

There are four binary truth tables with these diagonal values. When the two off-diagonal values agree, they give NAND or NOR. When they differ, $f(p,q)$ is either $\neg p$ or $\neg q$. In either of the latter cases, induction shows that every term is a literal, so the connective is not functionally complete. □

Note 173. Complete atomic Boolean algebras as power sets

Statement. *For a complete atomic Boolean algebra B , the map $f : B \rightarrow \mathcal{P}(\text{Atom}(B))$ defined by*

$$f(x) = \{a \in \text{Atom}(B) : a \leq x\}$$

is an isomorphism onto the power-set Boolean algebra $\mathcal{P}(\text{Atom}(B))$. In particular, a finite Boolean algebra with n atoms has 2^n elements.

Proof. (Injection) Suppose that $x \neq y$. Then there is $z \in \text{Atom}(B)$ such that either $z \leq x$ and $z \not\leq y$ or $z \not\leq x$ and $z \leq y$. In both cases $f(x) \neq f(y)$.

(Order preserving) Suppose that $x \leq y$. Then for every $z \in \text{Atom}(B)$ we have that if $z \leq x$ then also $z \leq y$. Thus $f(x) \subseteq f(y)$. For the converse direction, suppose that $x \not\leq y$. Then $y \wedge x \neq x$. Since $x \wedge y \leq x$, there exist $z \in \text{Atom}(B)$ such that $z \leq x$ and $z \not\leq x \wedge y$. So we must have that $z \not\leq y$ and therefore $f(x) \not\subseteq f(y)$.

Note that this already shows that there is always an order preserving embedding of any atomic Boolean algebra to the power-set algebra.

(Surjection) Let $X \subseteq \text{Atom}(B)$. Since B is complete, we know that $\sup X$ exists in B . We claim that $f(\sup X) = X$. Clearly $f(\sup X) \supseteq X$, since $\sup X \geq z$ for every $z \in X$. Conversely, pick $z \in f(\sup X)$ so $z \in \text{Atom}(B)$ and $z \leq \sup X$. Suppose that $z \notin X$. Then for every $u \in X$ we have that $0 = u \wedge z$ or equivalently

$$\neg z = (u \wedge z) \vee \neg z = (u \vee \neg z) \wedge (z \vee \neg z) = u \vee \neg z$$

In other words, for every $u \in X$ we have $u \leq \neg z$ and therefore $\sup X \leq \neg z$. Since $z \leq \sup X$, we get $z \leq \neg z$ and so $z = z \wedge \neg z = 0$, which is a contradiction, because z was an atom.

As a corollary we get that every finite Boolean algebra $(B, \leq)$ has size 2^n , where n is the number of atoms in B . This follows from the above isomorphism together with the fact that every finite Boolean algebra is clearly atomic and complete. □

Note 174. Density in an atomless Boolean algebra

Statement. *The order of an atomless Boolean algebra is dense: whenever $x < y$, there is z with $x < z < y$. The Boolean algebra of finitary propositional formulas over infinitely many propositional variables, modulo equivalence, is atomless.*

Proof. Fix $x < y$ in B . We claim that $y \wedge \neg x \neq 0$. If $y \wedge \neg x = 0$, then $x \geq y > x$, a contradiction. Thus $y \wedge \neg x \neq 0$. Now, since B is atomless, there exists $z \neq 0$ such that $z < y \wedge \neg x$. In particular $z < y$ and $z \leq \neg x$. We claim that

$$x < x \vee z < y$$

We start by considering the first inequality. Since $x \leq x \vee z$, it suffices to show that $x \vee z \neq x$. To see this, note that otherwise $z = x \wedge z$ and therefore

$$z = z \wedge \neg x = z \wedge x \wedge \neg x = 0$$

which is not possible since we chose z so that $z \neq 0$. For the second inequality, again it suffices to show that $x \vee z \neq y$ as clearly $x \vee z \leq y$. If $x \vee z = y$, then

$$y \wedge \neg x = (x \vee z) \wedge \neg x = (x \wedge \neg x) \vee (z \wedge \neg x) = z \wedge \neg x = z,$$

which is a contradiction as $z \neq y \wedge \neg x$. Thus $x < x \vee z < y$.

Note that if we consider propositional logic over an infinite set of proposition symbols, then we have an atomless Boolean algebra with elements

$$[\phi] := \{\psi \mid \phi \equiv \psi\}$$

and $[\phi] \leq [\psi]$ defined by $\phi \models \psi$. To see that this is atomless, note that if $\phi \neq \perp$, then $\phi \wedge p \neq \perp$ where p is a proposition symbol not occurring in ϕ . Thus $0 < [\phi \wedge p] < [\phi]$. However, if instead of propositional logic we consider infinitary propositional logic, the resulting algebra is complete and atomic. □

Note 175. Boolean rings and Boolean algebras

Statement. *Let B be a ring with identity in which every element is idempotent. Then $x + x = 0$ for all x , multiplication is commutative, and*

$$x \wedge y = xy, \quad x \vee y = x + y + xy, \quad \neg x = 1 + x$$

define a Boolean algebra. Conversely, a Boolean algebra becomes a Boolean ring by taking multiplication to be meet and addition to be symmetric difference.

Proof. We have $x + x = 0$, for every x . Indeed

$$\begin{aligned} x + x &= (x + x)(x + x) \\ &= x \cdot x + x \cdot x + x \cdot x + x \cdot x \\ &= x + x + x + x \end{aligned}$$

and subtracting $x + x$ from both sides gives us $0 = x + x$. This immediately implies that $+$ is commutative, since $x + y = x + y + (y + x) + (y + x) = y + x$. Furthermore, multiplication is also commutative, since

$$x + y = (x + y)^2 = x^2 + xy + yx + y^2 = x + xy + yx + y,$$

whence $xy + yx = 0$.

Now we define

$$x \wedge y := x \cdot y, \quad x \vee y := x + y + xy \text{ and } \neg x = 1 + x$$

Since $+, \cdot$ are commutative, it is trivial to check that $(B, \wedge, \vee, 0, 1)$ is indeed a Boolean algebra.

Conversely, suppose that we are given a Boolean algebra $(B, \wedge, \vee, 0, 1)$. If we define $x \cdot y := x \wedge y$ and $x + y := (x \wedge \neg y) \vee (\neg x \wedge y)$, then $(B, +, \cdot, 0, 1)$ is a Boolean ring. □

6. Logic and computation

6.1 Computability and arithmetic

Note 101. First-order satisfiability and the word problem

Statement. *First-order satisfiability is undecidable. In particular, the word problem for finitely presented semigroups reduces to deciding whether a first-order sentence describing the presentation and a word inequality has a model. The first-order theory of semigroups is also undecidable.*

Proof. Let $\mathcal{T}$ be a Thue system. Let $g_1, \dots, g_m$ be its generators and $W_1 = V_1, \dots, W_r = V_r$ a finite set of word equations. Now each word W can be written as term in the language $\{g_1, \dots, g_m\} \cup \{\cdot\}$, where $g_1, \dots, g_m$ are constant symbols and $\cdot$ is a binary function symbol. Consider the following set of first-order axioms.

- (1) $\forall x \forall y \forall z ((x \cdot y) \cdot z = x \cdot (y \cdot z))$
- (2) $W_1 = V_1 \wedge \dots \wedge W_r = V_r$
- (3) $\forall x \forall y (\exists z_1 \exists z_2 \exists w_1 \exists w_2 (x = z_1 \cdot w_1 \cdot z_2 \wedge y = z_1 \cdot w_2 \cdot z_2 \wedge w_1 = w_2) \rightarrow x = y)$
- (4) $\neg V = W$

Let φ be a conjunction of (1) - (4). Clearly φ is satisfiable iff the word equation $V = W$ does not hold in $\mathcal{T}$. The fact that the first-order theory of semigroups is undecidable also follows from the previous construction, because we can replace the constants $g_1, \dots, g_m$ by existentially quantifying m distinct elements. (Axiom (3) is redundant.)

□

Note 120. Parity requires quantifiers in additive arithmetic

Statement. *The set of even natural numbers, defined by $\exists y (x = y + y)$, is not definable by a quantifier-free formula in $(\mathbb{N}, +, 0, 1)$. Indeed, every quantifier-free definable subset in one variable is finite or cofinite.*

Proof. To prove this, we show that every quantifier-free formula in one variable defines either a finite or a co-finite set. First, note that an atomic formula over $\{+, 0, 1\}$ is equivalent over $(\mathbb{N}, +, 0, 1)$ with an equation of the form

$$ax + b = cx + d.$$

Such an equation is either an identity, has no solution, or has exactly one solution. It therefore defines a finite or cofinite set. Finite and cofinite sets form a Boolean algebra, so the claim follows.

□

Note 121. Initial segments of computable ordinals

Statement. *If an ordinal α has a computable presentation, then every $\beta < \alpha$ also has a computable presentation. Here a computable presentation is a computable linear order on a computable subset of $\mathbb{N}$ that is isomorphic to the ordinal.*

Proof. Let $(A, <)$ be a computable presentation of α , and let $f : (A, <) \rightarrow (\alpha, \in)$ be an isomorphism. Since $\beta < \alpha$, there is a unique $b \in A$ such that $f(b) = \beta$. Define

$$A_\beta := \{a \in A : a < b\}.$$

Then

$$(A_\beta, < \upharpoonright (A_\beta \times A_\beta)) \cong (\beta, \in).$$

Moreover, A_β is computable, because A and $<$ are computable and b is fixed. The restricted ordering is computable as well. Therefore β is computable. □

Note 123. A diagonal noncomputable function

Statement. *Let $(M_n)_{n \in \mathbb{N}}$ enumerate the Turing machines. The total function*

$$U(n) = \begin{cases} 0, & M_n(n) \text{ halts with output } 1, \\ 1, & \text{otherwise} \end{cases}$$

is not computable.

Proof. Aiming for a contradiction, suppose that it is. Then there is $n \in \mathbb{N}$ such that M_n computes U . Now observe that $U(n) = 1$ iff $M_n(n) = 0$, because M_n halts on every input. Hence

$$M_n(n) = 1 \Leftrightarrow U(n) = 1 \Leftrightarrow M_n(n) = 0,$$

a contradiction. □

Note 150. A derivability proof of second incompleteness

Statement. *Let Σ be a consistent first-order theory with numerals for Gödel codes, and let $\text{der}(x)$ satisfy the following derivability conditions for all sentences φ, ψ :*

A1 *If $\Sigma \vdash \varphi$, then $\Sigma \vdash \text{der}(\ulcorner \varphi \urcorner)$.*

A2 *$\Sigma \vdash (\text{der}(\ulcorner \varphi \urcorner) \wedge \text{der}(\ulcorner \varphi \rightarrow \psi \urcorner)) \rightarrow \text{der}(\ulcorner \psi \urcorner)$.*

A3 *$\Sigma \vdash \text{der}(\ulcorner \varphi \urcorner) \rightarrow \text{der}(\ulcorner \text{der}(\ulcorner \varphi \urcorner) \urcorner)$.*

If some sentence φ_0 satisfies $\Sigma \vdash \varphi_0 \leftrightarrow \neg \text{der}(\ulcorner \varphi_0 \urcorner)$, then

$$\Sigma \not\vdash \neg \text{der}(\ulcorner 0 \neq 0 \urcorner).$$

Thus Σ does not prove the consistency statement expressed by this derivability predicate.

Proof. We will need the following auxiliary facts:

F1 For every $\varphi, \psi \in \text{FO}[\tau]$, we have that if $\Sigma \vdash \varphi \rightarrow \psi$, then

$$\Sigma \vdash \text{der}(\ulcorner \varphi \urcorner) \rightarrow \text{der}(\ulcorner \psi \urcorner).$$

F2 $\Sigma \vdash \text{der}(\ulcorner \varphi_0 \urcorner) \rightarrow \text{der}(\ulcorner \neg \varphi_0 \urcorner)$.

To prove F1, note that we can simply use A1 and then A2. To prove F2, observe first that, by the defining property of φ_0 ,

$$\Sigma \vdash \text{der}(\ulcorner \varphi_0 \urcorner) \rightarrow \neg \varphi_0.$$

By F1, we therefore have

$$\Sigma \vdash \text{der}(\ulcorner \text{der}(\ulcorner \varphi_0 \urcorner) \urcorner) \rightarrow \text{der}(\ulcorner \neg \varphi_0 \urcorner).$$

Hence, by A3,

$$\Sigma \vdash \text{der}(\ulcorner \varphi_0 \urcorner) \rightarrow \text{der}(\ulcorner \text{der}(\ulcorner \varphi_0 \urcorner) \urcorner).$$

Combining these two implications gives

$$\Sigma \vdash \text{der}(\ulcorner \varphi_0 \urcorner) \rightarrow \text{der}(\ulcorner \neg \varphi_0 \urcorner),$$

as required.

Now, by F1 and F2, we get

$$\Sigma \vdash \text{der}(\ulcorner \varphi_0 \urcorner) \rightarrow \text{der}(\ulcorner \varphi_0 \wedge \neg \varphi_0 \urcorner).$$

From this it follows that

$$\Sigma \vdash \text{der}(\ulcorner \varphi_0 \urcorner) \rightarrow \text{der}(\ulcorner \neg 0 = 0 \urcorner).$$

Indeed, we have

$$\Sigma \vdash \varphi_0 \wedge \neg \varphi_0 \rightarrow \neg 0 = 0,$$

and therefore, by F1,

$$\Sigma \vdash \text{der}(\ulcorner \varphi_0 \wedge \neg \varphi_0 \urcorner) \rightarrow \text{der}(\ulcorner \neg 0 = 0 \urcorner).$$

Hence

$$\Sigma \vdash \text{der}(\ulcorner \varphi_0 \urcorner) \rightarrow \text{der}(\ulcorner \neg 0 = 0 \urcorner).$$

Now suppose, towards a contradiction, that

$$\Sigma \vdash \neg \text{der}(\ulcorner \neg 0 = 0 \urcorner).$$

Then

$$\Sigma \vdash \neg \text{der}(\ulcorner \varphi_0 \urcorner).$$

By the defining property of φ_0 , this implies that

$$\Sigma \vdash \varphi_0.$$

But then, by A1,

$$\Sigma \vdash \text{der}(\ulcorner \varphi_0 \urcorner).$$

Using again the defining property of φ_0 , we get

$$\Sigma \vdash \neg \varphi_0.$$

This contradicts the consistency of Σ . Therefore

$$\Sigma \not\vdash \neg \text{der}(\ulcorner \neg 0 = 0 \urcorner).$$

□

Note 156. Provability of true existential arithmetic statements

Statement. If $\varphi(x)$ is a Σ_1^0 arithmetic formula and $\mathbb{N} \models \varphi(n)$, then $\text{PA} \vdash \varphi(\mathbf{n})$. In particular, using the standard arithmetized proof predicate,

$$\text{ZFC} \vdash \psi \iff \text{PA} \vdash \text{Prov}_{\text{ZFC}}(\ulcorner \psi \urcorner).$$

Proof. Let $\varphi(x) =: \exists \bar{y} \psi(x, \bar{y})$. Since $\mathbb{N} \models \varphi(n)$, there is $\bar{m} \in \mathbb{N}^*$ such that $\mathbb{N} \models \psi(n, \bar{m})$. Thus it is enough to show that $\text{PA} \vdash \psi(\mathbf{n}, \bar{\mathbf{m}})$. Now the axioms of PA clearly imply even without induction schemas that for each $N \in \mathbb{N}$ we have that

$$\text{If } \mathcal{A} \models \text{PA}, \text{ then } \mathcal{A} \upharpoonright \{0^{\mathcal{A}}, \dots, N^{\mathcal{A}}\} \cong \mathbb{N} \upharpoonright \{0, \dots, N\}.$$

Here

$$\mathcal{A} \upharpoonright \{0^{\mathcal{A}}, \dots, N^{\mathcal{A}}\} \text{ and } \mathbb{N} \upharpoonright \{0, \dots, N\}$$

denote partial structures in the sense that $+$ and $\cdot$ are undefined for inputs whose output is too large.

Let N be a natural number larger than any number that occurs when we evaluate $\psi(n, \bar{m})$. Note that such an N exists, because ψ does not contain unbounded quantifiers. Assume now that $\text{PA} \not\vdash \psi(\mathbf{n}, \bar{\mathbf{m}})$. Thus there is a model $\mathcal{A}$ such that $\mathcal{A} \models \text{PA}$ and $\mathcal{A} \models \neg \psi(\mathbf{n}, \bar{\mathbf{m}})$. Now we get a clear contradiction, because

$$\mathcal{A} \models \neg \psi(\mathbf{n}, \bar{\mathbf{m}}) \text{ iff } \mathcal{A} \upharpoonright \{0^{\mathcal{A}}, \dots, N^{\mathcal{A}}\} \models \neg \psi(\mathbf{n}, \bar{\mathbf{m}})$$

$$\text{iff } \mathbb{N} \upharpoonright \{0, \dots, N\} \models \neg \psi(n, \bar{m}) \text{ iff } \mathbb{N} \models \neg \psi(n, \bar{m}).$$

Hence $\text{PA} \vdash \varphi(\mathbf{n})$.

As an application we get the following. Let $\text{Prov}_{\text{ZFC}}(\ulcorner \varphi \urcorner)$ be a Σ_1^0 -formula which states that φ can be proved using ZFC. Then

$$\text{ZFC} \vdash \varphi \text{ iff } \mathbb{N} \models \text{Prov}_{\text{ZFC}}(\ulcorner \varphi \urcorner) \text{ iff } \text{PA} \vdash \text{Prov}_{\text{ZFC}}(\ulcorner \varphi \urcorner)$$

That is, if PA proves that φ is provable from ZFC, then ZFC proves φ .

□

Note 157. Limits of existential second-order self-description

Statement. *Over standard arithmetic, neither of the following sets of Gödel codes is definable by an existential second-order formula:*

- (i) *codes of ESO sentences $\psi \vee \chi$ such that $\mathbb{N} \models \neg\psi \wedge \chi$;*
- (ii) *codes of ESO sentences $\exists x \psi(x)$ such that $\mathbb{N} \models \neg\psi(0) \wedge \psi(1)$.*

Both sets consist entirely of true ESO sentences.

Proof. The set T of Gödel codes of true ESO sentences over standard arithmetic is Σ_1^1 -complete, so its complement is not Σ_1^1 .

Suppose the first set were ESO-definable. The computable map $\varphi \mapsto \varphi \vee \top$ sends a sentence into that set exactly when $\mathbb{N} \models \varphi$. Its inverse image would therefore make the complement of T ESO-definable, a contradiction.

For the second set, use the computable map

$$\varphi \mapsto \exists x ((x = 0 \wedge \varphi) \vee x = 1).$$

The matrix is true at 1 and is false at 0 exactly when φ is false. Thus ESO-definability of the second set would again make the complement of T Σ_1^1 , giving the same contradiction. $\square$

6.2 Fragments, games and model checking**Note 96. Boolean combinations of bounded static logic**

Statement. *Let $\mathcal{L} = \text{Bool}(\text{BndSCL})$ be the effective logic of finite Boolean combinations of bounded static computational sentences, with the semantics described below. Its satisfiability and validity sets are arithmetical. Moreover, every satisfiable sentence of $\mathcal{L}$ has a countable model.*

Proof. Let $\mathcal{L} := \text{Bool}(\text{BndSCL})$. That is, $\mathcal{L}$ consist of sentences which are (finite) Boolean combinations of sentences of BndSCL. More formally, a **sentence** of $\mathcal{L}$ is a pair

$$(\phi, f),$$

where ϕ is a propositional formula with variables $p_1, \dots, p_n$ and

$$f : \{p_1, \dots, p_n\} \rightarrow \text{BndSCL}$$

is an injection. The semantics of $\mathcal{L}$ is defined as follows. Let $\varphi = (\phi, f) \in \mathcal{L}[\tau]$ and $\mathfrak{A}$ a τ -structure. We define φ^f to be the formula obtained from ϕ by replacing each propositional symbol p with

$$\bigvee_{n \in \mathbb{N}} \Phi_{f(p)}^n.$$

We then define $\mathfrak{A} \models_{\mathcal{L}} \varphi$ iff $\mathfrak{A} \models \varphi^f$. Clearly $\mathcal{L}$ is an effectively regular logic. We claim that the set of valid sentences of $\mathcal{L}$ is arithmetical and that it has the countable downwards Löwenheim-Skolem property.

For the latter claim we note that trivially $\mathcal{L} \leq \mathcal{L}_{\omega_1\omega}$. Since the latter logic has the countable downwards Löwenheim-Skolem property, the claim follows.

Next we show that the set of valid sentences of $\mathcal{L}$ is arithmetical. Since $\mathcal{L}$ is effectively closed under negation, it suffices to show that the set of satisfiable sentences of $\mathcal{L}$ is arithmetical. Consider a sentence $\varphi = (\phi, f) \in \mathcal{L}[\tau]$. We want to determine whether φ has a model. First, we guess a valuation $v : \{p_1, \dots, p_n\} \rightarrow \{0, 1\}$, where $p_1, \dots, p_n$ contains all the proposition symbols in ϕ . If $v(\phi) = 0$, then we reject.

Otherwise for each p for which $v(p) = 1$ we guess $n_p \in \mathbb{N}$. We then check whether or not

$$\bigwedge_{v(p)=1} \Phi_{f(p)}^{n_p} \wedge \bigwedge_{v(p)=0} \neg f(p)$$

is satisfiable. This is equivalent to checking whether or not

$$\bigvee_{v(p)=1} \neg \Phi_{f(p)}^{n_p} \vee \bigvee_{v(p)=0} f(p)$$

is not valid. Since the validity problem for BndSCL is in Σ_1^0 , this latter problem is in Π_1^0 . Hence the whole problem is in Σ_2^0 , i.e., it is arithmetical. $\square$

Note 80. A first-moment bound for modal CNF formulas

Statement. Let $m_\exists \geq 1$, $m_\forall \geq 0$, and

$$\varphi = \bigwedge_{i=1}^{m_\exists} \Diamond \psi_i^\exists \wedge \bigwedge_{j=1}^{m_\forall} \Box \psi_j^\forall.$$

Each ψ_i^Q is an independently sampled k -CNF over a set Φ of Boolean variables, with m independent clauses. Each clause uses k distinct uniformly chosen variables and independent uniform signs. Then

$$\Pr(\varphi \text{ is satisfiable}) \leq 2^{m_\exists |\Phi|} (1 - 2^{-k})^{m(m_\exists + m_\forall)}.$$

Proof. It is easy to see that φ is satisfiable iff there exists a function $f : [m_\exists] \rightarrow \mathcal{P}(\Phi)$ such that the following conditions are satisfied.

1. $f(i) \models \psi_i^\exists$, for every $1 \leq i \leq m_\exists$.
2. $f(i) \models \bigwedge_{j=1}^{m_\forall} \psi_j^\forall$, for every $1 \leq i \leq m_\exists$.

Such a function f is called a **witness** for φ . Suppose that the formulas ψ_i^Q , for $Q \in \{\exists, \forall\}$, are sampled independently at random. We are interested in giving an upper bound on

$$\Pr[\text{There exists a witness for } \varphi].$$

Applying the union bound we get that

$$\begin{aligned} \Pr[\text{There exists a witness for } \varphi] &\leq \sum_f \Pr[f \text{ is a witness for } \varphi] \\ &= \sum_f \Pr[\forall i \in [m_\exists] : f(i) \models \psi_i^\exists \wedge \bigwedge_{j=1}^{m_\forall} \psi_j^\forall] \end{aligned}$$

Since the formulas ψ_i^Q are sampled independently, we get that

$$\Pr[\forall i \in [m_\exists] : f(i) \models \psi_i^\exists \wedge \bigwedge_{j=1}^{m_\forall} \psi_j^\forall] = \Pr[\forall i \in [m_\exists] : f(i) \models \psi_i^\exists] \cdot \Pr[\forall i \in [m_\exists] : f(i) \models \bigwedge_{j=1}^{m_\forall} \psi_j^\forall].$$

Again, since the formulas ψ_i^Q are sampled independently, we get that

$$\begin{aligned} \Pr[\forall i \in [m_\exists] : f(i) \models \psi_i^\exists] &= \prod_{i \in [m_\exists]} \Pr[f(i) \models \psi_i^\exists] \\ &= \prod_{i \in [m_\exists]} (1 - 2^{-k})^m \\ &= (1 - 2^{-k})^{m \cdot m_\exists}. \end{aligned}$$

Now

$$\begin{aligned} \sum_f \Pr[\forall i \in [m_\exists] : f(i) \models \psi_i^\exists \wedge \bigwedge_{j=1}^{m_\forall} \psi_j^\forall] \\ \leq (1 - 2^{-k})^{m \cdot m_\exists} \sum_f \Pr[\forall i \in [m_\exists] : f(i) \models \bigwedge_{j=1}^{m_\forall} \psi_j^\forall] \end{aligned}$$

Each summand is maximized by witness candidates f which map the whole set $[m_\exists]$ to a single assignment. Given such a witness candidate we get that

$$\Pr[\forall i \in [m_\exists] : f(i) \models \bigwedge_{j=1}^{m_\forall} \psi_j^\forall] = \Pr[f(1) \models \bigwedge_{j=1}^{m_\forall} \psi_j^\forall] = (1 - 2^{-k})^{m \cdot m_\forall}.$$

Hence we get that

$$\begin{aligned} (1 - 2^{-k})^{m \cdot m_\exists} \sum_f \Pr[\forall i \in [m_\exists] : f(i) \models \bigwedge_{j=1}^{m_\forall} \psi_j^\forall] \\ \leq (1 - 2^{-k})^{m(m_\exists + m_\forall)} \cdot 2^{m_\exists |\Phi|}. \end{aligned}$$

□

Note 128. Model checking a relational-algebra fragment

Statement. *For finite structures whose input relations are listed explicitly, the model-checking problem for $\text{GRA}(\neg, \cap, \exists)$ is decidable in polynomial time in the combined input size. Here model checking asks whether a given tuple belongs to the relation denoted by the input term.*

Proof. We first note that for every $\mathcal{T} \in \text{GRA}(\neg, \cap, \exists)$ there exist $\mathcal{T}' \in \text{GRA}(\setminus, \cap, \cup, \exists, \forall)$ such that $\mathcal{T} \equiv \mathcal{T}'$ or $\neg \mathcal{T} \equiv \mathcal{T}'$. Furthermore, $\mathcal{T}'$ can be computed from $\mathcal{T}$ in polynomial time. This can be established by using structural induction and De-Morgan laws. Thus it suffices to give a polynomial time algorithm for computing $\mathcal{T}^\mathfrak{A}$, for a given structure $\mathfrak{A}$ and a term $\mathcal{T} \in \text{GRA}(\setminus, \cap, \cup, \exists, \forall)$.

Let

$$\mathcal{T}_1, \dots, \mathcal{T}_m$$

be an enumeration of the subterms of $\mathcal{T}$ which satisfies the following condition: if $\mathcal{T}_i$ is a proper subterm of $\mathcal{T}_j$, then $i < j$. The algorithm performs a for-loop over $i = 1, \dots, m$ as follows.

1. If $\mathcal{T}_i$ is R , we compute $\mathcal{T}_i^{\mathfrak{A}}$.
2. If $\mathcal{T}_i$ is $\mathcal{T}_j \setminus \mathcal{T}_k$, then we know that we have already computed $\mathcal{T}_j^{\mathfrak{A}}$ and $\mathcal{T}_k^{\mathfrak{A}}$. We can then compute $\mathcal{T}_i^{\mathfrak{A}}$ by removing from $\mathcal{T}_j^{\mathfrak{A}}$ all the tuples that are also in $\mathcal{T}_k^{\mathfrak{A}}$.
3. If $\mathcal{T}_i$ is $\mathcal{T}_j \cap \mathcal{T}_k$, then we know that we have already computed $\mathcal{T}_j^{\mathfrak{A}}$ and $\mathcal{T}_k^{\mathfrak{A}}$. We can then compute $\mathcal{T}_i^{\mathfrak{A}}$ by taking the intersection of $\mathcal{T}_j^{\mathfrak{A}}$ and $\mathcal{T}_k^{\mathfrak{A}}$.
4. If $\mathcal{T}_i$ is $\mathcal{T}_j \cup \mathcal{T}_k$, then we know that we have already computed $\mathcal{T}_j^{\mathfrak{A}}$ and $\mathcal{T}_k^{\mathfrak{A}}$. We can then compute $\mathcal{T}_i^{\mathfrak{A}}$ by taking the union of $\mathcal{T}_j^{\mathfrak{A}}$ and $\mathcal{T}_k^{\mathfrak{A}}$.
5. If $\mathcal{T}_i$ is $\exists \mathcal{T}_j$, then we know that we have already computed $\mathcal{T}_j^{\mathfrak{A}}$. We can then compute $\mathcal{T}_i^{\mathfrak{A}}$ by removing the last element from every tuple in $\mathcal{T}_j^{\mathfrak{A}}$.
6. If $\mathcal{T}_i$ is $\forall \mathcal{T}_j$, then we know that we have already computed $\mathcal{T}_j^{\mathfrak{A}}$. We can then compute $\mathcal{T}_i^{\mathfrak{A}}$ as follows: we include those tuples $\bar{a}$ for which $\bar{a}b \in \mathcal{T}_j^{\mathfrak{A}}$, for every $b \in A$.

Clearly the above algorithm runs in polynomial time with respect to the input. □

Note 129. Model checking with signed relations

Statement. *The polynomial-time model-checking algorithm for $\text{GRA}(\neg, \cap, \exists)$ can operate directly on signed relations $(X, \#)$, where $\# \in \{+, -\}$ records whether the denoted relation is X or its complement. Every intermediate stored relation has size polynomial in the explicit input, so expansion into the full complement is unnecessary.*

Proof. The idea is that we compute for each term $\mathcal{T}$ a pair

$$(X, \#),$$

where X is a relation over A and $\# \in \{+, -\}$. Let $\bar{a}$ be a tuple over A . Intuitively speaking, $\bar{a}$ satisfies $\mathcal{T}$ if the following equivalence holds:

$$\bar{a} \in X \Leftrightarrow \# = +.$$

Let

$$\mathcal{T}_1, \dots, \mathcal{T}_m$$

be an enumeration of the subterms of $\mathcal{T}$ which satisfies the following condition: if $\mathcal{T}_i$ is a proper subterm of $\mathcal{T}_j$, then $i < j$. The algorithm performs a for-loop over $i = 1, \dots, m$ as follows.

1. If $\mathcal{T}_i$ is R , we return $(R^{\mathfrak{A}}, +)$.
2. Suppose that $\mathcal{T}_i$ is $\neg \mathcal{T}_j$. Let $(X, \#)$ be the pair we computed for $\mathcal{T}_j$. If $\# = +$, we return $(X, -)$ and otherwise we return $(X, +)$.
3. Suppose that $\mathcal{T}_i$ is $\mathcal{T}_j \cap \mathcal{T}_k$. Let $(X, \#_1)$ and $(Y, \#_2)$ be the pairs we computed for $\mathcal{T}_j$ and $\mathcal{T}_k$ respectively. We have several cases.
 - a) If $\#_1 = \#_2 = +$, we return $(X \cap Y, +)$.
 - b) If $\#_1 = \#_2 = -$, we return $(X \cup Y, -)$.
 - c) If $\#_1 = +$ and $\#_2 = -$, we return $(X \setminus Y, +)$.

- d) If $\#_1 = -$ and $\#_2 = +$, we return $(Y \setminus X, +)$.
4. Suppose that $\mathcal{T}_i$ is $\exists\mathcal{T}_j$. Let $(X, \#)$ be the pair we computed for $\mathcal{T}_j$. If $\# = +$, we return $(\exists X, +)$. Otherwise we return $(\forall X, -)$.

To see that the above algorithm runs in polynomial time with respect to the input, note that for every pair $(X, \#)$ computed by the algorithm we have that X has size polynomial with respect to the size of the input model $\mathfrak{A}$.

□

Note 132. Defining successors with two-variable counting

Statement. *There is a sentence ψ of two-variable first-order logic with counting, in the vocabulary $\{S, L\}$, such that for every finite nonempty structure in which L is a strict linear order,*

$$\mathfrak{A} \models \psi \iff S^{\mathfrak{A}} \text{ is the immediate-successor relation of } L^{\mathfrak{A}}.$$

Proof. Indeed, we can take ψ to be the conjunction of the following sentences.

$$\begin{aligned} \psi_1 &:= \forall x \forall y (S(x, y) \rightarrow L(x, y)) \\ \psi_2 &:= \forall x \exists^{\leq 1} y S(x, y) \\ \psi_3 &:= \forall x \exists^{\leq 1} y S(y, x) \\ \psi_4 &:= \exists^{\neq 1} x \forall y \neg S(x, y) \\ \psi_5 &:= \exists^{\neq 1} x \forall y \neg S(y, x) \end{aligned}$$

To see this, suppose that $\mathfrak{A} \models \psi$ and $L^{\mathfrak{A}}$ is a linear order. Let a and b be the smallest and largest elements of $L^{\mathfrak{A}}$ respectively. If $a = b$, then we are done, so suppose that $a \neq b$. Now, it is clear that

$$\mathfrak{A} \models \forall y \neg S(b, y),$$

Thus every element besides b needs to have a unique S -successor. Starting from a , we can find a sequence

$$c_1, \dots, c_n,$$

where $c_1 = a$, $c_n = b$ and for every $1 \leq i < n$ the element c_{i+1} is the unique S -successor of c_i . It suffices to show that every element of $\mathfrak{A}$ is c_i , for some $1 \leq i \leq n$. Aiming for a contradiction, suppose that there exist d such that d is not c_i for any $1 \leq i \leq n$. Now there is an S -path

$$e_1, \dots, e_m$$

from d to b . Let $1 \leq j < m$ be the largest index for which $e_j \neq c_i$, for any $1 \leq i \leq n$. Thus $e_{j+1} = c_i$, for some $1 \leq i \leq n$. Now $c_i \neq a$, because e_{j+1} has an S -predecessor, so $i > 1$. Thus $\{(e_j, c_i), (c_{i-1}, c_i)\} \subseteq S^{\mathfrak{A}}$, a contradiction.

□

Note 136. A bijection game on equivalence relations

Statement. Let $\kappa < \lambda$ be infinite cardinals. Let $\mathcal{M}$ be an equivalence relation with κ classes of size λ , and $\mathcal{N}$ one with λ classes of size λ . Duplicator wins every finite-round bijection game between these structures. Nevertheless, the Magidor–Malitz quantifier Q_λ^{MM} , expressing the existence of a pairwise satisfying set of size λ , separates them by the property of having at least λ equivalence classes.

Proof. Suppose that the players have played the game for m rounds and the current position is

$$\begin{aligned} a_1 &\mapsto f_1(a_1) \\ a_2 &\mapsto f_2(a_2) \\ &\vdots \\ a_m &\mapsto f_m(a_m) \end{aligned}$$

Since Duplicator has not lost the game, for every $i < j$ we have that $a_i \sim a_j$ iff $f_i(a_i) \sim f_j(a_j)$. The Duplicator now picks some bijection

$$f_{m+1} : M \rightarrow N$$

which bijectively maps $[a_i]_\sim$ to $[f_i(a_i)]_\sim$ with $a_i \mapsto f_i(a_i)$. Let $a_{m+1} \in M$ be the element picked by the Spoiler. We claim that $a_{m+1} \sim a_i$ iff $f_{m+1}(a_{m+1}) \sim f_i(a_i)$. If $a_{m+1} \sim a_i$, then we know that $f_{m+1}(a_{m+1}) \sim f_i(a_i)$, since $[a_{m+1}]_\sim = [a_i]_\sim$ is mapped to $[f_i(a_i)]_\sim$. Conversely, if $f_{m+1}(a_{m+1}) \sim f_i(a_i)$, then there is $b \in [a_i]_\sim$ such that $f_{m+1}(b) = f_{m+1}(a_{m+1})$. Since f_{m+1} is an injection, we get that $b = a_{m+1}$ and hence $a_{m+1} \sim a_i$.

Hence the Duplicator has a winning strategy in the bijective-game between $\mathcal{M}$ and $\mathcal{N}$. This means that no extension of FO with monotone unary quantifiers can separate the two structures. On the other hand, a Magidor–Malitz quantifier Q_λ^{MM} can separate them. Indeed, we have that

$$\mathcal{A} \models Q_\lambda^{MM} xy (\neg x = y \rightarrow \neg x \sim y)$$

iff there exists a set $B \subseteq A$ of cardinality λ such that for every distinct $x, y \in B$ we have that $\neg x \sim y$. That is, iff there exist at least λ many distinct equivalence classes. This is clearly true in $\mathcal{N}$ and false in $\mathcal{M}$. □

Note 148. Defining finite ordered structures with few variables

Statement. Let τ be a finite relational vocabulary containing a strict linear order, and let $k = \max_{R \in \tau} \text{ar}(R) \geq 2$. Relative to linearly ordered structures, each finite τ -structure is characterized up to isomorphism by an $\text{FO}^k[\tau]$ -sentence. Consequently, every class of finite ordered τ -structures is definable in $\mathcal{L}_{\infty, \omega}^k[\tau]$.

Proof. It suffices to show that each finite linearly ordered τ -structure can be defined in $\text{FO}^k[\tau]$, where $k := \max\{\text{ar}(R) \mid R \in \tau\}$. Note that $k \geq 2$, because $< \in \tau$. Fix such a τ -structure $\mathfrak{A}$ of size n . First, there is a sentence θ_n of $\text{FO}^2[<]$ which states that the size of a linearly ordered structure is n .

To see this, note that we can clearly say in $\text{FO}^2[<]$ that there is a $<$ -path of length at least $n - 1$ ¹ and therefore we can also say that there is no $<$ -path longer than that.

To express the relational structure of $\mathfrak{A}$, we use for each $1 \leq i \leq n$ an $\text{FO}^2[<]$ formula $\psi_i(x)$ which states that x is the i th element of $\mathfrak{A}$. Again, this formula can be constructed by saying that there is a $<$ -path of length at least $i - 1$ ending at x and there is no $<$ -path longer than that.

Now, $\mathfrak{A}$ is defined by

$$\begin{aligned} \varphi_{\mathfrak{A}} := & \theta_n \wedge \bigwedge_{R \in \tau} \bigwedge_{(i_1, \dots, i_m) \in R^{\mathfrak{A}}} \exists x_1 \dots \exists x_m \left(\bigwedge_{j=1}^m \psi_{i_j}(x_j) \wedge R(x_1, \dots, x_m) \right) \\ & \wedge \bigwedge_{R \in \tau} \forall x_1 \dots \forall x_m \left(R(x_1, \dots, x_m) \rightarrow \bigvee_{(i_1, \dots, i_m) \in R^{\mathfrak{A}}} \bigwedge_{j=1}^m \psi_{i_j}(x_j) \right), \end{aligned}$$

which is a formula of $\text{FO}^k[\tau]$. Therefore a given model class $\mathcal{K}$ is simply defined by

$$\bigvee_{\mathfrak{A} \in \mathcal{K}/\cong} \varphi_{\mathfrak{A}},$$

which is a formula of $\mathcal{L}_{\infty, \omega}^k[\tau]$. □

6.3 Fixed points and expressive power

Note 142. Translating least fixed points into second-order logic

Statement. *Every least fixed-point formula has an effectively constructible equivalent second-order formula. For a positive defining formula $\varphi(R, \bar{x})$, membership in its least fixed point is expressed by*

$$\forall R \left(\forall \bar{x} (\varphi(R, \bar{x}) \rightarrow R\bar{x}) \rightarrow R\bar{a} \right).$$

If every fixed-point operator occurs in positive polarity, the resulting sentence can be expressed in universal second-order logic.

Proof. To see this, consider a formula

$$[\text{lfp}_{R, \bar{x}} \varphi(R, \bar{x})](\bar{a})$$

Assume that R occurs only positively in φ . Suppose that there exists a formula $\varphi^*(R, \bar{x})$ of SO which is equivalent to $\varphi(R, \bar{x})$. Now

$$[\text{lfp}_{R, \bar{x}} \varphi(R, \bar{x})](\bar{a})$$

is equivalent to

$$\forall R \left(\forall \bar{x} (\varphi^*(R, \bar{x}) \rightarrow R\bar{x}) \rightarrow R\bar{a} \right).$$

Thus we can recursively and effectively translate sentences of LFP to equivalent sentences of SO. Furthermore, the same argument shows that sentences of LFP where fixed point

¹Length of path is the number of edges on that path.

operator is not in the scope of an odd number of negations can be translated effectively to sentences of USO.

□

Note 143. The countable model property for least fixed points

Statement. *Least fixed-point logic over relational vocabularies has the countable model property: every satisfiable sentence has a model of cardinality at most $\aleph_0$.*

Proof. We work in a relational vocabulary and argue using the game-theoretic semantics of LFP in terms of parity games.

Let $\mathfrak{A} \models \varphi$, where φ is a sentence of LFP. Then the Verifier has a winning strategy in the parity game

$$\mathcal{G}(\mathfrak{A}, \varphi).$$

Since parity games are positionally determined, we may assume that this winning strategy σ is positional. Let T_σ be the labelled tree generated by σ . Thus T_σ starts from the initial position of $\mathcal{G}(\mathfrak{A}, \varphi)$, and each node is labelled by a position of the game. At positions controlled by the Verifier, the tree contains only the successor prescribed by σ , while at positions controlled by the Falsifier, the tree contains all possible successors.

Without loss of generality we can assume that φ is in negation normal form. For $B \subseteq A$, define $T_\sigma \upharpoonright B$ to be the labelled subtree obtained from T_σ by restricting universal moves to elements of B . More precisely, whenever a node is labelled by a position of the form

$$(s, \forall y \psi(\bar{x}, y)),$$

we keep only those children labelled by positions of the form

$$(s(a/y), \psi(\bar{x}, y))$$

where $a \in B$.

Define $W(T_\sigma \upharpoonright B)$ to be the set of all elements $a \in A$ such that $T_\sigma \upharpoonright B$ has a node labelled by a position of the form

$$(s, \exists y \psi(\bar{x}, y))$$

whose σ -successor is labelled by

$$(s(a/y), \psi(\bar{x}, y)).$$

Thus $W(T_\sigma \upharpoonright B)$ is the set of existential witnesses chosen by σ after universal choices have been restricted to B .

We note that if B is countable, then $T_\sigma \upharpoonright B$ is countable. Indeed, $T_\sigma \upharpoonright B$ has height ω , and every node has at most countably many children. Consequently $W(T_\sigma \upharpoonright B)$ is also countable.

We now define sets $B_n \subseteq A$ recursively. Choose some $b_0 \in A$, and set

$$B_0 := \{b_0\}.$$

Given B_n , define

$$B_{n+1} := B_n \cup W(T_\sigma \upharpoonright B_n).$$

Finally, let

$$B := \bigcup_{n < \omega} B_n.$$

By induction, each B_n is countable, and hence B is countable.

We claim that

$$\mathfrak{A} \upharpoonright B \models \varphi.$$

The strategy σ restricts to a winning strategy for the Verifier in

$$\mathcal{G}(\mathfrak{A} \upharpoonright B, \varphi).$$

Indeed, universal moves in the restricted game use only elements of B . Moreover, suppose that a finite play compatible with σ has so far used only elements of B . Then all elements occurring in the current position belong to some B_n . If the current position is existential, then the witness chosen by σ belongs to

$$W(T_\sigma \upharpoonright B_n) \subseteq B_{n+1} \subseteq B.$$

Thus the strategy σ never leaves the domain B .

Therefore every play of

$$\mathcal{G}(\mathfrak{A} \upharpoonright B, \varphi)$$

compatible with the restricted strategy is also a play of

$$\mathcal{G}(\mathfrak{A}, \varphi)$$

compatible with σ . Since σ is winning in $\mathcal{G}(\mathfrak{A}, \varphi)$, the parity condition is satisfied on every such play. Hence the Verifier wins

$$\mathcal{G}(\mathfrak{A} \upharpoonright B, \varphi).$$

By the game-theoretic semantics of LFP, it follows that

$$\mathfrak{A} \upharpoonright B \models \varphi.$$

Since B is countable, this proves that LFP has the countable model property. □

Note 144. Second-order Craig interpolation

Statement. *Second-order logic has Craig interpolation. For sentences $\varphi \in \text{SO}[\tau]$ and $\psi \in \text{SO}[\sigma]$ with $\varphi \models \psi$, there is $\theta \in \text{SO}[\tau \cap \sigma]$ such that $\varphi \models \theta \models \psi$. The same projection argument applies to a logic $\mathcal{L}$ closed under existential second-order quantification.*

Proof. Consider the sentence

$$\theta := \exists(R)_{R \in \tau \setminus \sigma} \varphi$$

of $\text{SO}[\tau \cap \sigma]$. We claim that θ is the desired interpolant. First, we clearly have that $\varphi \models \theta$. Furthermore, we have that $\theta \models \psi$. Indeed, if $\mathfrak{A} \models \theta$, then $\mathfrak{A}$ has an expansion $\mathfrak{A}^+$ such that $\mathfrak{A}^+ \models \varphi$. Now $\mathfrak{A}^+ \models \psi$, because $\varphi \models \psi$, and therefore $\mathfrak{A} \models \psi$. □

Note 145. First-order expressibility inside second-order logic

Statement. *The set*

$$\{\ulcorner \varphi \urcorner : \varphi \in \text{SO} \text{ is equivalent to a first-order sentence}\}$$

does not belong to any finite level of the analytical hierarchy.

Proof. We prove this by giving an effective reduction from the problem of determining whether a sentence of $\text{SO}[\{+, \cdot\}]$ is true on $(\mathbb{N}, +, \cdot)$. Let $\theta_{\mathbb{N}}$ be a sentence of $\text{SO}[\{+, \cdot\}]$ that characterizes $(\mathbb{N}, +, \cdot)$ up to isomorphism. For any sentence φ of $\text{SO}[\{+, \cdot\}]$ we claim that

$$\theta_{\mathbb{N}} \wedge \neg \varphi \text{ is equivalent to a sentence of FO} \Leftrightarrow (\mathbb{N}, +, \cdot) \models \varphi.$$

If $(\mathbb{N}, +, \cdot) \not\models \varphi$, then $\theta_{\mathbb{N}} \wedge \neg \varphi \equiv \theta_{\mathbb{N}}$ and therefore $\theta_{\mathbb{N}}$ is not equivalent to any sentence of FO as FO cannot characterize $(\mathbb{N}, +, \cdot)$ up to isomorphism. Conversely, if $(\mathbb{N}, +, \cdot) \models \varphi$, then $\theta_{\mathbb{N}} \wedge \neg \varphi$ is equivalent to a contradiction and a fortiori equivalent to a sentence of FO. □

Note 147. Limits on decidable second-order fragments

Statement. *There is no computable fragment $\mathcal{L} \subseteq \text{SO}$ that has decidable satisfiability and contains an equivalent sentence for every second-order sentence. The complexity classes Σ_2, Π_2 used in the argument are those of the Lévy hierarchy of set theory.*

Proof. Aiming for a contradiction, suppose that this is the case. Since $\mathcal{L}$ is computable and it has a decidable satisfiability problem, there is a Σ_1^0 -formula $\varphi_{\mathcal{L}\text{-SAT}}(x)$ such that

$$V \models \varphi_{\mathcal{L}\text{-SAT}}(n) \Leftrightarrow n \text{ is } \ulcorner \varphi \urcorner, \text{ for some } \varphi \in \mathcal{L} \text{ which is satisfiable.}$$

Now, it is known that the validity problem for SO is in Π_2 . Since two sentences of φ, ψ of SO are equivalent if and only if $\varphi \leftrightarrow \psi$ is valid, there is a Π_2 -formula $\varphi_{\text{equiv}}(x, y)$ such that

$$V \models \varphi_{\text{equiv}}(n, m) \Leftrightarrow n, m \text{ are } \ulcorner \varphi \urcorner, \ulcorner \psi \urcorner, \text{ for some } \varphi, \psi \text{ such that } \varphi \equiv \psi.$$

Consider now the following Π_2 -formula

$$\varphi_{\text{SAT}}(n) := \exists m \in \omega (\varphi_{\mathcal{L}\text{-SAT}}(m) \wedge \varphi_{\text{equiv}}(n, m)).$$

Then

$$V \models \varphi_{\text{SAT}}(n) \Leftrightarrow n \text{ is } \ulcorner \varphi \urcorner, \text{ for some } \varphi \text{ such that } \varphi \text{ is satisfiable.}$$

Hence the satisfiability problem of SO would be in Π_2 , which is a contradiction because the satisfiability problem of SO is Σ_2 -complete and therefore not in Π_2 . $\square$

Note 149. The compact topology of first-order model classes

Statement. Fix a countable vocabulary τ , and let X consist of all τ -structures whose domain is $[n]$ for some $n \geq 1$, or $\mathbb{N}$. For a first-order sentence φ , put $X_\varphi = \{\mathfrak{A} \in X : \mathfrak{A} \models \varphi\}$. The family $(X_\varphi)_\varphi$ is a clopen basis for a compact topology on X .

Proof. To check the basis property, observe that $X_\top = X$ and

$$X_\varphi \cap X_\psi = X_{\varphi \wedge \psi}.$$

Each basic set is also closed because $X \setminus X_\varphi = X_{\neg\varphi}$.

For compactness, it suffices to consider a cover by basic open sets $\{X_\varphi : \varphi \in \Sigma\}$. If no finite subfamily covered X , every finite subset of $\{\neg\varphi : \varphi \in \Sigma\}$ would have a model in X . First-order compactness gives a model of the whole theory, and downward Löwenheim–Skolem gives a countable model. An isomorphic representative belongs to X and lies outside every member of the supposed cover, a contradiction. $\square$

Note 155. Failure of finite Beth definability

Statement. Beth’s definability theorem fails over finite structures: a first-order theory can implicitly determine a relation on every finite reduct without that relation being explicitly first-order definable on the finite models. Alternating membership in a unary predicate along a finite linear order gives such an example.

Proof. Aiming for a contradiction, suppose that it holds. Consider the vocabulary $\{<, P\}$, where $<$ is binary and P is unary. Consider the theory Σ which states that

1. $<$ is a linear order
2. smallest element of $<$ is in P
3. a successor of an element in P is not in P
4. a successor of an element not in P is in P

Clearly Σ defines P implicitly over finite models. Now, let $\theta(x)$ be a formula of $\text{FO}[\{<\}]$ such that

$$\Sigma \models \forall x (P(x) \leftrightarrow \theta(x))$$

in the finite and

$$\varphi := \exists x \forall y ((x = y \vee y < x) \wedge \theta(x)).$$

We now claim that if $\mathfrak{A}$ is a finite linear order, then

$$\mathfrak{A} \models \varphi \text{ iff } \mathfrak{A} \text{ has odd length,}$$

which results in a contradiction because FO cannot define this property over the class of finite linear orders.

Suppose first that $\mathfrak{A}$ is a finite linear order of odd length. Clearly it has an expansion $\mathfrak{A}^*$ to $\{<, P\}$ such that $\mathfrak{A}^* \models \Sigma$, namely the expansion where starting from the first element every other element is added to the interpretation of P . Thus the largest element of $\mathfrak{A}$ is in the interpretation of P and therefore $\mathfrak{A}^* \models \varphi$, because $\mathfrak{A}^* \models \Sigma$. But since P does not occur in φ , $\mathfrak{A} \models \varphi$. Conversely, suppose that $\mathfrak{A} \models \varphi$.

Again, $\mathfrak{A}$ has an expansion $\mathfrak{A}^*$ to $\{<, P\}$ such that $\mathfrak{A}^* \models \Sigma$. Since P does not occur in φ , $\mathfrak{A}^* \models \varphi$. Thus the largest element of $\mathfrak{A}^*$ satisfies $\theta(x)$. Since $\mathfrak{A}^* \models \Sigma$, we have that the largest element of $\mathfrak{A}^*$ is in the interpretation of P and therefore $\mathfrak{A}$ must be of odd length.

The reason why this argument does not work when infinite structures are included is that Σ no longer defines P implicitly. Consider for instance the structure $(\mathbb{N} + \mathbb{Z}, <)$, so natural numbers followed by the integers.

This structure has two different expansions to a model of Σ : in both expansion every other natural number starting from 0 is in the interpretation of P , but in the first expansion only the even integers are in the interpretation of P while in the second expansion it is the odd integers.

□

Note 176. Unbounded finite-cardinality quantifiers

Statement. Let $S \subseteq \mathbb{N}$ be unbounded, and interpret the generalized quantifier Q_S by

$$\mathfrak{A} \models Q_S x \varphi(x) \iff |\{a \in A : \mathfrak{A} \models \varphi(a)\}| \in S.$$

The validities of $\text{FO}[Q_S]$, allowing relational vocabularies with a binary relation, are not recursively enumerable.

Proof. Indeed, if they were, then in particular we could enumerate sentences of the form

$$\left(Q_S x x = x \wedge \exists x U(x) \right) \rightarrow \varphi^U$$

which would contradict Trakhtenbrot's theorem. Indeed, the antecedent implies that U is a non-empty subset of a finite set.

□

Note index

2	Variable-length random clauses	8	33	Covering the Boolean cube by translates	43
3	Specified letters in a random word	8			
4	Connectivity under the pure literal rule	9	34	Randomized NP protocols with polynomial advice	44
5	Cycles in sparse random graphs	9			
6	Common neighbours in dense random graphs	10	35	Functions of independent discrete variables	1
7	The threshold for isolated vertices	10	37	Covering numbers in doubling metric spaces	21
8	Oscillation of a symmetric random walk	15	38	Lower mass bounds for doubling measures	22
9	The mean of Brownian motion	24			
10	The variance of Brownian motion	24	39	From a doubling measure to a doubling metric	22
11	The covariance of Brownian motion	25			
12	Mean-square Brownian increments	25	41	The expectation and variance of a sample mean	47
13	The Gaussian characteristic function	26	42	The bias of the sample variance	47
14	Brownian motion as a martingale	26	43	A lower bound involving absolute moments	48
15	Quadratic variation in mean square	26			
16	A Brownian stochastic sum	27	44	Scaling Brownian hitting times	28
16a	Shannon entropy as a Rényi limit	40	45	Conditional expectation as a projection	28
17	Cross-isometry for stochastic step integrals	29	46	Moments of the exponential distribution	37
18	Invariant sets and indicator functions	17	47	The mean of a Poisson process	38
19	Periodic points of rational rotations	18	48	Strictly increasing Poisson arrival times	38
20	Periodic points of the doubling map	18	49	Divergence of Poisson arrival times	39
21	The tent map preserves Lebesgue measure	19	50	Coincidences of independent binomial variables	11
22	Pointwise recurrence on the unit interval	20	51	Separation of independent binomial variables	12
23	Sine coefficients of an even function	20	52	Kraft's inequality	41
24	An integral bound from Cauchy–Schwarz	20	53	The parity of a Bernoulli sum	13
25	The Fourier series of the identity function	21	54	A key-length obstruction to perfect secrecy	45
26	Information per symbol and the strong law	40	55	Predicting one bit through a one-time pad	45
27	Typical empirical letter frequencies	16	56	A PSPACE-complete oracle collapses P and NP	44
28	Infinitesimals in elementary extensions	74	57	Monotone paths in a random grid	13
29	An entropy identity for finite partitions	41	58	A threshold for selecting a perfect square	14
30	Zero-error randomized computation	42			
31	Monte Carlo integration and the strong law	16	59	Expected Bayes error under a uniform type prior	52

60	Mean absolute deviation of a beta variable	53	91	Oligomorphic groups and quantifier elimination	71
61	Expected Bayes error under a Dirichlet prior	54	92	Radical ideals and algebraic sets	76
62	Stochastic step integrals are martingales	29	94	Local-to-global existence by compactness	71
63	Square-integrability of Brownian motion	30	95	Types and prime ideals over a field	76
64	The stochastic integral of Brownian motion	31	96	Boolean combinations of bounded static logic	90
65	Integration by parts against Brownian motion	31	98	Parameter-free subsets of a vector space	77
66	Square-integrability of a stochastic integral	33	99	A supremum that is a successor ordinal	66
67	The exponential Brownian martingale	33	101	First-order satisfiability and the word problem	86
68	An affine Ornstein–Uhlenbeck equation	34	102	Counting structures of an infinite cardinality	67
69	A linear SDE with a time-dependent drift	35	103	Bayesian inference from two bags of fruit	3
70	A linear SDE with a vanishing diffusion	35	104	An exponential decision-tree lower bound	58
71	Moments of geometric Brownian motion	36	105	Local minima of a convex function	59
72	Borel sets in a product of separable metric spaces	7	106	Cancellation of ordinal successors	67
73	Conditioning on a disjoint union	2	107	Mean distance from the centre of a disk	49
74	Expectation under a push-forward law	3	108	Mean-square error implies consistency . . .	50
75	Packing and covering numbers	23	109	Conditional expectation and mean-square error	50
76	The median-of-means estimate	48	110	Variance decomposition in linear regression	50
77	Expected disagreement under a Dirichlet prior	56	111	Maximum likelihood for a Poisson parameter	51
78	The variance of a biased random walk	16	112	Comparing the one-norm and the two-norm	59
79	Margin perceptron and hinge-loss descent	57	113	The error of retaining the largest coordinates	60
80	A first-moment bound for modal CNF formulas	91	114	Quadratic regularization and strong convexity	61
81	Majority forests versus DNF formulas	58	115	Measurable functions preserve independence	1
82	Injective polynomial maps over closed fields	75	116	Factoring expectations of independent variables	1
83	Model completeness and universal formulas	68	117	Scaling Lebesgue integrals	4
84	Ordering torsion-free Abelian groups	82	118	The tail integral formula for expectation	4
85	Characteristic zero in existential monadic logic	75	119	Images of inverse images under a surjection	5
87	Well-orders inside the real line	66	120	Parity requires quantifiers in additive arithmetic	86
88	Countability of finite Cartesian powers	66	121	Initial segments of computable ordinals	87
89	Obstructions to universal second-order definitions	70	123	A diagonal noncomputable function	87
90	First-order Craig interpolation	70	126	The law of total expectation	5

127	The Bayes rule for asymmetric binary loss	63	153	Pigeonhole principles for infinite cardinals	67
128	Model checking a relational-algebra fragment	92	155	Failure of finite Beth definability	100
129	Model checking with signed relations ..	93	156	Provability of true existential arithmetic statements	89
130	The expected image size of a random function	14	157	Limits of existential second-order self-description	90
131	Adaptive bit prediction through a one-time pad	45	158	The two functionally complete binary connectives	83
132	Defining successors with two-variable counting	94	159	Winning decimal games and perfect subsets	68
133	Primitive powers in a finite field	78	160	Rational operations on irrational numbers	17
134	Total probability conditional on an event	6	162	Chromatic number forces edges	15
136	A bijection game on equivalence relations	95	163	An infinitesimal from a Dedekind cut ..	78
138	Scaled dot-product attention, entry by entry	62	164	Polynomial derivatives through infinitesimals	79
139	An affine Gaussian observation model ..	51	165	A saturated model at an inaccessible cardinal	73
140	Permutation equivariance of self-attention	63	166	Definability of the inverse Frobenius map	80
141	The hardness of finding a sparse solution	61	167	The algebraic closure of a finite field ..	80
142	Translating least fixed points into second-order logic	96	168	Optimizing confusion-matrix objectives	64
143	The countable model property for least fixed points	97	169	Optimizing the minimum of precision and recall	65
144	Second-order Craig interpolation	98	171	Maximal ideals over an algebraically closed field	81
145	First-order expressibility inside second-order logic	99	172	One-variable definable sets in closed fields	81
147	Limits on decidable second-order fragments	99	173	Complete atomic Boolean algebras as power sets	83
148	Defining finite ordered structures with few variables	95	174	Density in an atomless Boolean algebra	84
149	The compact topology of first-order model classes	100	175	Boolean rings and Boolean algebras ...	85
150	A derivability proof of second incompleteness	87	176	Unbounded finite-cardinality quantifiers	101
151	Elementary equivalence without partial isomorphism	72	177	Symmetric difference bounds probability differences	6
152	Continuum many countable models of true arithmetic	72	178	Bayesian confirmation is reciprocal	7
			179	Groups with finitely many subgroups ..	82