Nonstandard Mathematics

Boris Haase

14th Edition

Nonstandard Mathematics

Boris Haase

May 10, 2026

Gratefully dedicated to L

Mathematics Subject Classification: 03H05 [Nonstandard models in mathematics], 03E15 [Descriptive set theory], 54G20 [General topology: counterexamples], 11-XX [Number theory], 51M05 [Euclidean geometries (general)], 90CXX [Mathematical programming], 65-XX [Numerical analysis] and 68-XX [Computer science].

Abstract

This paper deepens the understanding of quantifying mathematical infinity. Open and closed sets turn out to be not tenable. The correct usage of bijections makes results obsolete based on the concept of cardinality. Elements of infinite sets are counted in a novel way. The Fueter-Pólya conjecture must be corrected. The fundamental theorem of set theory and the restricted Jacobian conjecture hold, but the hairy ball theorem does not. Archimedes' theorem replaces the axiom of the same name. The measure problem is solved. Exact integrals and differentials are also valid for non-continuous functions and (conventionally un-) measurable sets. They are leading to the key theorems and in some cases do not require the notion of holomorphicity. The concepts of convergence and continuity are reformulated. This reveals that the concepts of uniform convergence and (Hölder) continuity are superfluous. The Cauchy product is corrected. Liouville's theorem and the Picard theorem are proven to be false, as well as the identity theorem and the Riemann rearrangement theorem. The (generalised) Riemann hypothesis proves to be untenable. The prime number theorem is tightened and Singmaster's as well as Giuga's conjecture are elementarily proven. The greatest-prime criterion turns out to be efficient. Finite and infinite algebraic numbers are distinguished and determined including their distances, approximations and asymptotics. Algebraic numbers are characterised by means of the bounding and the coefficient theorem. Furthermore, Roth's theorem is proven to be false. Then Goldbach's strong conjecture is proven followed by the ones of Alaoglu, Bunyakovsky, Wilson and Erdős resp. the one by Collatz. The three-cube theorem, Beal's and Brocard's theorem as well as the Catalan conjecture are elementarily proven. The conventionally provable Littlewood conjecture is rejected in nonstandard mathematics. Several axioms of Euclidean geometry are reviewed revealing to be false or redundant. Toeplitz' conjecture is refuted and Fickett's conjecture is proven. After proving the diameter theorem, fast matrix multiplication and LUX, intex and merit method follow, which may accelerate linear algebra significantly. Discrete Fourier transform allows to differentiate, integrate and revert Taylor series in a novel way to determine zeros of polynomials - also by a fixed-point method. If derivatives are replaced by function values at the so-called crown points, the approximation theorem may be applied and analytic differential equations can be solved efficiently. The conclusion is the sorting method bit sort, proving the halting theorem and that complexity classes P and NP are equal. Finally, applications of the Euler–Maclaurin formulas in theoretical physics are presented.

14th improved and extended edition, from German by translated.net

Contents

1 Introduction 3

2 Set Theory 6

3 Topology 9

4 Nonstandard Analysis 10

5 Number Theory 20

6 Euclidean Geometry 24

7 Linear programming 25

8 Scientific Computing 33

9 Theoretical Informatics 38

10 Theoretical Physics 41

Curriculum Vitae 55

Appendix: FastShift Algorithm 56

List of Symbols 58

Bibliography 59

Index of Abbreviations

AD antiderivative

AN algebraic number

CO crossover

CP closed path

CS crown series

FFT fast Fourier transform

FSR fast-shift-recursive

FTD fast-Toeplitz decomposition

GPC greatest-prime criterion

GS geometric series

LI line integral

LP linear programme

NR neighbourhood relation

TS Taylor series

Everything finds its roots in the grace of God.

From the book "Relil"

1 Introduction

The following results in the branches set theory, topology, nonstandard analysis, number theory, Euclidean geometry, linear programming, scientific computing and theoretical informatics are brilliant achievements! Known statements and elementary concepts such as axiom, field, etc., are as given in the relevant literature or on Wikipedia. Hence, only deviating or clarifying definitions will follow.

Unlike conventional usage in which brackets denote a more detailed explanation, bracketed parts of a statement are both optional and valid. To avoid trivialities and for reasons such as the consequences of fame, proofs are not always fully carried out and are terminated with □, whereas definitions are terminated with △. The largest exponent of the argument of a polynomial is its degree. Minimal polynomials may be specified by integer (instead of rational) coefficients as throughout in the following text.

Preliminary consideration: The finiteness of the world causes certain difficulties in dealing with the infinite (where the infinitesimal is the reciprocal and the finite the dual term). A reference to philosophical concepts is justified, since mathematics cannot be explained metalinguistically from itself alone and abstractions provide the most general statements due to the principle of scientificity. The complete decomposition of infinite sets involves a successive process that cannot be completed.

The abrupt transition from finite to infinite numbers, which is difficult to convey, requires mid-finite ones. The existence of actual or only potentially infinite sets remains open, since the transcendence of the infinite fails to provide a proof. From the almost "wasteful" size or expansion of the universe, it can be concluded that there is an availability that has no recognisable limits.

However, such inferences are weaker than abduction. If a finite line is broken down into an infinite number of parts, the infinite is finitely limited. If, in addition, the number of parts of both infinities is the *same*, there is mathematically an isomorphism: The enlargement of the infinitely small parts of the finite segment to finite ones results in infinity in the conventional sense in relation to the whole. That easily deduces a bijection in the mathematical sense.

Rational numbers are real, put a minimum polynomial of degree 1 to 0 and some have periodic decimal places. If the degree is ≥ 2, these are purely algebraic numbers with infinite numerators and denominators. Finite real fractions, together with the infinite or purely complex fractions, already form all real or complex numbers. Real continued fractions that do not terminate as finite fractions are algebraic (conventionally transcendent!) since they have an infinite denominator.

(Conventionally and infinite natural) induction can show, starting with the set of conventionally natural numbers, that it can be diagonalised up to any power according to Cantor, so that all(!) infinite sets are equipotent to the set of conventionally natural numbers if Hilbert's translations are used as an aid. However, the numbers of the sets of natural and integers differ almost exactly by a factor of 2.

Algebra shows that sum, difference, product, and quotient of two conventional algebraic numbers of natural degree m or n are algebraic of degree at most mn, and that the $1/m$-th power of a conventional algebraic number of degree n is also algebraic of degree at most mn. If the algebraicity of a number is to be investigated and the remainder is given by the limit of a zero sequence (a_n), omitting the sequence values for large n is not permitted. They are crucial.

An infinite real number can consist of a finite continued fraction with an infinitely large last denominator. If it were equated to a conventionally rational number by removing the last partial fraction, it would also be the solution to a linear equation with (infinite) whole coefficients. No (infinite) subset of complex numbers is closed. The (finite) definition offers significant advantages over axioms in terms of handling and is traditional.

Treating bijections correctly concerning the number, there are infinite many sets between the set of conventionally natural numbers and the one of conventionally real numbers. Thus, the continuum hypothesis gets a new answer. The definition of real numbers via Dedekind cuts is thus just as redundant as its definition via equivalence classes of real Cauchy sequences. The whole set theory is of course more extensive since only essential (new) ideas are presented here.

A disk without its boundary conventionally represents an open set, because then each point of it has a conventional neighbourhood that lies completely in this set. If the points are considered on a half-line, starting at the centre of the disk, there must be always a real neighbourhood for each point on this half-line towards the boundary.

This idea so far negates, however, "the end of the flagpole" whereby every such half-line must contain a point in the interior of that disk without any conventional neighbourhood there. Hence, the term openness for sets is inept (cf. [10], p. 36). If the unit disk is considered around the origin of ordinates, so the last point of the half-line $[0, 1[$, dually represented, is the point $0.\overline{1}_2$ and the next point is the boundary point 1. There is no other point between these two points.

For this reason, the disk without boundary is also closed, since the considered end points of the half-line just form the closure as boundary. Because the neighbourhoods do not exist on their boundary, the term closure is meaningless for sets in Euclidean space. At the same time, every open set is also closed there. This absurdity is unsettling when infinitesimal quantities are considered in a differentiated manner, thus in particular the numbers 1 (rational!) and $0.\overline{1}_2$ (algebraic!) are not equated.

The absurd shows itself also by an infinite intersection of open sets such as by all open concentric disks forming a closed set, more precisely: the common centre of the disk. An infinite union of closed sets can build an open set as an open disk does, as a union of all its points as closed sets. A 0-dimensional set (point) is therefore open, because every neighbourhood also consists of one point.

Hence, the empty set $\emptyset$ is also closed, and consequently the complete Euclidean space is closed, what can easily be generalised to higher dimensions for spheres. This special case makes also the general one absurd and meaningless: considering open or closed sets is not suitable for metric and topological spaces. Particularly, the definition of a conventional topological space appears oddly content-free and arbitrary, and becomes dubious.

Permitting infinitesimal radiuses makes the terms inner and outer point as well as boundary point meaningful, however. The conventional irrationality proof of $\sqrt{2}$ is problematic since the square of the related rational number does not exist (of which numerator and denominator are infinite). The subject is also about conventionally unmeasurable, mid-finite and infinite sets as well as discontinuous functions. Every probabilistic statement is only valid when all its relevant possibilities were verified.

When integrating identical paths in opposite positive and negative directions, the counter-directional rule for integrals is adopted, stating that when following the path in the negative direction, the same (!) function value of both possible ones must be chosen if the function is too discontinuous, implying that the integral sums to 0 to prevent a significantly different value. Stokes' theorem may be proven generally (cf. [17], p. 625 f.). Function values may skip values, which are not measured then.

Conventional differentiation and integration lose the ability to distinguish between rationality and pure algebraicity in the conventional process of taking limits. This is e.g. problematic when setting about determining the roots of a polynomial exactly. Therefore, the conventional analysis cannot be preserved in its existing form and requires practicable alternatives. Period 2π must define sine and cosine since their power series only converge for finite arguments.

The exact volume integral is the easiest to handle compared to conventional ones. Its improper form arises analogously. In some cases, suitable Landau notation may be useful. Combining function values to finitely many continuous functions, integrals may be calculated even for discontinuous functions. Here, one of both Euler-Maclaurin formulas may be used. Note that the continuity of neither integral nor derivative is assumed. An appropriate definition makes this possible.

If the result of differentiation lies outside of the domain D, the closest number within the domain D should replace it. If this is not uniquely determined, the result can either be given as the set of all such numbers, or the preferred result may be selected (e.g. according to a uniform rule). The exact integral is more general than Riemann, Lebesgue(-Stieltjes) integral and other types of integral. The latter exist only in conventionally measurable sets. Extrapolations may detect bends.

Using $1/\infty$ instead of 0 avoids a division by 0 and any vague notions of limits but requires considering carefully where this replacement makes sense and how may be exponentiated to invoke no contradiction by switching the symbols. This also allows to define integrals and differentials for each operation on real and complex numbers in such a way that every function is at least directionally integrable and differentiable wherever the function values are defined.

The definition of the exact integral via a rectangular rule may require error estimations (cf. [7], [8] and [26]). Actual integration as the inverse operation to differentiation only makes sense for continuous functions if there is a wish to go beyond simple summation. The term predecessor, which is dual to successor, is usually not specifically mentioned again, but must be kept in mind. Only three examples illustrate the superiority of nonstandard analysis and the strength of using infinitesimal and infinite values.

Associative, commutative, and distributive laws allow to arbitrarily rearrange sums if care is taken to calculate them correctly (using Landau symbols). The Riemann series theorem is false, since the commutative law cannot be avoided, even at infinity. Arbitrariness is given if single summands are neglected or not considered at infinity.

Coefficients of Taylor series may be computed as equivalent crown series very easily (for complex numbers) by using discrete Fourier transform (DFT). This allows an efficient number representation and calculation. Functions can be determined as a vector or product of the Fourier matrix by a vector of fixed function values at so-called crown points, which are arranged in a circle around a crown centre to be chosen effectively.

The product mentioned can be stored before the calculation with a freely selectable precision, which for computer calculations is best the reciprocal of a power of two. Since it can be calculated quickly and Taylor series can be realised using the Horner scheme, this method (especially as an fast Fourier transform version) is very efficient. Here tensors replace matrices. Moduli are considered by using the signum function or sections are formed.

Suitable matching methods enable smooth transitions between individual sections. The necessary reversibility of operators as in the Adomian decomposition is no longer necessary. The Taylor series of the integrated logarithm, as a geometric series, allows a simple implementation of the multiplicative inverse. The exact derivation rules need not be known. Linear algebra can solve analytic differential equations numerically with comparable efficiency using crown series.

This book is based on ISO 80000-2:2019 (quantities and units - mathematics). The conventional notation of sums (Σ), products (Π), differentials (d and ∂), integrals ($\int$) and roots ($\sqrt{}$) is prevented: It is too sweeping and historically less intuitive. The minimality theorem explains the choice of 2 as basis (also for digital computers that are working most often within a binary system). In practice (of computer science), sufficiently small formal systems often help.

The combined LUX method, which was developed in 35 years and answers (e.g. by using Gomory cuts) Hilbert's tenth problem positively, can solve every solvable linear programme in quadratic time. Details are only published when no misuse for non-transparent or bad decisions must be feared. Fast matrix multiplication surpasses even the Strassen algorithm and allows solving systems of linear equations quickly.

Beauty and elegance in mathematics can be ensured by adequately thinking through what is to be presented and without being stingy, reducing it to the clear essence, which justifies both and is a hallmark of the true. Unfortunately, there is a lot of ugly length in the mathematical world. It can only be hoped that this book can provide a lot of pleasure with the nonstandard mathematics and gives insight into the real good and beautiful. Who likes it, may realise both of it!

2 Set Theory

Definition: An *entity* is everything that is distinguishable as a *being*. A *set* S is an entirety of entities, called *elements*, whose number is denoted by $|S|$. Only the *empty set* $\emptyset$ contains no elements. A set $S \neq \emptyset$ is called *finite* (otherwise *infinite*) if and only if it can be completely divided into subsets, where each subset contains the rounded half of the remaining elements. Let ν the largest *finite* real number, ω in between the largest *mid-finite* and ι the smallest positive one.$\triangle$

Conventionally, sufficiently well-understood axioms exist that define the real numbers as a linearly ordered field and the complex numbers as a field with the imaginary unit $\underline{1}$ (read as "im 1"). Analogously, addition, multiplication, and their inverse operations may be extended to the largest and by definition closed field extensions $\mathbb{R}$, and $\mathbb{C} := \mathbb{R} + \underline{\mathbb{R}}$ (with further operations like exponentiation). Every number from $\mathbb{C}^*$ is said to be *infinitesimal* if the absolute value of its reciprocal is infinite.

Presenting Peano axioms and field axioms is deliberately omitted here to remain clear and to emphasise the new content. Two numbers must have a minimum distance, since *all* different numbers of $\mathbb{R}$ are separated by a distance. Assuming the opposite of the minimum not being fixed leads to the contradiction that every number from $\mathbb{R}$ must have at least one nearest neighbour as a number, which is itself in $\mathbb{R}$ (cf. the isomorphism above).

The set $\mathbb{R}$ of all real numbers is isomorphic to a set of (hyper-) natural or integer numbers. It has both a fixed minimum element and a fixed maximum element, since $\mathbb{R}$ is viewed holistically and completely, from which its closure is obtained together with the field axioms. Otherwise, the former theory would have to be repeatedly adapted to the circumstances. Numbers noted as sequences are plain unwieldy (cf. [19], p. 15 f.).

Cantor regarded many sets as of the same cardinal number. His distinction between merely countable and uncountable sets is not differentiated enough. If instead bijections are correctly considered, this gives a different picture. The fuss made so far over ordinal and cardinal numbers is omitted because there is an easier way: the *consistent* extension of the real numbers.

Definition: Successively adding 1 to 0 gives the *set of natural numbers* $\mathbb{N} := \mathbb{N}^* \cup \{0\}$. Numbers from $\underline{\mathbb{R}}^*$ are called *purely complex*. Excluding all composite numbers from $\mathbb{N}_{\geq 2}$ gives the set of *prime numbers* $\mathbb{P}$. The set of *integers* $\mathbb{Z}$ is given by $\mathbb{Z} := \mathbb{N} \cup -\mathbb{N}^*$ and the set of *real numbers* $\mathbb{R}$ by the set of fractions with numerator from $\mathbb{Z}$ and denominator from $\mathbb{N}^*$. The *minimal distance* to 0 is ι. $\mathbb{R}$ includes both the conventionally hyperreal and surreal numbers. The numbers ν, ω and $1/\iota$ have the form 2^n for $n \in \mathbb{N}$.$\triangle$

Definition: Let $\mathbb{B} = \{0, 1\}$ the *Boolean domain*. *Decrement* and *increment* of $a \in \mathbb{C}$ are given by $\acute{a} := a - 1$ and $\grave{a} := a + 1$ and the reciprocal of $u \in \mathbb{C}^*$ is $\tilde{u} := 1/u$. They are pronounced "a dec", "a inc" and "turn uc'. Furthermore, $\hat{a} := 2a, \check{a} := a/2$ and $a\text{-} := -a$ are pronounced "a hat", "half a" and "a neg". The logarithm of $a \in {}^{\omega}\mathbb{C} \setminus {}^{\omega}\mathbb{R}_{\leq 0}$ and base $b \in {}^{\nu}\mathbb{R}_{>0}$ is denoted by ${}_{b}a$ and pronounced "b log a". If $\sigma \in \{\nu, \omega\}$ and $S^* = S^{*-1}$, let ${}^{\sigma}S := S \cap [-\sigma, \sigma]$ for $S \subseteq \mathbb{R}, {}^{\sigma}\mathbb{C} := {}^{\sigma}\mathbb{R} + {}^{\sigma}\underline{\mathbb{R}} \subset \mathbb{C}$ and ${}^{\sigma}S := S \cap {}^{\sigma}\mathbb{C}$ for $S \subseteq \mathbb{C}$.$\triangle$

Definition: Concerning all $s \in {}^{\sigma}_{\tau}S \subset \mathbb{C}$, the maximum value of their numerators is specified by σ and that of their denominators by τ. The sets $\mathbb{M}_{\mathbb{R}} := {}^{\omega}\mathbb{R} \setminus {}^{\nu}\mathbb{R}$ and $\mathbb{M}_{\mathbb{C}} := \mathbb{M}_{\mathbb{R}} + \underline{\mathbb{M}_{\mathbb{R}}}$ result in the *mid-finite* numbers. For real sets, subscript $_{\mathbb{R}}$ may be omitted in the notation. Two sets are *equal* if and only if they contain the same elements (extensionality). The set Y is called *union* of the set X if it contains exactly the elements of the elements of X as elements. Let $\mathcal{P}(X) := \{Y : Y \subseteq X\}$ be the *power set* of the set X.$\triangle$

Remark: For subsets of S, this holds analogously. For $T := {}^{\sigma}S \notin \{\emptyset, \{0\}\}$, it holds that $T + T \nsubseteq T$ or even $T \cdot T \nsubseteq T$. Sets with leading $^{\omega}$ only contain non-infinite elements. Sets ${}^{\nu}S$ are related to the conventional sets S. Mid-finite sets close the gap between finite and infinite ones.

Extreme theorem: Every linear order has exactly both a maximum and a minimum element, since all other constellations lead to a contradiction to totality or to that mentioned above.$\square$

Remark: Only the unique construction of a non-finite set allows to determine the number of its elements and to relate it to ${}^{\omega}\mathbb{N}$ as basis thanks to its simple construction. If there are multiple possible constructions, the most plausible one is selected. The relation $\in$ is irreflexive and asymmetric, whereas $\subseteq$ is partial order.

Definition: The sum $p(z) = \boldsymbol{+}_{k=0}^{\acute{m}} a_k z^k$ for $z \in \mathbb{C}$ and $m \in \mathbb{N}^*$ is called an *m-polynomial*, if the number of *coefficients* with $a_k \in {}^{\nu}\mathbb{Z}$ or $a_k \in {}^{\omega}\mathbb{Z}$ where $k \in \mathbb{N}_{<m}$ and $a_k \neq 0$ is finite, else *m-series*. Then $\deg(p) := \acute{m}$ for $a_k \neq 0$ is called the *degree* of the polynomial or series p. For the *zero polynomial* $p = 0$, $\deg(p) := -1$ holds. The $z \in \mathbb{C}$ setting $p(z) = 0$ are called *zeros* and are each an *m-algebraic number (AN)*. Alternating sums begin with $\boldsymbol{\pm}$ negating the second summand whereas $\boldsymbol{\mp}$ negates the first one.△

Definition: Let $\min |p(z)| = |p(\alpha)|$ for the minimal polynomial $p(z)$ of $\alpha \in \mathbb{C}$ where $0 \neq p(\alpha) \approx 0$. The corresponding sets are denoted $A := {}^{m}\mathbb{A}_{\mathbb{R}}$ in the real case and ${}^{m}\mathbb{R}_{\mathbb{C}} \supset {}^{m}\mathbb{A}_{\mathbb{C}} \supset {}^{m}_{m}\mathbb{R}_{\mathbb{C}} \supset {}^{m}\mathbb{Z}_{\mathbb{C}} \supset {}^{m}\mathbb{N}_{\mathbb{C}}$ in the complex one. For $a_{\deg(p)} = 1$, *m-algebraic integers* are given. The notation for m-ANs is $(m, a_{\acute{k}}, a_{k-2}, ..., a_1, a_0; r, i; \#n, \&q; v, p)_s$. All k-minimal polynomials have the sign $<$ as specification s, all k-minimal series have $>$.

Definition: For $r \in {}^{\nu}\mathbb{N}^*(-{}^{\nu}\mathbb{N}^*)$ exists the r-th largest ($|r|$-th smallest) zero with real part > 0 (< 0), where $r = 0, i \in {}^{\nu}\mathbb{N}^*(-{}^{\nu}\mathbb{N}^*)$ denotes a non-real zero with the i-th largest ($|i|$-th smallest) imaginary part > 0 (< 0), and the other ANs have analogous notations. The value $\#n$ gives the quantity $n \in {}^{\nu}\mathbb{N}^*$ of zeros. When at least one a_j is taken as a variable, $\&q$ gives the number $q \in {}^{\nu}\mathbb{N}$ of repeated zeros.△

Remark: Here r takes precedence over i and $r = i = a_0 = 0$ represents the number 0. The numerical value v has the precision p. Not distinguishing between repeated zeros allows the zeros of k-polynomials or k-series with integer coefficients to be endowed with a strict total ordering. The information $r, i, \#n, \&q, v, p$, and s may optionally be omitted as e.g. for fractions. The $(\nu + 2)$-tuple $(0, ..., 0, a_{\acute{k}}, ..., a_0; r, i)_<$ where every $a_j \in {}^{\nu}\mathbb{N}$ gives a strict lexical well-ordering of the ANs.

Examples: The numbers $(\nu; 1, 0, 0, 0, -1)_<$ are $1, -1, \underline{1}$, and $-\underline{1}$. The golden ratio $\Phi := \check{1} + 5^{\tilde{2}}/2$ may be written as $(\nu; 1, -1, -1; 1, 0; 1.618033, 10^{-6})_<$. Precision $p = 10^{-\omega}$ implies $\tilde{9} \neq 0.\overline{1} = 0.1...1$, since $9 \times 0.\overline{1} = 0.9...9 = 1 - 10^{-\omega} \neq 1$. Thus $0.\overline{1}$ is not ω-algebraic and may be written as $(\omega, 9 \times 10^{\omega}, 1 - 10^{\omega})$.

Remark: Let $m \in {}^{\nu}\mathbb{N}$ be the maximum polynomial degree and $n \in {}^{\nu}\mathbb{N}$ the maximum absolute value that the integer coefficients a_k can take of the polynomials $a_m x^m + a_{\acute{m}} x^{\acute{m}} + ... + a_1 x + a_0$ with $k \in {}^{\nu}\mathbb{N}_{\leq m}$. This makes sense due to the symmetry of the a_k. The number of ANs is the number of zeros of the normalised irreducible polynomials specified by the conditions: greatest common divisor gcd of the coefficients is equal to 1, $a_m > 0$, and $a_0 \neq 0$.

Inclusion theorem: Every set $S \neq \emptyset$ falsifies $S \in S, \mathcal{P}(S) \subseteq S$ and any *isomorphism* $T \cong S$ for $T \subset S$.

Proof: Every set differs from its elements, since it comprises the latter. Thus $\emptyset \neq \{\emptyset\}$. Its relative complement shows the elements with missing partner element for the bijection.□

Conclusion: Especially, this contradicts Dedekind-infiniteness and Cantor's first diagonal argument (see [3], p. 80 f.), since $\mathbb{N}$ is a proper subset of $\mathbb{R}$. Putting numbers excluded from the beginning to the end refutes the second one. The same is true for the Cantor, the Russell-Zermelo and the Banach-Tarski paradox. A translation of an infinite set always departs from the original set as the successor function $s : {}^{\omega}\mathbb{N} \to {}^{\omega}\mathbb{N}^* \cup \{\grave{\omega}\}$ shows for ${}^{\omega}\mathbb{N}$. This contradicts Hilbert's hotel.

Definition: *Cycle freedom* denotes the absence of cyclic sequences of sets, each containing one as an element in the previous one. *Replaceability* denotes the possible transition from a set X by uniquely replacing each element of X by an arbitrary set. If set Y contains exactly one element from each element of X (postulation of selectability), it is called selection of pairwise disjoint nonempty sets from a set X.△

Lemma: Because of $\tilde{\nu}m \leq 1 \leq a$ for all $m \in {}^{\nu}\mathbb{N}$ and $a \in {}^{\nu}\mathbb{R}_{\geq 1}$, the Archimedean axiom is invalid.□

Archimedes' theorem: There exists $m \in {}^{\nu}\mathbb{N}$ such that $a < bm$ if and only if $a < b\nu$ whenever $a > b$ for $a, b \in \mathbb{R}_{>0}$, since $\nu = \max {}^{\nu}\mathbb{N}$ holds.□

Claim: The Cantor polynomial $P(m, n) := \tilde{2}((m + n)^2 + 3m + n)$ bijectively maps ${}^{\omega}\mathbb{N}^2$ to ${}^{\omega}\mathbb{N}$.

Refutation: It holds that $P(\omega, \omega) = \hat{\omega}\grave{\omega} > \omega = \max {}^{\omega}\mathbb{N}$.□

Remark: Similarly, the Fueter-Pólya conjecture is refuted. If the set ${}^{\omega}\mathbb{N}^2$ is replaced by $\{(m, n) \in {}^{\omega}\mathbb{N}^2 : m + n \leq k \in {}^{\omega}\mathbb{N}\}$ for $k(k + 3) = \acute{\omega}$, the claim holds. Since infinitely many sets exist whose number of elements lie between $|{}^{\nu}\mathbb{N}|$ and $|{}^{\nu}\mathbb{R}|$, the continuum hypothesis is also wrong: By mapping too many (infinite) sets onto each other, they cannot be longer told apart.

Foundation theorem: Putting $X := \{x_m : x_0 := \{\emptyset\}, x_{\omega} := \{x_1\}$ and $x_{\acute{n}} := \{x_n\}$ for $m \in {}^{\omega}\mathbb{N}$ and $n \in {}^{\omega}\mathbb{N}_{\geq 2}\}$, only postulating the axiom of foundation guarantees cycle freedom that every nonempty subset $X \subseteq Y$ contains an element x_0 such that X and x_0 are disjoint.□

Remark: Setting $x_{\omega} := \{x_0\}$ instead of $x_{\omega} := \{x_1\}$, X becomes an infinite chain. All above definitions determine the set theory represented here, which does not require proper classes (cf. [3], p. 117 f.) since only meaningful sets are considered. The existing set $\mathbb{W}$ of all worlds cannot be changed, because otherwise it would contradict completeness. However, this does not mean that changes in the worlds do not exist or that everything is determined. It helps to look at time spatially.

Definition: Let $\mathbb{U}$ the *fuzzy set*, $\mathbb{Q}$ the set of *quaternions*, $\mathbb{O}$ the set of *octonions* and $\mathbb{S}$ the set of *sedenions*. Let $\mathbb{K} \in \{\mathbb{C}, \mathbb{R}\}$. Two different points x and y in a subset $M \subseteq \mathbb{K}^n$ where $n \in \mathbb{N}^*$ are said to be *neighbours* if $||x-y|| \leq \max\{||x-z||, ||y-z||\}$ holds for all points $z \in M$, where $||\cdot||$ denotes the Euclidean norm. If all neighbouring points of $M \subseteq \mathbb{K}^n$ have the minimal distance $h \in \mathbb{R}_{>0}$, M is said to be *h-homogeneous* and is denoted by *h-M*. The case $h = \iota$ makes M *gapless*. The isomorphism $\mathbb{C}^n \cong \mathbb{R}^{\acute{n}}$ holds.△

Definition: Let $a\!\upharpoonright_{f(,w)}$ for $a \in \mathbb{R}^n_{\geq 0}$ the f-mean of $a_1, \ldots, a_n$ with function f (maybe with weight vector $w \in [0, 1]^n$) or Hölder mean for level $r \in {}^{\nu}\mathbb{R}^*$ instead of f. Read $\upharpoonright$ as "mean". Let $u\urcorner_n := (u, ..., u)^{\top} \in {}^{\omega}\mathbb{K}^n$. Read $u\urcorner_n$ as "u rep n". A subset $M \subseteq \mathbb{K}^n$ is called *dense* in $\mathbb{K}^n$ if there is a point $y \in M$ for every $x \in \mathbb{K}^n$ with $||x - y|| = \iota$. For the point-symmetric $\dot{\mathbb{K}}^n$ related to $\mathbb{K}^n$, an *n-cube* with edge length σ is given by the product ${}^{\sigma}\mathbb{R}^n := \bigtimes_{k=1}^{n} {}^{\sigma}\mathbb{R}$ and an *n-ball* by ${}^{\sigma}\dot{\mathbb{R}}^n$ with radius σ.△

Remark: To h-homogenise a set, move h away from the origin in each dimension. Round elements in between to the nearest h-homogeneous elements. The maximum number of leading and also fractional digits of elements of $\tilde{\nu}$-${}^{\nu}\mathbb{R}$, $\tilde{\omega}$-${}^{\omega}\mathbb{R}$ and (ι-) $\mathbb{R}$ is given by the logarithms to base 2 (see Nonstandard Analysis) ${}_2\nu, {}_2\omega$ or ${}_2\tilde{\iota}$. The following theorem solves the Erdős-Ulam problem and reveals irrational numbers to be non-existent, and $\mathbb{N}, \mathbb{Z}, \mathbb{R}$ as well as $\mathbb{C}$ to be h-homogeneous.

Fundamental theorem of set theory: The set $\mathbb{R}$ is a maximal, linearly ordered, closed, continuous and ι-homogeneous field giving $|\mathbb{R}| = 2\tilde{\iota}^2 + 1$, since the distance ι guarantees continuity, although the exact value of ι is undetermined, and $\mathbb{R}$ can be "glued together" by identifying $\min \mathbb{R}$ and $\max \mathbb{R}$.□

Remark: Glue subsets of $\mathbb{K}^n$ analogously, which is then closed for certain calculations. Since reconstructing real numbers is unambiguous with necessarily periodic fractional expansion, h-homogeneity does not restrict significantly. A complete homogeneity (i. e. in all directions) of higher-dimensional real spaces is unfortunately impossible, as a simple consideration of circles or balls shows – also when one-dimensional spaces are rolled up.

Conclusion: A linear order can be defined for every non-empty set.□

Remark: Unlike the conventional way, immediate predecessors and successors can be specified. Vividly, this is a threading of pearls into a chain where the number of dimensions does not matter. Individual points from $\mathbb{R}^n$ can be linearly ordered with $n \in {}^{\omega}\mathbb{N}^*$, for example, dimension by dimension: The total chain length is then $\iota|\mathbb{R}|^n$. Another linear order is the spiral order starting at $0\urcorner_n$ according to the Euclidean norm for individual points.

Remark: Only malicious gossip has it that is wrong what is not trivially true in Cantor's set theory. What is certain, however, is that much of the latter cannot be maintained so far, but still contains enough interesting ideas to be pursued. A more detailed evaluation is omitted here for reasons of fairness.

3 Topology

In the following section, Set Theory is presupposed with the Euclidean norm $\|\cdot\|$.

Definition: A family of sets $\Upsilon \subseteq \mathcal{P}(X)$ is called *topology* on $X \subseteq R$ if every intersection and union of sets of Υ belongs apart from $\emptyset$ and X to Υ. The pair (X, Υ) is called *topological space*. If $\Upsilon = \mathcal{P}(X)$, the topology is called *discrete*. A set $B \subseteq \Upsilon$ is called a *base* of Υ if every set of Υ can be written as union of any number of sets of B. Every irreflexive relation $N \subseteq A^2$ founds a *neighbourhood relation (NR)* in $A \subseteq X$ for the underlying set X.△

Definition: If $(a, b) \in N$, a is called *neighbour* of or *neighbouring* to b. Particularly, an element $x \in A \subseteq X$ is called neighbour of an element $y \in A$, where $x \neq y$ if for all $z \in X$ and a mapping $d : X^2 \to \mathbb{R}_{\geq 0}$ holds: (1) $d(x, y) \leq \max\{\min\{d(x, z), d(z, x)\}, \min\{d(y, z), d(z, y)\}\}$ and (2) $d(z, z) = 0$. Here d is called *neighbourhood metric*. Let $P = R \cup V$ be the set of all *points* partitioned into *actual* points R and *virtual* points V for $R, V \neq \emptyset = R \cap V$. When R or V is clear from context, it can be omitted.△

Definition: The set $A^c := R \setminus A$, where $A \subseteq R$, is called *complement* of A in R. A^c can be called the *exterior* of A. All points of V (A) that have a neighbour in R $(A^c \cup V)$ form the *(inner) boundary* ∂V (∂A) of V (A). Here ^{c} takes precedence over ∂. When ∂ is applied successively beyond that, the argument is assumed to be without complement. The set $A^\circ := A \setminus \partial A$ is called the *interior* of A. The boundary (of width $b \in \mathbb{R}_{>0}$) of A is given by the set $\partial^b A := \{x \in A : d(x, y) \leq b, y \in A^c\}$ and is called *annulus* for disks.△

Definition: A set $S \subseteq R$ (V) is said to be *connected* if there is for every partition of S into $Y \cup Z$ such that $Y, Z \neq \emptyset = Y \cap Z$: $\partial Y^c \cap \partial Z \neq \emptyset \neq \partial Z^c \cap \partial Y$. $S \subseteq R$ is moreover said to be *simply connected* if holds: Both $\partial Y^c \cap \partial Z \cup \partial Z^c \cap \partial Y$ for every partition into connected Y and Z and $S^c \cup (\partial)V$ for S^c as complement of S in R are connected for a connected $(\partial)V$. Let P and R be simply connected. Every $U \subseteq R$ is called *neighbourhood* of $x \in R$ if $x \in U^\circ$.△

Definition: If $\emptyset \neq \mathbb{D} \subseteq (X, \Upsilon)$ holds, a connected $\mathbb{D}$ is called *domain*. The set of h-$S \subseteq \mathbb{R}^m$ for $m \in \mathbb{N}^*$ is *n-dimensional*, where $m \geq n \in \mathbb{N}^*$, if and only if it contains at least one n-cube with edge length $h \in \mathbb{R}_{>0}$ and maximum n. Let ${}^1\dot{\mathbb{R}}^n$ the *unit ball* with the special case *unit disk* ${}^1\dot{\mathbb{R}}^2$. *Midpoints a* of n-balls and n-cubes may be denoted in brackets as (a) after those. A set has dimension ${}_\iota n$ if its elements consist of n maximal cubes of edge length ι.△

Examples: The base for $\mathbb{N}, \mathbb{Z}, \mathbb{A}_{\mathbb{R}}, \mathbb{A}_{\mathbb{C}}, \mathbb{R}$ and $\mathbb{C}$ is precisely each related discrete topology. The boundary of every n-ball with $n \geq 2$ is only connected, itself it is simply connected. For $n = 1$ both are not connected. For $n \geq 2$ and $r \in \mathbb{R}_{>0}$, the n-Torus ${}^r\mathbb{T}^n := \left(\partial^r \dot{\mathbb{R}}\right)^n$ is only connected.

Theorem: For $n \in \mathbb{N}_{\geq 2}$, no finite decomposition of an n-ball can be reassembled giving an n-cube, since finitely many convex boundaries cannot have the same order of concave counterparts.□

Theorem: Traversing up to nine cells of a tic-tac-toe grid renders the singleton the only connected set with fixed point property (invalid hairy ball theorem).□

Definition: A function between two topological spaces is said to be *continuous* if for every point that can be mapped holds: for every neighbourhood of the image of this point there is a neighbourhood of the point whose image lies completely in the neighbourhood of the image of this point.△

Remark: The suggestive terms compactness and countability (possibly misleading) are not used since they are not given for infinite sets like $\mathbb{R}$ and $\mathbb{N}$ resp. For all contracting deformations, points are to be removed from the target set as the contraction specifies. Only then the (generalised) Poincaré conjecture holds.

Remark: An *exact* topology is represented, in which different infinitesimal distances establish an infinitesimal-discrete structure that can model Weyl and higher symmetries. It differs from standard topologies by making separation axioms obsolete and by not being Hausdorff. It can, for example, mitigate the wave-particle duality by using a unified model and nonstandard analytic methods to reflect reality.

4 Nonstandard Analysis

Preliminary remarks: In the following section, the definitions from set theory and topology are used. Since the mapping concept requires replacing every element not in the image set by the neighbouring element in the target set, the following statements are accurate to a maximum of ι and are thus no longer unambiguous. The following can also be easily generalised to other sets and norms.

From Norm to Differential

Definition: Let $\emptyset \neq A \subseteq {}^{(\omega)}\mathbb{K}^n$ and usually $m, n \in {}^{\omega}\mathbb{N}^*$. The function $\|\cdot\| : \mathbb{V} \to {}^{(\omega)}\mathbb{R}_{\geq 0}$ where $\mathbb{V}$ is a vector space over ${}^{(\omega)}\mathbb{K}$ is called a *norm*, if for all $x, y \in \mathbb{V}$ and $\lambda \in {}^{(\omega)}\mathbb{K}$, it holds that: $\|x\| = 0 \Rightarrow x = 0$ (*definiteness*), $\|\lambda x\| = |\lambda|\,\|x\|$ (*homogeneity*), and $\|x + y\| \leq \|x\| + \|y\|$ (*triangle inequality*). The *dimension* $\dim \mathbb{V}$ represents the maximal number of linearly independent vectors. The norms $\|\cdot\|_a$ and $\|\cdot\|_b$ are said to be *equivalent* if there exist $s, t \in [\tilde{\nu}, \nu]$ such that $s\|x\|_b \leq \|x\|_a \leq t\|x\|_b$ for all $x \in \mathbb{V}$. Let N be the set of all norms in $\mathbb{V}$.△

Theorem: Norms are equivalent if and only if $\|x\|_a/\|x\|_b \in [\tilde{\nu}, \nu]$ for all $\|\cdot\|_a, \|\cdot\|_b \in N$ and all $x \in \mathbb{V}^*$ by putting $s := \min\{\|x\|_a/\|x\|_b : x \in \mathbb{V}^*\}$ and $t := \max\{\|x\|_a/\|x\|_b : x \in \mathbb{V}^*\}$.□

Minimality theorem: For every $r \in \mathbb{R}$ with $n \in \mathbb{N}$ positions after the floating point, the geometric series (GS) yields a unique b-adic expansion and $\min\{b \in \mathbb{R}_{>1}\} = 2$, where $(1 - b^{-n})/(1 - b^{-1}) < 2$ holds.□

Remark: For $b \in]1, 2[$, the uniqueness is lost and two digits (0 and 1), intended for the dual representation, must be used: *One* digit only makes sense for base $b = 1$.

Definition: For the *characteristic function* χ, let $\chi_A(a) := 1$ for $a \in A$ and $\chi_A(a) := 0$ for $a \notin A$. Let $A_\infty := A \cup \{\pm\infty\}$ for $A \subseteq \mathbb{R}$ and $\infty \gg \tilde{\iota}^2$ as scalable constant. Furthermore, let $\operatorname{sgn}(z) := \tilde{z}|z|\chi_{\mathbb{C}^*}(z)$ for complex z and $\operatorname{sgn}(x) := \tilde{x}|x|\chi_{\mathbb{R}^*_\infty}(x)$ for real x. The area of ${}^1\dot{\mathbb{R}}^2$ gives *pi* π. *Euler's number* ϵ (read briefly as "eps") is defined as the solution of $x^{\underline{\pi}} = -1$. Then the *logarithm function* ln is given by $\epsilon^{\ln z} = z$ and the *power function* by $z^s = \epsilon^{s \ln z}$ for $s, z \in \mathbb{C}$. This allows giving a formal definition of *exponentiation*.△

Remark: If ± 0 is replaced by $\pm\widetilde{\infty}$, calculations become unique and consistent. The definition of ϵ is $O(\tilde{\nu})$ larger than that by $(1 + \tilde{\nu})^\nu$ (calculate with approximations!): The exponential series justifies the former when being exactly differentiated with as many terms as possible.

Definition: The function $\mu_h : A \to \mathbb{R}_{\geq 0}$ where $A \subseteq {}^{(\omega)}\mathbb{C}^n$ is an m-dimensional set with $h \in \mathbb{R}_{>0}$ less than or equal the minimal distance of the points in $A, m \in {}^{\omega}\mathbb{N}_{\leq \hat{n}}$, $\mu_h(A) := |A|h^m$ and $\mu_h(\emptyset) = |\emptyset| = 0$ is called the *exact h-measure* of A being *h-measurable*. Let the *exact standard measure* μ_ι (ι may be omitted).△

Remark: Answering positively the measure problem, the union A of pairwise disjoint h-homogeneous sets A_i for $i \in I \subseteq \mathbb{N}$ clearly additively and uniquely results in $\mu_h(A) = \sum_{i \in I} \mu_h(A_i)$. Its strict monotony follows for h-homogeneous sets $A_1, A_2 \subseteq {}^{(\omega)}\mathbb{K}^n$ satisfying $A_1 \subset A_2$ from $\mu_h(A_1) < \mu_h(A_2)$. If h is not equal on all considered sets A_i, the minimum of all h is chosen and the homogenisation follows as described in Set Theory. In the following, let $\|\cdot\|$ the Euclidean norm.

Examples: Consider the set $A \subset [0, 1[^n$ of points, whose least significant bit is 1 (0) in all $n \in {}^{\omega}\mathbb{N}^*$ coordinates. Then $\mu_\iota(A) = \tilde{2}^n$. Since A is an infinite and conventionally uncountable union of individual points without the neighbouring points of $[0, 1[^n$ in A, and these points are Lebesgue null sets, A is not Lebesgue measurable, however it is exactly measurable. Domains from ${}^{(\omega)}\mathbb{K}^n$ that are more densely pushed together have no smaller (larger) intersection (union) than previously.

Remark: The exact h-measure is optimal, since it only considers the NRs of points, i.e. in the extreme case distances of points parallel to the coordinate axes. Concepts such as σ-algebras and null sets are dispensable, since the empty set $\emptyset$ is null set enough.

Definition: Neighbouring points in A are described by the irreflexive symmetric *NR* $B \subseteq A^2$. The function $\gamma : C \to A \subseteq \mathbb{C}^n$, where $C \subseteq \mathbb{R}$ is h-homogeneous and h is infinitesimal, is called a *path* if $\|\gamma(x) - \gamma(y)\|$ is infinitesimal for all neighbouring points $x, y \in C$ and $(\gamma(x), \gamma(y)) \in B$. Let $z_0 \in A \subseteq \mathbb{K}^n$ and $f : A \to {}^{(\nu)}\mathbb{K}^m$. NRs are systematically written as (predecessor, successor) with the notation $(z_0, \vec{z}_0)$ or $(\overleftarrow{z}_0, z_0)$ pronouncing $\rightharpoonup$ as "post" and $\leftharpoonup$ as "pre". Clear A or B remain unmentioned.△

Definition: If $\|f(\vec{z}_0) - f(z_0)\| < \alpha$ for infinitesimal $\alpha \in {}^{(\omega)}\mathbb{R}_{>0}$, f is defined *α-successor-continuous* in z_0 in the direction $\vec{z}_0$. If the exact modulus of α does not matter, α may be omitted in the notation. If f is α-successor-continuous for all z_0 and $\vec{z}_0$, it simply is called α-continuous. It holds that α is the *degree* of continuity. If the inequality only holds for $\alpha = \tilde{\nu}$, f simply is called (successor-) continuous.△

Example: The function $f : \mathbb{R} \to \{\pm 1\}$ with $f(x) = \underline{1}^{\tilde{\iota}\hat{x}}$ is nowhere successor-continuous on $\mathbb{R}$, but its modulus is (cf. Number Theory). Here, $\tilde{\iota}x$ is an integer since $\mathbb{R}$ is ι-homogeneous. Setting $f(x) = 1$ for finite fractions x and $= -1$ otherwise, then $f(x)$ is partially ι-successor-continuous on infinite fractions, unlike the conventional notion of continuity.

Example of a Peano curve (cf. [30], p. 188): Consider the even, periodic function $g : {}^{\omega}\mathbb{R} \to {}^{\omega}\mathbb{R}$ with period 2 in $I := [0, 1]$ given by $g(s) = \chi_{[1,\check{3}]}(\tilde{s}) + \chi_{]\check{3},3[}(\tilde{s})(3s - 1)$. Now let $\phi : I \to {}^{\omega}\mathbb{R}^2$ be defined by

$$\phi(s) = \left(+_{n=0}^{\omega}\tilde{2}^{\hat{n}}g(4^{\hat{n}}s), +_{n=0}^{\omega}\tilde{2}^{\hat{n}}g(4^{\hat{n}+1}s)\right).$$

The function ϕ is at least continuous since the sums are ultimately locally linear functions in s. It would however be an error to believe that I can be bijectively mapped onto I^2 in this way: the powers of four in g, and the values 0 and 1 taken by g in two sub-intervals thin out I^2 so much that a bijection is clearly impossible. Restricting the proof only to finite fractions is simply insufficient.

Definition: For $f : A \to {}^{(\omega)}\mathbb{K}^m$, $\downarrow f(\vec{z}) := f(\vec{z}) - f(z)$ is called to be *successor-differential* of f in the direction $\vec{z}$ for $z \in A$. If $\dim A = n$, then $\downarrow f(\vec{z})$ stands for a successor-derivative in every variable. Mixed differentials are specified by several arrows. Unimportant ones are omitted. If $f(z) = z$, then $\downarrow\vec{z}$ can be written instead of $\downarrow f(\vec{z})$. Read $\downarrow$ as "down". If $|f(\vec{x}) - f(x)| > \tilde{\omega}$ holds for x of $f : A \subseteq {}^{\omega}\mathbb{R} \to {}^{\omega}\mathbb{R}$, x is called a *jump discontinuity*. Choose $g(\vec{x}) = \vec{g}(x)$ in the following text.△

Definition: If the modulus of the successor-differential of f in the direction $\vec{z}$ at $z \in A$ is smaller than α and infinitesimal, then f is also rated as α-successor-continuous there. A function $f : A \subseteq {}^{(\omega)}\mathbb{K}^n \to {}^{\omega}\mathbb{R}$ is said to be *convex (concave)* (written $f \in Con(A)$) if the line segment between any two points on the graph of the function lies above (below) or on the graph. Let it *strictly* convex (concave) if "or on" can be omitted.△

Theorem: For $A \subseteq {}^{(\omega)}\mathbb{K}^n$, all $f \in Con(A)$ are α-successor-continuous and successor-differentiable.□

Definition: The m arithmetic means of all $f_k(\vec{z})$ of $f(z)$ give the *m averaged normed tangential normal vectors* of m unique hyperplanes, defining the mn continuous derivatives of the Jacobian matrix of f, which is not necessarily continuous. The hyperplanes are taken to pass through $f_k(\vec{z})$ and $f(z)$ translated towards 0. A quite simple linear programme (LP) (cf. Linear Programming) minimises the moduli of their coefficients.△

Definition: For the (maybe dropped) *control variable* $m \in \mathbb{N}^*_{\le n}$ and the *concatenation operator* $\complement$, the *derivative* in the direction $\vec{z}_k$ of $F : A \to {}^{(\omega)}\mathbb{K}$ at $z = \left(\complement_{(m=)1}^{n} z_m\right) := (z_1, ..., z_n) \in A \subseteq {}^{(\omega)}\mathbb{K}^n$ is given by

$${}^1F_m(\vec{z}) := \downarrow F(z)/\downarrow z_m = (F(z_1, ..., \vec{z}_m, ..., z_n) - F(z))/(\vec{z}_m - z_m).△$$

Definition: The derivative of a function $f : A \to {}^{(\omega)}\mathbb{R}$, where $A \subseteq {}^{(\omega)}\mathbb{R}$, is said to be 0 if and only if 0 lies in the interval given by the boundaries of the left and right exact derivatives or where f is discontinuous. Let $\text{num}(x) = p \in \mathbb{Z}$ the *numerator function* and $\text{den}(x) = |q| \in \mathbb{N}^*$ the *denominator function* of $x = p/q \in \mathbb{R}$ for coprime p and q (in short: $p \perp q$).

With this notation, if the function f satisfies $f = \left(\complement_1^n f_m\right) : A \to {}^{(\omega)}\mathbb{K}^n$ with $z \in A \subseteq {}^{(\omega)}\mathbb{K}^n$

$$f(\vec{z}) = \left(\frac{F(\vec{z}_1, \complement_2^n z_m) - F(\complement_1^n z_m)}{\vec{z}_1 - z_1}, ..., \frac{F(\complement_1^{\acute{n}} z_m, \vec{z}_n) - F(\complement_1^n z_m)}{\vec{z}_n - z_n}\right) = \left(\frac{\downarrow F_1(z)}{\downarrow z_1}, \; ... \;, \frac{\downarrow F_n(z)}{\downarrow z_n}\right),$$

then $f(\vec{z}) = \nabla F(\vec{z})$ with the *Nabla operator* ∇ is said to be *exact successor-derivative* ${}^1F(\vec{z})$ or the *exact successor-gradient* grad $F(\vec{z})$ of the function $F(z)$, which is called *exactly differentiable* at z in direction $\vec{z}$.

If this definition is satisfied for every $z \in A$, then F is said to be an *exactly differentiable antiderivative (AD)* of f. If all directions have the same value, *holomorphicity* is obtained ($A = {}^{\nu}\mathbb{C}$ and $n = 1$ make F holomorphic). On a domain $\mathbb{D}$, let $\mathcal{O}(\mathbb{D}) \subseteq \mathcal{C}(\mathbb{D}) \subseteq \mathbb{C}$ be the *ring of holomorphic resp. continuous functions*.△

Chain rule: From $x \in A \subseteq {}^{(\omega)}\mathbb{R}, g : A \to B \subseteq {}^{(\omega)}\mathbb{R}, f : B \to C \subseteq {}^{(\omega)}\mathbb{R}$, it follows that ${}^{1}f(g(x)) = {}^{1}f(g(x))\,{}^{1}g(x)$ due to

$${}^{1}f(g(x)) = \frac{f(g(\vec{x})) - f(g(x))}{g(\vec{x}) - g(x)}\,\frac{g(\vec{x}) - g(x)}{\vec{x} - x} = \frac{f(\vec{g}(x)) - f(g(x))}{\vec{g}(x) - g(x)}\,{}^{1}g(x) = {}^{1}f(g(x))\,{}^{1}g(x).\square$$

Product rule: Adding and subtracting $f(\vec{x})g(x)$ or $f(x)g(\vec{x})$ in the numerator yields

$${}^{1}(fg)(x) = {}^{1}f(x)g(x) + f(\vec{x})\,{}^{1}g(x) = {}^{1}f(x)g(\vec{x}) + f(x)\,{}^{1}g(x).\square$$

Quotient rule: The same for $f(x)g(x)$ and $f(\vec{x})g(\vec{x})$ yields for denominators $\neq 0$ of the following quotients

$${}^{1}\left(\tfrac{f}{g}\right)(x) = \tfrac{{}^{1}f(x)\,g(x) - f(x)\,{}^{1}g(x)}{g(x)\,g(\vec{x})} = \tfrac{{}^{1}f(x)\,g(\vec{x}) - f(\vec{x})\,{}^{1}g(x)}{g(x)\,g(\vec{x})}.\square$$

Remark: Arguments and function values must belong to a smaller level of infinity than $\tilde{\iota}$ forcing f and g to be sufficiently (α-) continuous at $x \in A$. I. e. α must be sufficiently small to allow $\vec{x}$ to be replaced by x. An analogous principle holds for infinitesimal arguments. The exact derivative of the reciprocal function reads ${}^{1}f^{-1}(y) = 1/{}^{1}f(x)$ from $y = f(x)$ and identity $x = f^{-1}(f(x))$ by the chain rule and the same precision. L'Hôpital's rule makes sense for (α-) continuous functions f and g, and follows for $f(v) = g(v) = 0$ where $v \in A$ and $g(\vec{v}) \neq 0$ from

$$f(\vec{v})/g(\vec{v}) = (f(\vec{v}) - f(v))/(g(\vec{v}) - g(v)) = {}^{1}f(v)/{}^{1}g(v).$$

Remark: If the exact derivative can be replaced by ${}^{1}F(\overleftrightarrow{v}) := F(\vec{v}) - F(\overleftarrow{v})/(\vec{v} - \overleftarrow{v})$ (numerator $\neq 0$), this has the advantage viewing ${}^{1}F(\overleftrightarrow{v})$ as the "tangent slope" at the point v, especially when F is α-continuous at v. This applies beyond that when $\vec{v} - v = v - \overleftarrow{v}$, and the combined derivatives both have the same sign. An extension to (conventional) complex numbers exists analogously.

Theorem, improving Froda's one: A monotone function $f : [a, b] \to {}^{\omega}\mathbb{R}$ has at most $2\omega^2 - 1$ jump discontinuities, since at most $2\omega^2$ jump discontinuities are possible between $-\omega$ and ω with a jump of $\widetilde{\omega}$ if the function does not decrease at non-discontinuities, like a step function.$\square$

Intermediate value theorem: Let $I := [\min f(x), \max f(x)]$ and $f : [a, b] \to {}^{(\omega)}\mathbb{R}$ α-continuous in $[a, b]$. Then $f(x)$ takes in I every value with precision $< \alpha$. If f is continuous in ${}^{\omega}\mathbb{R}$, it takes every value of ${}^{\nu}\mathbb{R}$ in I.

Proof: A gapless chain of overlapping α-environments exists in I where $f(x)$ is centre, since otherwise there would be a contradiction to the α-continuity of f. The second part of the claim follows from the fact that a deviation $|f(\vec{x}) - f(x)| < \tilde{\nu}$ in ${}^{\nu}\mathbb{R}$ falls below the conventional resolution maximally permitted.$\square$

Remark to the extreme value theorem: The continuous function $f(x) := \hat{\omega}\sin(\omega x)$ attains for $x \in [-1, 1]$ the minima $-\hat{\omega}$ and the maxima $\hat{\omega}$ as infinite values.

Example: The $\hat{\iota}$-continuous function $f : {}^{(\omega)}\mathbb{R} \to \{0, \iota\}$ defined by $f(x) := \check{\iota}(\underline{1}^{\hat{x}\tilde{\iota}} + 1)$ consists of only the local minima 0 and the local maxima ι, and has the left and right exact derivatives ± 1.

Examples: In ([4], p. 160), it holds that $r_1 = r_2 = 3$. For $q = \operatorname{den}(x)$ and $f(x) := -\chi_{{}^{\omega}_{\omega}\mathbb{R}}(x)\underline{1}^{\hat{q}}\tilde{q}\acute{q}$, it has $f : [0, 1] \to [\acute{\iota}, -\acute{\iota}]$ the two relative extrema $\pm\acute{\iota}$ (cf. ib. p. 24).

Definition: The direction $w := \vec{z}$ gives the *second derivative* of $f : A \to {}^{(\omega)}\mathbb{K}$ at $z \in A \subseteq {}^{(\omega)}\mathbb{K}$ by

$${}^{2}f(\vec{z}) := \tfrac{\downarrow^{2} f(\vec{z})}{(\downarrow\vec{z})^{2}} = \tfrac{f(\vec{w}) - \hat{f}(\vec{z}) + f(z)}{(\downarrow\vec{z})^{2}}.\triangle$$

Remark: Higher derivatives are defined analogously. Every number $m_n \in {}^{\omega}\mathbb{N}$ for $n \in {}^{\omega}\mathbb{N}^{*}$ of derivatives is written as an exponent after the n-th variable to be differentiated. The exponent to be specified in the numerator is the sum of all m_n. Regarding $1/(-1)! = 0$ implies then for g like f and $p \in {}^{\omega}\mathbb{N}^{*}$ the Leibniz product rule:

$${}^{p}(fg) = \textstyle\sum_{m+n=p}\binom{p}{m}\,{}^{m}f\,{}^{n}g.$$

Proof: For $p = 1$, the product rule mentioned above holds. Induction step from p to $\grave{p}$:

$$\begin{aligned}{}^{\grave{p}}(fg) &\underset{p}{=} \textstyle\sum_{m+1+n=\grave{p}}\left(\binom{p}{m} + \binom{p}{\grave{m}}\right){}^{\grave{m}}f\,{}^{n}g + \textstyle\sum_{m+1+n=\grave{p}}\binom{p}{m}\,{}^{m}f\,{}^{\grave{n}}g - \textstyle\sum_{m+1+n=\grave{p}}\binom{p}{\grave{m}}\,{}^{\grave{m}}f\,{}^{n}g\\ &= {}^{p}\left({}^{1}(fg)\right) \underset{1}{=} {}^{p}\left({}^{1}fg + f\,{}^{1}g\right) = {}^{p}\left({}^{1}fg\right) + {}^{p}\left(f\,{}^{1}g\right) = \textstyle\sum_{m+n=\grave{p}}\binom{\grave{p}}{m}\,{}^{m}f\,{}^{n}g.\square\end{aligned}$$

From Integral to Stirling Formula

Definition: Given $z \in A \subseteq {}^{(\omega)}\mathbb{K}^n$, it is $\uparrow_{z \in A} f(z) \downarrow z := \boldsymbol{+}_{z \in A} f(z)(\vec{z} - z)$ called the *exact successor integral* of the *vector field* $f = (f_1, ..., f_n) : A \rightarrow {}^{(\omega)}\mathbb{K}^n$ on A and $f(z)$ is said to be *exactly integrable*. If this requires removing at least one point from A, then the exact integral is called *improper*. Read $\uparrow$ as "up". For $\gamma : G = [a, b[\, \cap C \rightarrow A \subseteq {}^{(\omega)}\mathbb{K}^n, C \subseteq \mathbb{R}$ and $f = (f_1, ..., f_n) : A \rightarrow {}^{(\omega)}\mathbb{K}^n, \uparrow_\gamma f(\zeta) \downarrow \zeta = \uparrow_{s \in G} f(\gamma(s))\, {}^1\gamma(s) \downarrow s$ where $\downarrow s > 0, \vec{s} \in \,]a, b] \cap C$, choosing $\vec{\gamma}(s) = \gamma(\vec{s})$, since $\zeta = \gamma(s)$ and $\downarrow \zeta = \gamma(\vec{s}) - \gamma(s) = {}^1\gamma(s) \downarrow s$ (i.e. for $C = \mathbb{R}, B$ maximal in $\mathbb{C}^2$, and D maximal in $\mathbb{R}^2$), is called the *exact line integral (LI)* of the vector field f along the path γ. Improper exact LIs are founded analogously to exact integrals, except that only interval end points may be removed from G.△

Remark: The (linear) exact LI on ${}^{(\nu)}\mathbb{K}$ f does not need a continuous f, exists always and is usually consistent with the conventional LI. It is linear and monotone in the (conventional) (infinite) real case.

Laisant's theorem: For $c \in {}^{\omega}\mathbb{R}$, the product rule yields $\uparrow f(x) \downarrow x \big|_{f^{-1}(y)} = \uparrow y \frac{\downarrow x}{\downarrow y} \downarrow y = y f^{-1}(y) - \uparrow f^{-1}(y) \downarrow y + c.\square$

Definition: For all $x \in V$ of an h-homogeneous n-volume $V \subseteq [a_1, b_1] \times ... \times [a_n, b_n] \subseteq {}^{(\omega)}\mathbb{R}^n$ with $B = B_1 \times ... \times B_n, B_k \subseteq [a_k, b_k]^2$ and $|\downarrow x_k| = h$ for all $k \in \mathbb{N}^*_{\leq n}$ such that $f(x) := 0$ for all $x \in {}^{(\omega)}\mathbb{R}^n \setminus V$,

$$\uparrow_{x \in V} f(x) \downarrow x := \uparrow_{x \in V} f(x) \downarrow (x_1, ..., x_n) := \uparrow_{a_n}^{b_n} ... \uparrow_{a_1}^{b_1} f(x) \downarrow x_1 ... \downarrow x_n$$

is called the *exact volume integral* of the *volume integrable* function $f : {}^{(\omega)}\mathbb{R}^n \rightarrow {}^{(\omega)}\mathbb{R}$.△

Remark: The isomorphy of $\mathbb{C}$ and $\mathbb{R}^2$ provides something similar for the complex case and $\uparrow_{x \in V} 1 \downarrow x = \mu_h(V)$.

Example: Using the exact volume integral in contrast to the Lebesgue integral,

$$||f||_p := (\uparrow_{x \in V} ||f(x)||^p \downarrow x)^{\tilde{p}}$$

satisfies for arbitrary $f : {}^{(\omega)}\mathbb{R}^n \rightarrow {}^{(\omega)}\mathbb{R}^m$ and $p \in [1, \omega]$ all the properties of a norm, also definiteness.

Example: Let $[a, b[\, \cap h^{\omega}\mathbb{Z} \neq \emptyset$ be an h-homogeneous subset of $[a, b[^{\omega}\mathbb{R}$, and write $B \subseteq [a, b[\, \cap h^{\omega}\mathbb{Z} \times]a, b] \cap h^{\omega}\mathbb{Z}$. Now let T an AD of a not necessarily convergent Taylor series (TS) t on $[a, b[\, \cap h^{\omega}\mathbb{Z}$, and define $f(x) := t(x) + \varepsilon \underline{1}^{\hat{x}/h}$ for conventionally real x and $\varepsilon \geq \tilde{\nu}$. For $h = \tilde{\nu}$, f is nowhere continuous, and thus is conventionally nowhere differentiable or integrable on $[a, b[\, \cap h^{\omega}\mathbb{Z}$, but for all h holds

$${}^1 f(x) = {}^1 t(x) - \widetilde{\downarrow x} \hat{\varepsilon} \underline{1}^{\hat{x}/h}$$

and

$$\uparrow_{x \in [a, b[\, \cap h^{\omega}\mathbb{Z}} f(x) \downarrow x = T(b) - T(a) + \check{\varepsilon} \left(\underline{1}^{\hat{a}/h} - \underline{1}^{\hat{b}/h} \right).$$

Example: The conventionally non-measurable middle-thirds Cantor set $C_{\check{3}}$ has measure $\mu_\iota(C_{\check{3}}) = \check{3}^{-\omega}$. Consider the function $c : [0, 1] \rightarrow \{0, \check{3}^{\omega}\}$ defined by $c(x) = \check{3}^{\omega} \chi_{C_{\check{3}}}(x)$. Then

$$\uparrow_{x \in [0,1]} c(x) \downarrow x = \boldsymbol{+}_{x=0}^{1} c(x) \downarrow x = \check{3}^{\omega} \mu_\iota (C_{\check{3}}) = 1.$$

Definition: A *sequence* (a_k) with *members* a_k is a mapping from ${}^{(\omega)}\mathbb{Z}$ to ${}^{(\omega)}\mathbb{C}^m : k \mapsto a_k$. A *series* is a sequence (s_k) with $m \in {}^{(\omega)}\mathbb{Z}$, *radius of convergence* r and *partial sums* $s_k = \boldsymbol{+}_{j=m}^{k} a_j$. A sequence (a_k) with $k \in {}^{(\omega)}\mathbb{N}^*, a_k \in {}^{(\omega)}\mathbb{C}$ and $\alpha \in]0, \tilde{\nu}]$ is called *α-convergent* to $a \in {}^{(\omega)}\mathbb{C}$ if there exists $m \in {}^{(\omega)}\mathbb{N}^*_{\leq k}$ where $|a_k - a| < \alpha$ for all a_k such that $k - m$ is not too small. The set α-A of all such a is called *set of α-limit values* of (a_k). A uniquely determined representative of this set (e.g. the final value or mean value) is called the *α-limit value α-a*. For the case $a = 0$, the sequence is called a *zero sequence*. If the inequality only holds for $\alpha = \tilde{\nu}$, the α- is omitted. Usually, k will be chosen maximal and α minimal.

Example: The alternating harmonic series implies $\pm_{n=1}^{\omega} (\tilde{n} - \omega) = {}_{\epsilon}2$.

Remark: Conventional limit values are hardly more precise than $O(\widetilde{\omega})$. Their actual rationality or pure algebraicity is seldom regarded! To avoid the exclusive relevance of the largest index of each sequence (cf. [9], p. 144) the conventional definition requires the completion that infinitely many or almost all members of the sequence have an arbitrarily small distance from the limit value. Only finitely many may have a larger distance. Then only the monotone convergence is valid (cf. [9], p. 155).

Remark: The fundamental theorem of set theory makes the representation of each positive number by a determined, unique, infinite decimal fraction baseless (cf. p. 27 f.). Putting $\varepsilon := \iota$ any proof claiming that, for $\varepsilon \in {}^{(\omega)}\mathbb{R}_{>0}$ - especially for all $\varepsilon \in {}^{(\nu)}\mathbb{R}_{>0}$ - there exists a real number $\varepsilon\tilde{r}$ with real $r \in {}^{(\omega)}\mathbb{R}_{>1}$, is false. Otherwise, an infinite regression may occur. The $\varepsilon\delta$-definition of the limit value (it is questionable that δ exists, p. 235 f.) requires ε as a specific multiple of ι making that of continuity also true (see p. 215).

Remark: Consider for example the real function that doubles every real value but is not even uniformly continuous. Uniform continuity need not be considered, since in general $\delta := \iota$ and ε accordingly larger. If two function values do not satisfy the conditions, then the function is not continous at that point. Thus, continuity is equivalent to uniform continuity, by choosing the largest ε from all admissible infinitesimal values. Easily, continuity is equivalent to Hölder continuity.

Remark: Here infinite real constants may be allowed. The same is true for uniform convergence, since the maximum of the indices may be chosen such that each argument as the index satisfies everything, and $\acute{\omega}$ is sufficient in every case. Otherwise, pointwise convergence also fails. Thus, uniform convergence is equivalent to pointwise convergence, by choosing the largest of all admissible infinitesimal values.

Fubini's theorem: For $X, Y \subseteq {}^{(\omega)}\mathbb{K}$ and $f : X \times Y \to {}^{(\omega)}\mathbb{K}$, a reordering of integral sums shows

$$\uparrow_Y \uparrow_X f(x,\ y) \downarrow x \downarrow y = \uparrow_{X \times Y} f(x,\ y) \downarrow (x,\ y) = \uparrow_X \uparrow_Y f(x,\ y) \downarrow y \downarrow x.\square$$

Example: By the principle of latest substitution (see below), $r_{\pm}^2 := x^2 \pm y^2$ results for

$$\uparrow_{[a,\,b[\times[c,\,d[} \tilde{r}_+^4 r_-^2 \downarrow (x,\ y) = \uparrow_a^b \tilde{r}_+^2 y \downarrow x\Big|_c^d = -\uparrow_c^d \tilde{r}_+^2 x \downarrow y\Big|_a^b = \arctan \tfrac{d}{b} - \arctan \tfrac{c}{b} + \arctan \tfrac{d}{a} - \arctan \tfrac{c}{a}$$

in the (improper) integral

$$I(a,b) := \uparrow_{[a,\,b[^2} \tilde{r}_+^4 r_-^2 \downarrow (x,\ y) = \arctan \tfrac{b}{b} - \arctan \tfrac{a}{b} + \arctan \tfrac{b}{a} - \arctan \tfrac{a}{a} = \check{\pi} - \check{\pi} = 0 \qquad \text{instead of}$$

$$I(0,1) = \uparrow_0^1 \uparrow_0^1 \tilde{r}_+^4 r_-^2 \downarrow y \downarrow x = \uparrow_0^1 \tfrac{\downarrow x}{1+x^2} = \tfrac{\pi}{4} \neq -\tfrac{\pi}{4} = -\uparrow_0^1 \tfrac{\downarrow y}{1+y^2} = \uparrow_0^1 \uparrow_0^1 \tilde{r}_+^4 r_-^2 \downarrow x \downarrow y = I(0,1).$$

Exchange theorem: (Transfinite) induction shows the result of multiple derivatives of a function $f : A \to {}^{(\omega)}\mathbb{K}$ as independent of the order of differentiation, provided that variables are only evaluated and limits are only computed at the end, if applicable (*principle of latest substitution*).$\square$

Example: For $f : {}^{\omega}\mathbb{R}^2 \to {}^{\omega}\mathbb{R}$, $f(0,0) = 0$ and $f(x,y) = \tilde{r}_+^2 x y^3$ where $r_{\pm}^2 := x^2 \pm y^2$, it holds that

$$\tfrac{\downarrow^2 f}{\downarrow x \downarrow y} = \tilde{r}_+^6 (y^6 + 6x^2y^4 - 3x^4y^2) = \tfrac{\downarrow^2 f}{\downarrow y \downarrow x}$$

with value $\tilde{2}$ at the point $(0,0)$, even though having y for $x = 0$ on the left and 0 on the right for $y = 0$ in

$$\tfrac{\downarrow f}{\downarrow x} = -\tilde{r}_+^4 r_-^2 y^3 \neq \tilde{r}_+^4 (xy^4 + 3x^3y^2) = \tfrac{\downarrow f}{\downarrow y},$$

where differentiating with respect to the other variable gives on the left $1 \neq 0$ on the right.

First fundamental theorem of exact differential and integral calculus for LIs: The function $F(z) = \uparrow_\gamma f(\zeta) \downarrow \zeta$ where $\gamma : [d, x[\, \cap C \to A \subseteq {}^{(\omega)}\mathbb{K}, C \subseteq \mathbb{R}, f : A \to {}^{(\omega)}\mathbb{K}, d \in G = [a, b[\, \cap C$, and choosing $\vec{\gamma}(x) = \gamma(\vec{x})$ is exactly differentiable, and ${}^1F(z) = f(z)$ holds for all $x \in G$ and $z = \gamma(x)$.

Proof:
$$\begin{aligned} \downarrow F(z) &= \uparrow_{s \in [d,x] \cap C} f(\gamma(s))\, {}^1\gamma(s) \downarrow s - \uparrow_{s \in [d,x[\, \cap C} f(\gamma(s))\, {}^1\gamma(s) \downarrow s &&= \uparrow_x f(\gamma(s)) \tfrac{\gamma(\vec{s}) - \gamma(s)}{\vec{s} - s} \downarrow s \\ &= f(\gamma(x))\, {}^1\gamma(x) \downarrow x = \qquad f(\gamma(x))(\vec{\gamma}(x) - \gamma(x)) &&= f(z) \downarrow z.\square \end{aligned}$$

Second fundamental theorem of exact differential and integral calculus for LIs: Conditions above imply with $\gamma : G \to {}^{(\omega)}\mathbb{K}$ that

$$F(\gamma(b)) - F(\gamma(a)) = \uparrow_\gamma {}^1F(\zeta) \downarrow \zeta.$$

Proof:
$$\begin{aligned} F(\gamma(b)) - F(\gamma(a)) &= \mathop{\Large +}_{s \in G} F(\vec{\gamma}(s)) - F(\gamma(s)) &&= \mathop{\Large +}_{s \in G} {}^1F(\gamma(s))(\vec{\gamma}(s) - \gamma(s)) \\ &= \uparrow_{s \in G} {}^1F(\gamma(s))\, {}^1\gamma(s) \downarrow s &&= \uparrow_\gamma {}^1F(\zeta) \downarrow \zeta.\square \end{aligned}$$

Corollary: If f has an AD F on a closed path (CP) γ, it holds with the conditions above that $\uparrow_\gamma f(\zeta) \downarrow \zeta = 0.\square$

Integral formula: The last corollary shows that for $f : A \to {}^{(\omega)}\mathbb{C}$ and the CP $\gamma([a, b[) \subseteq A \to {}^{(\omega)}\mathbb{C}$ the equation (see below) $f(z) \operatorname{ind}_\gamma(z) = \widetilde{\underline{\hat{\pi}}} \uparrow_\gamma \widetilde{\zeta - z} f(\zeta) \downarrow \zeta$ holds, if and only if $g(\zeta) = \widetilde{\zeta - z}(f(\zeta) - f(z))$ implies that $\uparrow_\gamma g(\zeta) \downarrow \zeta = 0$. This is especially true if g has an AD on $\gamma([a, b[)$.□

Remark: The conventionally real case of both fundamental theorems may be established analogously. Given $u, v \in G, u \neq v$ and $\gamma(u) = \gamma(v)$, it may be the case that $\vec{\gamma}(u) \neq \vec{\gamma}(v)$.

Remark: In the first fundamental theorem, the derivative $\downarrow(F(z))/\downarrow z$ can be tightened to the arithmetic mean $\tilde{2}(f(z) + f(\vec{z}))$ resp. $f(\tilde{2}(z + \vec{z}))$, and similarly, in the second fundamental theorem, $F(\gamma(b)) - F(\gamma(a))$ can be tightened to $\tilde{2}(F(\gamma(b)) + F(\vec{\gamma}(b))) - \tilde{2}(F(\gamma(a)) + F(\vec{\gamma}(a)))$ resp. $F(\tilde{2}(\gamma(b) + \vec{\gamma}(b))) - F(\tilde{2}(\gamma(a) + \vec{\gamma}(a)))$. This yields approximately the original results when f and F are sufficiently α-continuous at the boundary.

Transformation theorem: If the Jacobian $D\varphi(x)$ exists, linear algebra shows for $f : \varphi(A) \to {}^{(\omega)}\mathbb{R}^n$ and $A \subseteq {}^{\omega}\mathbb{R}^n$ (cf. [17], p. 519):

$$\uparrow_{\varphi(A)} f(y) \downarrow y = \uparrow_A f(\varphi(x)) |\operatorname{eig}(D\varphi(x))| \downarrow x. \square$$

Leibniz integral rule: For $f : {}^{(\omega)}\mathbb{K}^{\grave{n}} \to {}^{(\omega)}\mathbb{K}, a, e : {}^{(\omega)}\mathbb{K}^n \to {}^{(\omega)}\mathbb{K}, \vec{x} := (s, x_2, ..., x_n)^\top$, and $s \in {}^{(\omega)}\mathbb{K} \setminus \{x_1\}$, choosing $\vec{a}(x) = a(\vec{x})$ and $\vec{e}(x) = e(\vec{x})$, it holds that

$$\frac{\downarrow}{\downarrow x_1}\left(\uparrow_{a(x)}^{e(x)} f(x, t) \downarrow t\right) = \uparrow_{a(x)}^{e(x)} \frac{\downarrow f(x,t)}{\downarrow x_1} \downarrow t + \frac{\downarrow e(x)}{\downarrow x_1} f(\vec{x}, e(x)) - \frac{\downarrow a(x)}{\downarrow x_1} f(\vec{x}, a(x)).$$

Proof:
$$\begin{aligned} \frac{\downarrow}{\downarrow x_1}\left(\uparrow_{a(x)}^{e(x)} f(x, t) \downarrow t\right) &= \left(\uparrow_{a(\vec{x})}^{e(\vec{x})} f(\vec{x}, t) \downarrow t - \uparrow_{a(x)}^{e(x)} f(x, t) \downarrow t\right) / \downarrow x_1 \\ &= \left(\uparrow_{a(x)}^{e(x)} (f(\vec{x}, t) - f(x, t)) \downarrow t + \uparrow_{e(x)}^{e(\vec{x})} f(\vec{x}, t) \downarrow t - \uparrow_{a(x)}^{a(\vec{x})} f(\vec{x}, t) \downarrow t\right) / \downarrow x_1 \\ &= \uparrow_{a(x)}^{e(x)} \frac{\downarrow f(x,t)}{\downarrow x_1} \downarrow t + \frac{\downarrow e(x)}{\downarrow x_1} f(\vec{x}, e(x)) - \frac{\downarrow a(x)}{\downarrow x_1} f(\vec{x}, a(x)). \square \end{aligned}$$

Remark: Complex integration allows a path whose start and end points are the limits of integration. If $\vec{a}(x) \neq a(\vec{x})$, then multiply the final summand by $(\vec{a}(x) - a(x))/(a(\vec{x}) - a(x))$. If $\vec{e}(x) \neq e(\vec{x})$, then multiply the penultimate summand by $(\vec{e}(x) - e(x))/(e(\vec{x}) - e(x))$. Let $n \in {}^{\omega}\mathbb{N}^*$ and $x \in [0, 1]$ in each case for the following examples (cf. [9], p. 540 - 543).

1. The sequence $f_n(x) = \sin(nx)/n^{\tilde{2}}$ does not tend to $f(x) = 0$ as $n \to \omega$, but instead to $f(x) = \widetilde{\omega^{\tilde{2}}} \sin(\omega x)$ with (continuous) derivative ${}^1f(x) = \omega^{\tilde{2}} \cos(\omega x)$ instead of ${}^1f(x) = 0$.

2. The sequence $f_n(x) = x - \tilde{n} x^n$ tends to $f(x) = x - \widetilde{\omega} x^\omega$ as $n \to \omega$ instead of $f(x) = x$ with (continuous) derivative ${}^1f(x) = 1 - x^{\acute{\omega}}$ instead of ${}^1f(x) = 1$. Conventionally, $f_n(x) = 1 - x^{\acute{n}}$ is discontinuous at $x = 1$.

3. The sequence $f_n(x) = nx(-\acute{x})^n$ does not tend to $f(x) = 0$ as $n \to \omega$, but to the continuous function $f(x) = \omega x(-\acute{x})^\omega$, and takes the value $\tilde{\epsilon}$ when $x = \widetilde{\omega}$.

Taylor's theorem: $+_m |{}^m f(a)| > \tilde{\nu}, {}^m f(a) \in {}^{\omega}\mathbb{C}, g(z) = (z - a)^\omega, |z - a| < \tilde{\epsilon}\omega$ and $z \to a \in {}^{\omega}\mathbb{C}$ imply

$$f(z) = T_\omega(z) := +_{m=0}^{\omega} \widetilde{m!} \, {}^m f(a)(z - a)^m.$$

Proof: From L'Hôpital's rule, it follows that Leibniz product rule gives

$$f(z) = \frac{(fg)(z)}{g(z)} = \frac{{}^1(fg)(z)}{{}^1g(z)} = ... = \frac{{}^{\acute{\omega}}(fg)(z)}{{}^{\acute{\omega}}g(z)} = \frac{{}^{\omega}(fg)(z)}{{}^{\omega}g(z)} = \widetilde{\omega!} \, {}^{\omega}(fg)(z) \quad \text{and}$$

$${}^{\omega}(fg)(z) = +_{m+n=\omega} \binom{\omega}{m} {}^m f(a) \, {}^{\omega-m}g(z) = {}^{\omega}g(z) +_{m=0}^{\omega} \widetilde{m!} \, {}^m f(a)(z - a)^m. \square$$

Conclusion: The second fundamental theorem implies for the remainder $R_n(z) := f(z) - T_n(z) = f(a) + \uparrow_a^z {}^1f(t) \downarrow t - T_n(z)$ by the mean value theorem where $\zeta \in {}^{a}\dot{\mathbb{C}}(z)$ and $p \in \mathbb{N}^*_{\le n}$

$$R_n(z) = \uparrow_a^z \widetilde{n!}(z - t)^n \, {}^{\grave{n}} f(t) \downarrow t = \widetilde{pn!}(z - \zeta)^{\grave{n} - p} \, {}^{\grave{n}} f(\zeta)(z - a)^p.$$

Proof by induction with integration by parts and induction step from $\acute{n}$ to n ($\acute{n} = 0$ see above):

$$f(z) = T_{\acute{n}}(z) + \widetilde{n!}(z - a)^n \, {}^n f(a) + \uparrow_a^z \widetilde{n!}(z - t)^n \, {}^{\grave{n}} f(t) \downarrow t = T_n(z) + R_n(z). \square$$

Remark: It holds that $(\epsilon^\iota - 1)/\iota = 1 = {}^1\exp(0)$ and thus $\downarrow_\epsilon y/\downarrow y = \tilde{y}$ from $\downarrow y/\downarrow x = y := \epsilon^x$ as well as $\downarrow x^n = \downarrow \epsilon^{n_\epsilon x} = n x^{\acute{n}} \downarrow x$ for $n \in {}^{\omega}\mathbb{N}^*$ by product and chain rule.

Remark: Unit circle and triangles easily show the relations $\sin \iota/1 = (\cos\iota - 1)/\iota$ and $\cos\iota/1 = -\sin\iota/\iota$. Hence, it holds $^1\sin(0) = \cos(0)$ and $^1\cos(0) = -\sin(0)$ as well as for $m \in {}^{\omega}\mathbb{N}$ and $n = \hat{k}$ de Moivre's formula:

$$(\cos z + \underline{\sin}\, z)^m = \epsilon^{\underline{mz}} = 1 + \mathop{\Large +}_{k=1}^{\check{\omega}} \left(\widetilde{\acute{n}!}(\underline{mz})^{\acute{n}} + \widetilde{n!}(\underline{mz})^{n}\right) = \cos(mz) + \underline{\sin}\,(mz).\square$$

Euler's sine formula: Zero and identity theorem (cf. [29], p. 41) for series plus the theorem above analogously yield $\Gamma(\tilde{2}) = \pi^{\tilde{2}}$ for the gamma function $\Gamma(z) := \omega!\omega^z/\mathop{\Large\times}_{k=0}^{\omega}(z+k)$ where $z \in {}^{\nu}\mathbb{C} \setminus -{}^{\nu}\mathbb{N}$ from

$$\frac{\epsilon^{\underline{\hat{\pi}z}} - 1}{\epsilon^{\underline{\pi z}}\underline{\hat{\pi}z}} = \frac{\epsilon^{\underline{\pi z}} - \epsilon^{-\underline{\pi z}}}{\underline{\hat{\pi}z}} = \frac{\sin(\pi z)}{\pi z} = \mathop{\Large +}_{k=0}^{\omega} \frac{(\underline{\pi z})^n}{\grave{n}!} = \mathop{\Large\times}_{k=1}^{\omega}(1 - z^2/k^2) = \frac{\tilde{z}}{\Gamma(z)\Gamma(-\acute{z})},$$

since all $\hat{\omega}$ zeros of the left- and right-hand side match due to $\epsilon^{\underline{\pi n}} = 1 + \sin 0.\square$

Functional equation of the gamma function: From $\Gamma(\grave{z}) = z\Gamma(z)\omega/(\omega + \grave{z})$ and $\Gamma(1) := 1$, it follows for sufficiently small $|z|$ integrating by parts that $\Gamma(\grave{z}) := \uparrow_0^{\omega} t^z \epsilon^{-t} \downarrow t = z\Gamma(z)$. This leads for $z = \tilde{2}$ and the substitution $x := t^{\tilde{2}}$ to the equation $\uparrow_0^{\omega} \epsilon^{-x^2} \downarrow x = \tilde{2}\pi^{\tilde{2}}$ relevant to statistics and ball computation.$\square$

Conclusions: The *Wallis product* is given by $\mathop{\Large\times}_{k=1}^{\omega} k^2/(k^2 - \tilde{4}) = \check{\pi}$. A logarithmic derivative (see [22], p. 324) shows $\uparrow_0^{\omega} \tilde{x} \sin x \downarrow x = \check{\pi}$. For the even function $\check{x}\cot\check{x} = \pm_{\check{n}=0}^{\check{\omega}} \widetilde{n!} B_n x^n$ where B_n are *Bernoulli numbers*, the GS yields by comparing coefficients in x $^1({}_{\epsilon}\sin(\pi x)) = \pi x \cot(\pi x) = 1 + 2x^2 \mathop{\Large +}_{n=1}^{\omega} \widetilde{n^2 - x^2}$ (see above) $\zeta(n) = -\widetilde{n!}\check{B}_n \underline{\hat{\pi}}^n$ where $s \in {}^{(\omega)}\mathbb{C}$, $\operatorname{Re} s > 1$ and $\zeta(s) := \mathop{\Large +}_{n=1}^{\omega} \tilde{n}^s$ is the *Riemann zeta function*.$\square$

Stirling formula: The asymptotic approximation $\tilde{\check{a}} < {}_{\epsilon}(\widetilde{\omega}^{\omega+\tilde{2}}\hat{\pi}^{-\tilde{2}}\omega!) + \omega < \tilde{a}$ applies for $a = 12\omega$.

Proof: The gamma function yields $\binom{\hat{\omega}}{\omega} = \frac{4^{\omega}}{(\pi\omega+\check{\pi})^{\tilde{2}}} \sim \frac{4^{\omega}}{(\pi\omega)^{\tilde{2}}}$ and $\mathop{\Large\times}_{n=\grave{\omega}}^{\hat{\omega}} n = (b\omega)^{\omega} \sim \omega! 4^{\omega} (\pi\omega)^{-\tilde{2}}$ with $b \in]1, 2[$. It follows for $c, d \in]\tilde{3}, 2[$ that

$$\omega! \sim c(\pi\omega)^{\tilde{2}}(d\omega)^{\omega}.$$

Subscripting $c_{\omega}^2/c_{\hat{\omega}}$ of c implies $c = 2^{\tilde{2}}$ and $d_{\grave{\omega}}^{\grave{\omega}}/d_{\omega}^{\omega}$ of d reveals $d = \tilde{\epsilon}$. The rest follows from [24].$\square$

Conclusion: Mathematical induction obtains $n! \in [\epsilon^{\tilde{n}[\tilde{t},\widetilde{12}]-n}\hat{\pi}^{\tilde{2}}n^{n+\tilde{2}}]$ for $n \in {}^{\omega}\mathbb{N}^*$ and $t = 12.004.\square$

From Green's Theorem to Stokes' one

Green's theorem: Given NRs $B \subseteq D^2$ for some h-domain $\mathbb{D} \subseteq {}^{(\omega)}\mathbb{R}^2$, sufficiently large $m \in \mathbb{N}^*$, infinitesimal $h = |{\downarrow}x| = |{\downarrow}y| = |\vec{\gamma}(s) - \gamma(s)| = \mathcal{O}(\tilde{\omega}^m)$, $(x, y) \in \mathbb{D}$, $\mathbb{D}^- := \{(x, y) \in \mathbb{D} : (x + h, y + h) \in \mathbb{D}\}$, and a CP $\gamma : [a, b[\to \partial\mathbb{D}$ followed anticlockwise, choosing $\vec{\gamma}(s) = \gamma(\vec{s})$ for $s \in [a, b[$, $A \subseteq [a, b]^2$, implies for sufficiently α-continuous functions $u, v : \mathbb{D} \to \mathbb{R}$ with not necessarily continuous ${\downarrow}u/{\downarrow}x$, ${\downarrow}u/{\downarrow}y$, ${\downarrow}v/{\downarrow}x$ and ${\downarrow}v/{\downarrow}y$

$$\uparrow_{\gamma}(u \downarrow x + v \downarrow y) = \uparrow_{(x,y)\in\mathbb{D}^-}\left(\tfrac{\downarrow v}{\downarrow x} - \tfrac{\downarrow u}{\downarrow y}\right)\downarrow(x, y).$$

Proof: Only $\mathbb{D} := \{(x, y) : r \le x \le s, f(x) \le y \le g(x)\}$, $r, s \in {}^{(\omega)}\mathbb{R}$, $f, g : \partial\mathbb{D} \to {}^{(\omega)}\mathbb{R}$ is proved, since the proof is analogous for each case rotated by ι. Every h-domian is union of such sets. It is sufficient to show $\uparrow_{\gamma} u \downarrow x = -\uparrow_{(x,y)\in D^-} \frac{\downarrow u}{\downarrow y}\downarrow(x, y)$, since the other relation is given analogously. Neglecting the regions of γ with respect to ${\downarrow}x = 0$ concerning the LI and regarding furthermore $s := h(u(r, g(r)) - u(t, g(t)))$ also shows that $-\uparrow_{\gamma} u \downarrow x - s = \uparrow_t^r u(x, g(x))\downarrow x - \uparrow_t^r u(x, f(x))\downarrow x = \uparrow_t^r \uparrow_{f(x)}^{g(x)} \frac{\downarrow u}{\downarrow y}\downarrow y \downarrow x = \uparrow_{(x,y)\in\mathbb{D}^-} \frac{\downarrow u}{\downarrow y}\downarrow(x, y).\square$

Finiteness criterion for series: Let $m, n, q, r \in \mathbb{N}$. Sum $S_r := \left|\mathop{\Large +}_{q=0}^{r} s_q\right|$ for $s_q \in {}^{(\omega)}\mathbb{C}$ is finite, if and only if $0 \le S_r = \left|\pm_{m=0}^{n} a_m\right| \le a_0$ for a sequence (a_m) such that $a_{\grave{m}} < a_m \in {}^{\nu}\mathbb{R}_{\ge 0}$ holds, since summands in sums can be arbitrarily sorted according to their signs and sizes, recombined or split into separate sums.$\square$

Theorem (series product): For $a_m, b_n \in {}^{(\omega)}\mathbb{K}$, replace the Cauchy product (see [29], p. 103) by

$$\mathop{\Large +}_{m=1}^{\omega} a_m \mathop{\Large +}_{n=1}^{\omega} b_n = \mathop{\Large +}_{m=1}^{\omega}\left(\mathop{\Large +}_{n=1}^{m}(a_n b_{m-\acute{n}} + a_{\omega-\acute{n}} b_{\omega-m+n}) - a_m b_{\omega-\acute{m}}\right).\square$$

Faulhaber's formula: The preceding theorem shows for $\acute{p} \in \mathbb{N}^*$ and Bernoulli numbers B_m (cf. [29], p. 163) using exponential series (see above) and GS by comparing coefficients $\mathop{\Large +}_{k=1}^{n} k^{\acute{p}} = \tilde{p} \mathop{\Large +}_{m=0}^{\acute{p}} \binom{p}{m} B_m n^{p-m}.\square$

Conclusion: $\mp_{k=1}^{n} k^{\acute{p}} = 2+_{k=1}^{\lfloor \check{n} \rfloor} \hat{k}^{\acute{p}} - +_{k=1}^{n} k^{\acute{p}} = \tilde{p} +_{m=0}^{\acute{p}} \binom{p}{m} B_m \left(2^p \lfloor \check{n} \rfloor^{p-m} - n^{p-m}\right).\square$

Example: The following series product has the finite value (cf. [4], p. 61 f.):

$$\left(\pm_{m=1}^{\omega} \widetilde{m}^{\tilde{2}}\right)^2 = +_{m=1}^{\omega}\left(\left(\frac{\widetilde{m}}{\omega - \acute{m}}\right)^{\tilde{2}} - \underline{1}^{\hat{m}} +_{n=1}^{m}\left(\left(\frac{\tilde{n}}{m - \acute{n}}\right)^{\tilde{2}} + \left(\frac{\widetilde{\omega - \acute{n}}}{\omega - m + n}\right)^{\tilde{2}}\right)\right) = 0.36590\ldots \quad \ll \frac{\zeta(\tilde{2})^2}{3+8^{\tilde{2}}}.$$

Example: The signum function sgn yields the following series product (cf. [4], p. 62):

$$+_{m=0}^{\omega} 2^{m^{\mathrm{sgn}(m)}} +_{n=0}^{\omega} \mathrm{sgn}(n - \gamma) = \acute{\omega} 2^{\grave{\omega}} \gg -2.$$

Remarks: If the moduli of $x \in \mathbb{C}$, $\downarrow x$ or $\widetilde{\downarrow x}$ have different orders of magnitude, the identity

$${}^{0}s(x) := \pm_{m=0}^{n} x^m = (1 - x\text{-}^{\grave{n}})/\grave{x}$$ yields by differentiating

$${}^{1}s(x) = \mp_{m=1}^{n} m x^{\acute{m}} = (\grave{n} x\text{-}^{n} - n x\text{-}^{\grave{n}} - 1)/\grave{x}^2.$$

The formulas above were sometimes miscalculated. For sufficiently small x, and sufficiently, but not excessively large n, the latter can be further simplified to $-\grave{x}^{-2}$, and remains valid when $x \geq 1$ is not excessively large. By successively multiplying ${}^{m}s(x)$ by x for $m \in {}^{\omega}\mathbb{N}^*$ and subsequently differentiating, other formulas can be derived for ${}^{\grave{m}}s(x)$, providing an example of divergent series. However, if ${}^{0}s(-x)$ is integrated from 0 to 1 putting $n := \omega$, an integral expression for ${}_{\epsilon}\omega + \gamma$ is obtained for Euler's constant γ.

L'Hôpital's rule solves the case of $x = -1$. Substituting $y := -\acute{x}$, by the binomial series a series is obtained with infinite coefficients (if ${}_{\epsilon}\omega$ is also expressed as a series, even an expression for γ is obtained). If the numerator of ${}^{0}s(x)$ is illegitimately simplified, finding incorrect results is risked, especially when $|x| \geq 1$. So ${}^{0}s(-\epsilon^{\underline{\pi}})$ is e.g. 0 for odd n, and 1 for even n, but not $\tilde{2}$.

Counter-directional theorem: If the path $\gamma : G = [a, b[\, \cap C \rightarrow V$ with $C \subseteq \mathbb{R}$ passes the edges of every n-cube of side length ι in the n-volume $V \subseteq {}^{(\omega)}\mathbb{R}^n$ with $n \in \mathbb{N}_{\geq 2}$ exactly once, where the opposite edges in all two-dimensional faces of every n-cube are traversed in reverse direction, but uniformly, then, for $D \subseteq \mathbb{R}^2, B \subseteq V^2, f = (f_1, ..., f_n) : V \rightarrow {}^{(\omega)}\mathbb{R}^n, \gamma(s) = x, \gamma(\vec{s}) = \vec{x}$ and $V_r := \{\vec{x} \in V : x \in V, \vec{x} \neq \overleftarrow{x}\}$, it holds that

$$\uparrow_{s \in G} f(\gamma(s))\, {}^{1}\gamma(s) \downarrow s = \uparrow_{\substack{(x, \vec{x}) \\ \in V \times V_r}} f(x) \downarrow x = \uparrow_{\substack{s \in G, \\ \gamma | \partial^{\acute{n}} V}} f(\gamma(s))\, {}^{1}\gamma(s) \downarrow s.$$

Proof: If two arbitrary squares are considered with common edge of length ι included in one plane, then only the edges of $V \times V_r$ are not passed in both directions for the same function value. They all, and thus the path to be passed, are exactly contained in $\partial^{\acute{n}} V.\square$

Theorem: Splitting $F : A \rightarrow {}^{(\omega)}\mathbb{C}$ into real and imaginary parts $F(z) := U(z) + \underline{V}(z) := f(x, y) := u(x, y) + \underline{v}(x, y)$, h-homogeneous $A \subseteq {}^{(\omega)}\mathbb{C}$ and $h = |\downarrow x| = |\downarrow y|$, with the NR $B \subseteq A^2$ for every $z = x + \underline{y} \in A, F$ is holomorphic if and only if the *Cauchy-Riemann differential equations* $\frac{\downarrow u}{\downarrow x} = \frac{\downarrow v}{\downarrow y}$ and $\frac{\downarrow v}{\downarrow x} = -\frac{\downarrow u}{\downarrow y}$ are satisfied.

Proof: The claim follows directly from $\frac{\downarrow u}{\downarrow x} + \frac{\downarrow \underline{v}}{\downarrow x} = \frac{\downarrow v}{\downarrow \underline{y}} - \frac{\downarrow u}{\downarrow \underline{y}} = \frac{\downarrow F}{\downarrow z} = \downarrow U(z) + \downarrow \underline{V}(z).\square$

Remark: The following necessary and sufficient condition is valid for F to be holomorphic:

$${}^{1}F(\bar{z}) = \frac{\downarrow f}{\downarrow x} = \frac{\downarrow f}{\downarrow \underline{y}} = \tilde{2}\left(\frac{\downarrow f}{\downarrow x} + \frac{\downarrow f}{\downarrow \underline{y}}\right) = \frac{\downarrow F}{\downarrow \bar{z}} = 0.$$

Goursat's integral lemma: If $f \in \mathcal{O}(\Delta)$ on a triangle $\Delta \subseteq {}^{(\omega)}\mathbb{C}$ but has no AD on Δ, then (cf. [22], p. 149 ff.)

$$I := \uparrow_{\partial \Delta} f(\zeta) \downarrow \zeta = 0.$$

Refutation of conventional proofs by means of a complete triangulation: Every minimal triangle $\Delta_s \subseteq \Delta$ wlog must either satisfy where κ, λ, and μ represent the vertices of Δ_s

$$I_s := \uparrow_{\partial \Delta_s} f(\zeta) \downarrow \zeta = f(\kappa)(\lambda - \kappa) + f(\lambda)(\mu - \lambda) + f(\kappa)(\kappa - \mu) = (f(\kappa) - f(\lambda))(\lambda - \mu) = 0 \qquad \text{or}$$

$$\begin{aligned}\uparrow_{\partial \Delta_s} f(\zeta) \downarrow \zeta &= f(\kappa)(\lambda - \kappa) + f(\lambda)(\mu - \lambda) + f(\mu)(\kappa - \mu) = (f(\kappa) - f(\lambda))\lambda + (f(\lambda) - f(\mu))\mu + (f(\mu) - f(\kappa))\kappa \\ &= {}^{1}f(\lambda)\left((\kappa - \lambda)\lambda - (\mu - \lambda)\mu + (\mu - \lambda)\kappa - (\kappa - \lambda)\kappa\right) = {}^{1}f(\lambda)\left((\mu - \lambda)(\kappa - \mu) - (\kappa - \lambda)^2\right) = 0.\end{aligned}$$

The direction in which $\partial\Delta$ is traversed is irrelevant. By holomorphicity and cyclic permutations, this can only happen for $f(\kappa) = f(\lambda) = f(\mu)$. Overall, it can be deduced that f must be constant, which contradicts the assumptions. This is because the term in large brackets is translation-invariant, since otherwise put $\mu := 0$ wlog, making this term 0, in which case $\kappa = \check{\lambda}(1 \pm 3\text{-}^{\tilde{2}})$ and $|\kappa| = |\lambda| = |\kappa - \lambda|$. However, since every horizontal and vertical line is homogeneous on ${}^{(\omega)}\mathbb{C}$, this cannot happen:

Otherwise, the corresponding sub-triangle would be equilateral and not isosceles and right-angled. Therefore, in both cases, $|I_s|$ must be at least $|{}^1f(\lambda)\mathcal{O}(\iota^2)|$, by selecting the vertices $0, \iota$ and $\underline{\iota}$ wlog. If L is the perimeter of a triangle, then it holds that $|I| \le 4^m|I_s|$ for an infinite natural number m, and also $2^m = L(\partial\Delta)/|\mathcal{O}(\iota^2)|$ since $L(\partial\Delta) = 2^m L(\partial\Delta_s)$ and $L(\partial\Delta_s) = |\mathcal{O}(\iota^2)|$. It holds that $|I| \le |{}^1f(\lambda)L(\partial\Delta)^2/\mathcal{O}(\iota^2)|$, causing the desired estimate $|I| \le |\mathcal{O}(\downarrow\zeta)|$ to fail, for example if $|{}^1f(\lambda)L(\partial\Delta)^2|$ is larger than $|\mathcal{O}(\iota^2)|$.□

Cauchy's integral theorem: Given the NRs $B \subseteq D^2$ and $A \subseteq [a, b]$ for some h-domain $\mathbb{D} \subseteq {}^{\omega}\mathbb{C}$, infinitesimal h, $f \in \mathcal{O}(\mathbb{D})$ and a CP $\gamma : [a, b[\rightarrow \partial\mathbb{D}$, choosing $\vec{\gamma}(s) = \gamma(\vec{s})$ for $s \in [a, b[$ gives $\uparrow_\gamma f(z)\downarrow z = 0$.

Proof: By the Cauchy-Riemann differential equations and Green's theorem, with $x := \operatorname{Re} z, y := \operatorname{Im} z, u := \operatorname{Re} f, v := \operatorname{Im} f$ and $\mathbb{D}^- := \{z \in \mathbb{D} : z + h + \underline{h} \in \mathbb{D}\}$, it holds that

$$\uparrow_\gamma f(z)\downarrow z = \uparrow_\gamma \left(u + \underline{v}\right)\left(\downarrow x + \downarrow \underline{y}\right) = \uparrow_{z \in \mathbb{D}^-}\left(\left(\tfrac{\downarrow \underline{u}}{\downarrow x} - \tfrac{\downarrow \underline{v}}{\downarrow y}\right) - \left(\tfrac{\downarrow v}{\downarrow x} + \tfrac{\downarrow u}{\downarrow y}\right)\right)\downarrow(x, y) = 0.\square$$

Remark: For $\widetilde{\omega} := 0$, the main theorem of Cauchy's theory of functions can be proven according to Dixon (as in [22], p. 228 f.), since the limit there shall be 0 resp. $\tilde{r}$ tends to 0 for $r \in {}^{\omega}\mathbb{R}_{>0}$ tending to ω. The in ${}^{\omega}\dot{\mathbb{C}} \subset {}^{\omega}\mathbb{C}$ (entire) functions $f(z) = {+}_{n=1}^{\omega} z^n \tilde{\omega}^{\tilde{n}}$ and $g(z) = \tilde{\omega}z$ give counterexamples to Liouville's (generalised) theorem and Picard's little theorem because of $|f(z)| < 1$ and $|g(z)| \le 1$. The function $f(\tilde{z})$ for $z \in {}^{\omega}\dot{\mathbb{C}}^*$ discounts Picard's great theorem.

Mean value equation: The integral formula yields for $\gamma([0, \hat{\pi}[) = \partial^r\dot{\mathbb{C}}(c), c \in {}^{(\omega)}\mathbb{C}$ and $r \in {}^{(\omega)}\mathbb{R}_{>0}$

$$f(c) = \widetilde{\hat{\pi}} \uparrow_0^{\hat{\pi}} f(c + r\epsilon^{\underline{\varphi}})\downarrow\varphi.$$

Remark: Substituting $z = c + r\epsilon^{\underline{\varphi}}$ implies the *mean value inequality* $|f(c)| \le |f|_\gamma$ (see [22], p. 160).

Example: Putting $f(x) := {+}_{n=1}^{\omega} \tilde{n}^2{}_\epsilon(1 + n^2x^2)$ implies ${}^1f(0) = \frac{f(\iota) - f(0)}{\iota - 0} = {+}_{n=1}^{\omega} \frac{\hat{x}}{1+n^2x^2}\Big|_0 = {+}_{n=1}^{\omega} \tilde{\iota}\tilde{n}^2{}_\epsilon(1 + n^2\iota^2) = \iota\omega = 0$, where the series expansion ${}_\epsilon\grave{x} = {\pm}_{n=1}^{\omega} \tilde{n}x^n$ for $x \in]-1, 1[$ was used differentiating term by term.

Definition: For a CP $\gamma : [a, b[\rightarrow {}^{(\omega)}\mathbb{C}$ and $z \in {}^{(\omega)}\mathbb{C}, \widetilde{\underline{\hat{\pi}}}\uparrow_\gamma \widetilde{\zeta - z}\downarrow\zeta$ is called *winding number* or *index* $\operatorname{ind}_\gamma(z) \in \mathbb{Z}$. The coefficients $a_{j,-1}$ of the function $f : A \rightarrow {}^{(\omega)}\mathbb{C}$ for $A \subseteq {}^{(\omega)}\mathbb{C}, n \in {}^{\omega}\mathbb{N}, a_{jk}, c_j \in {}^{(\omega)}\mathbb{C}$ and $f(z) = {+}_{j=0}^{n} {+}_{k=-\omega}^{\omega} a_{jk}(z - c_j)^k$ as well as pairwise different c_j are called *residues* $\operatorname{res}_{c_j} f$.△

Residue theorem: For γ and f as above, it holds that $\widetilde{\underline{\hat{\pi}}}\uparrow_\gamma f(\zeta)\downarrow\zeta = {+}_{j=0}^{n} \operatorname{ind}_\gamma(c_j)\operatorname{res}_{c_j} f$, since all $k \in {}^{\omega}\mathbb{Z} \setminus \{-1\}$ imply ${+}_{j=0}^{n} |\uparrow_\gamma a_{jk}(\zeta - c_j)^k\downarrow\zeta| = 0$ and $\widetilde{\underline{\hat{\pi}}}\uparrow_\gamma a_{jk}\widetilde{\zeta - c_j}\downarrow\zeta = \operatorname{ind}_\gamma(c_j)\operatorname{res}_{c_j} f$ is true for $k = -1$.□

Definition: A point $z_0 \in M \subseteq {}^{(\omega)}\mathbb{C}^n$ or of a sequence (a_k) for $k \in {}^{(\omega)}\mathbb{N}$ is called a *(proper) α-accumulation point* of M or of the sequence, if the ball ${}^{\alpha}\dot{\mathbb{C}}(z_0) \subseteq {}^{(\omega)}\mathbb{C}^n$ with centre z_0 and infinitesimal α contains infinitely many points from M or pairwise distinct members of $a_k \in {}^{(\omega)}\mathbb{C}^n$. Let α- be omitted for $\alpha = \widetilde{\omega}$.△

Remark: Choose the pairwise distinct zeros $c_k \in {}^{\widetilde{\omega}}\dot{\mathbb{C}} \subset \mathbb{D}$ for $z \in {}^{\omega}\mathbb{C}$ in $p(z) = \bigtimes_{k=0}^{\omega}(z - c_k)$ in such a way that $|f(c_k)| < \widetilde{\omega}$ for $f \in \mathcal{O}(\mathbb{D})$ on a domain $\mathbb{D} \subseteq \mathbb{C}$ where $f(0) = 0$. Let $\mathbb{D}$ contain ${}^{\widetilde{\omega}}\dot{\mathbb{C}}$ completely, which a coordinate transformation always achieves provided that $\mathbb{D}$ is sufficiently "large". The coincidence set $\{\zeta \in \mathbb{D} : f(\zeta) = g(\zeta)\}$ of $g(z) := f(z) + p(z) \in \mathcal{O}(\mathbb{D})$ contains an accumulation point at 0. Since $p(z)$ can take every conventional complex number, the deviation between f and g is non-negligible.

Since $f \ne g$, this contradicts the statement of the identity theorem like the (local) fact that all derivatives ${}^nu(z_0) = {}^nv(z_0)$ of two functions u and v can be equal at $z_0 \in \mathbb{D}$ for all n, but u and v may significantly differ further away maintaining to be holomorphic, since some holomorphic function has to be developed into a TS with approximated powers. The function $b(z) := \tilde{\nu}z$ for $z \in {}^{\nu}\dot{\mathbb{C}} \subset {}^{\nu}\mathbb{C}$ maps the simply connected ${}^{\nu}\dot{\mathbb{C}}$ holomorphicly to ${}^1\dot{\mathbb{C}}$.

A missing injectivity or surjectivity requires correcting the Riemann mapping theorem. Examples of such $f \in \mathcal{O}(\mathbb{D})$ include functions with $f(0) = 0$ that are restricted to ${}^{\tilde{\omega}}\dot{\mathbb{C}}$. Extending the upper limit from ω to $|\mathbb{N}^*|$ gives entire functions with an infinite number of zeros. The set of zeros is not necessarily discrete. Thus, the set of all functions $f \in \mathcal{O}(\mathbb{D})$ may contain zero divisors. The functions once again give counterexamples to Picard's little theorem since they omit at least $\acute{n}$ values in $\mathbb{C}$.

Theorem (binomial series): From $\alpha \in {}^{(\nu)}\mathbb{C}, \binom{\alpha}{n} := \widetilde{n!}\alpha\acute{\alpha}...(\grave{\alpha} - n)$ and $\left|\binom{\alpha}{\grave{m}}/\binom{\alpha}{m}\right| < 1$ for all $m \geq \nu$ where $\binom{\alpha}{0} := 1$, it follows for $z \in {}^{1-\tilde{\nu}}\dot{\mathbb{C}}$ (or $z \in {}^{(\omega)}\mathbb{C}$ for $\alpha \in {}^{(\omega)}\mathbb{N}$ the TS centred on 0) that $\grave{z}^{\alpha} = \mathop{\Large +}_{n=0}^{\omega}\binom{\alpha}{n}z^n$.□

Multinomial theorem: For $\zeta \in {}^{(\omega)}\mathbb{C}, z \in {}^{(\omega)}\mathbb{C}^k, k \in {}^{(\omega)}\mathbb{N}_{\geq 2}, m, n_j \in {}^{\omega}\mathbb{N}^*, |n| := \mathop{\Large +}_{j=1}^{k} n_j, z^n := \mathop{\Large\times}_{j=1}^{k} {z_j}^{n_j}$ and $\binom{m}{n} := \widetilde{n_1!...n_k!}m!$, it holds that

$$(1\uparrow_k^{\top} z)^m = \mathop{\Large +}_{|n|=m}\binom{m}{n}z^n.$$

Proof: Cases $k \in \{1, 2\}$ are clear. Induction step from k to $\grave{k}$ where $\binom{m}{n} = \binom{m}{n_1,...,n_{\acute{k}},p}\binom{p}{n_k,n_{\grave{k}}}$ and $p = n_k + n_{\grave{k}}$:

$$(1\uparrow_{\grave{k}}^{\top} z)^m\Big|_{\zeta_k = z_k + z_{\grave{k}}} = \mathop{\Large +}_{|n|=m}\binom{m}{n}z^n\Big|_{\eta_k! = n_k! n_{\grave{k}}!} = \mathop{\Large +}_{|n|=m}\binom{m}{n}z^n \qquad \text{resp. from } m \text{ to } \grave{m}$$

$$(1\uparrow_k^{\top} z)^{\grave{m}} = \grave{m}\uparrow_0^{z_j} (1\uparrow_k^{\top} z)^m\Big|_{z_j=\zeta}\downarrow\zeta + (1\uparrow_k^{\top} z)^{\grave{m}}\Big|_{z_j=0} = \grave{m}\uparrow_0^{z_j}\mathop{\Large +}_{|n|=m}\binom{m}{n}z^n\Big|_{z_j=\zeta}\downarrow\zeta + (1\uparrow_k^{\top} z)^{\grave{m}}\Big|_{z_j=0} = \mathop{\Large +}_{|\grave{n}|=\grave{m}}\binom{\grave{m}}{\grave{n}}z^{\grave{n}}.\square$$

General Leibniz formula: Putting $\downarrow^n := \downarrow_1^{n_1}...\downarrow_k^{n_k}$ and $\downarrow_j^{n_j} := \downarrow^{n_j}/\downarrow {z_j}^{n_j}$, it follows for $j, k, m, n \in {}^{(\omega)}\mathbb{N}$ and differentiable $f = f_1 \cdot ... \cdot f_k \in {}^{(\omega)}\mathbb{C}$ from the multinomial theorem $\downarrow^m f = \mathop{\Large +}_{|n|=m}\binom{m}{n}\downarrow^n f$.□

Taylor's theorem for several variables: For $n! := \mathop{\Large\times}_{j=1}^{k} n_j!, a, z \in {}^{(\omega)}\mathbb{C}^k$ and $(z - a)^n := \mathop{\Large\times}_{j=1}^{k}(z - a)^{n_j}$, it follows from the multinomial theorem also analogously to the proof of the simple TS for $n \in {}^{(\omega)}\mathbb{N}$

$$f(z) = T_{\omega}(z) := \mathop{\Large +}_{|n|=0}^{\omega}\widetilde{n!}\downarrow^n f(a)(z - a)^n.\square$$

Conclusion: Analogously to the simple TS, the remainder is for $\zeta \in {}^{a}\dot{\mathbb{C}}(z)$ and $k \in \mathbb{N}^*_{\leq \grave{n}}$

$$R_n(z) = (z - \zeta)^k/(1 - k/\grave{n})\mathop{\Large +}_{|m|=\grave{n}}\widetilde{m!}\downarrow^m f(\zeta)(z - a)^{m-k}.\square$$

Chain rule for several variables: For $z \in A \subseteq {}^{(\omega)}\mathbb{K}^k, g : A \to B \subseteq {}^{(\omega)}\mathbb{K}^m, f : B \to C \subseteq {}^{(\omega)}\mathbb{K}^n$, it holds for $k, m, n \in {}^{(\omega)}\mathbb{N}^*$ that:

$${}^1f(g(z)) = {}^1f(g(z))\,{}^1g(z).$$

Proof: Taylor's theorem for several variables implies with bounded $||r(z)||$ and $||s(g(z))||$

$$g(\vec{z}) = g(z) + {}^1g(z)\downarrow z + r(z)||\downarrow z||^2 \text{ and } f(\vec{g}(z)) = f(g(z)) + {}^1f(g(z))\downarrow g(z) + s(g(z))||\downarrow g(z)||^2.\square$$

Newton's method: Demanding above $f(\vec{z}) = f(z) + {}^1f(z)\downarrow z = 0$ implies $z_{\grave{n}} := z_n - {}^1f(z_n)^{-1}f(z_n)$ if ${}^1f(z_n)^{-1}$ is invertible resulting in quadratic convergence close to a zero.□

Stokes' theorem (cf. [17], p. 625 f.): If $-$ stands for sufficiently α-continuous functions $f_m : C \to {}^{\omega}\mathbb{R}$ above a term to be omitted for an alternating differential form $v := \mathop{\Large +}_{m=1}^{n} f_m \downarrow x_1 \wedge ... \wedge \overline{\downarrow x_m} \wedge ... \wedge \downarrow x_n$ of degree $\acute{n}$ on a cuboid $C = [a_1, b_1] \times ... \times [a_n, b_n] \subseteq {}^{\omega}\mathbb{R}^n$ where $\partial C := \mathop{\Large \mp}_{m=1}^{n}(F_{a,m} - F_{b,m})$ has the faces $F_{a,m} = [a_1, b_1] \times ... \times \{a_m\} \times ... \times [a_n, b_n]$ and $F_{b,m} = [a_1, b_1] \times ... \times \{b_m\} \times ... \times [a_n, b_n]$, then $\uparrow_C \downarrow v = \uparrow_{\partial C} v$.□

Proof: The second fundamental theorem and Fubinis theorem (see above) give

$$\uparrow_C \downarrow v = \mathop{\Large \mp}_{m=1}^{n}\uparrow_{a_n}^{b_n}...\overline{\uparrow_{a_m}^{b_m}}...\uparrow_{a_1}^{b_1}(f_m(x_1, ..., a_m, ..., x_n) - f_m(x_1, ..., b_m, ..., x_n))\downarrow x_1 \wedge ... \wedge \overline{\downarrow x_m} \wedge ... \wedge \downarrow x_n$$

and

$$\tfrac{\downarrow f_m}{\downarrow x_m}\downarrow x_m \wedge \downarrow x_1 \wedge ... \wedge \downarrow x_{\acute{m}} \wedge \downarrow x_{\grave{m}} \wedge ... \wedge \downarrow x_n = -\underline{1}^{\hat{m}}\tfrac{\downarrow f_m}{\downarrow x_m}\downarrow x_1 \wedge ... \wedge \downarrow x_n.\square$$

Remark: Stokes' theorem also holds for n-dimensional manifolds consisting of cuboids.

Examples: For $n \in \mathbb{N}^*$, $[-1, 1]^n$ and ${}^1\dot{\mathbb{R}}^n$ consist of $\check{\iota}^{-n}$ resp. $\tilde{\iota}^n\pi^{\check{n}}/\Gamma(\check{n} + 1)$ points (cf. [30], p. 254).

5 Number Theory

This section requires Set Theory, Topology and Nonstandard Analysis. Let $k \in \mathbb{N}$.

Prime number theorem: For $\pi(x) := |\{p \in \mathbb{P}_{\le x} : x \in {}^{\omega}\mathbb{R}\}|$, it holds that $\pi(\omega) = \widetilde{{}_{\epsilon}\omega}\ \omega + \mathcal{O}({}_{\epsilon}\omega\ \omega^{\tilde{2}})$.

Proof: From intervals of fix length $y \in {}^{\omega}\mathbb{R}_{>0}$, $\check{y}$ set-2-tuples of prime numbers are formed such that the first interval has the unchanged representative prime number density and the second interval is empty, then the interval with the second most prime number density is followed by the second least one etc. The Stirling formula (see Nonstandard Analysis) suggests the prime gap $n = \epsilon^{\sigma} = \mathcal{O}({}_{\epsilon}(n!))$ for $n \in {}^{\omega}\mathbb{N}_{\ge 2}$.

For induction basis $n = 2$ resp. 3, the hypothesis states the first interval to contain $x_n/{}_{\epsilon}x_n$ primes for $n \in {}^{\omega}\mathbb{N}_{\ge 2}$ and $x_4 \in [2, 4[$. Stepping from x_n to x_n^2 finds $\pi(x_n^2) = \pi(x_n)\check{x}_n$ primes only from $\pi(x_n) = x_n/{}_{\epsilon}x_n$. The average prime gap is ${}_{\epsilon}x_n$, but the maximal one ${}_{\epsilon}x_n^2$, and the maximal x_n^2 to x_n behaves like ω to $\omega^{\tilde{2}}$.□

Remark: Replacing 2 by $m \in {}^{\omega}\mathbb{N}_{>2}$ for $\widetilde{m}y^{\acute{m}}$ set-m-tuples gives the same result. Induction and the sieve of Eratosthenes show by the Dirichlet prime number theorem both infinitely many prime and composite Mersenne numbers $M_n := 2^n - 1$ for $n \in {}^{\omega}\mathbb{N}^*$ (see [25], p. 174 f. and 354 – 365).

Singmaster's theorem: There are maximally 8 distinct binomial coefficients of the same value > 1.

Proof: The existence is clear due to $\binom{3003}{1} = \binom{78}{2} = \binom{15}{5} = \binom{14}{6}$ and the structure of Pascal's triangle. With $p \in {}^{\omega}\mathbb{P}, a, b, c, d \in {}^{\omega}\mathbb{N}^*, \hat{a} \le r := p - b, \hat{a} < \hat{c} \le n := p - d, b < d$ and $s \notin \mathbb{P}$ for every $s \in [\max(r - \acute{a}, \grave{n}), r]$, Stirling's formula $n!^2 \sim \pi(\hat{n} + \tilde{3})(\tilde{\epsilon}n)^{\hat{n}}$ and the prime number theorem imply $\omega\binom{r}{a} \le {}_{\epsilon}\omega\binom{n}{c}$ for $p \to \omega$.□

Giuga's theorem: Every number $n \in {}^{\omega}\mathbb{N}_{\ge 2}$ is prime if and only if $+_{k=1}^{\acute{n}} k^{\acute{n}} \equiv -1 \mod n$.

Proof: Fermat's (little) theorem settles the case $n \in {}^{\omega}\mathbb{P} \cup {}^{\omega}2\mathbb{N}^*$. Otherwise, the harmonic and geometric mean H_n resp. G_n imply $+_{p \in {}^{\omega}\mathbb{P}}\ \tilde{p} - \times_{p \in {}^{\omega}\mathbb{P}}\ \tilde{p} = m/H_n - G_n^{-m} = c \in {}^{\omega}\mathbb{N}^*$ (cf. [1], p. 3 f.). This contradicts $c < 1$ due to $H_n(m) \ne H_n = H_n(m, n) = m/(c + \tilde{n}) < n^{\widetilde{m}} = G_n$.□

Prime gap theorem: For the set M_g of mediate prime gaps, $M_g \supset {}^{\omega}2\mathbb{N} \cup \{1\}$ holds.

Proof by induction: The claim states that beside 1 the mediate prime gaps m_p exist from 2 to $\acute{p}$. It is true for primes $p \in \{2, 3\}$. By stepping from $p \to p + 2$, the prime number theorem permits no greater prime gap than those occurred as m_{p+2}. Hence $\grave{p}$ exists, too.□

Goldbach's theorem: Every even whole number > 2 is the sum of two primes.

Proof: For $\hat{m} + \hat{n} = p_{m+r,n-r} + q_{m+r,n-r} + r, r \in \{0, 2, \ldots, \max(g(n))\}$, it follows alike $\hat{m} + \hat{n} = p_{m+s,n-s} + q_{m+s,n-s} + s$, $s \in \{0, 2, \ldots, \max(g(n)) + 2\}$. This implies $\hat{m} + \hat{n} + 2 = p_{\grave{m}+r,\grave{n}-r} + q_{\grave{m}+r,\grave{n}-r} + r, r \in \{0, 2, \ldots, \max(g(\grave{n}))\}$. Induction yields then the claim by the previous theorem.□

Theorem of second Hardy-Littlewood conjecture: For $m, n \in {}^{\omega}\mathbb{N}_{\ge 2}$, $\pi(m + n) \le \pi(m) + \pi(n)$ holds.

Proof by induction for n: Cases $\pi(m + \grave{n}) = \pi(m + n)$ resp. $\pi(\grave{n}) = \grave{\pi}(n)$ and $\pi(m + \grave{n}) = \grave{\pi}(m + n)$ are clear. For $\pi(\grave{n}) = \pi(n)$, the last case implies the claim for $m := n + k$ and $\pi(n) = \widetilde{\sigma}n + \mathcal{O}(\sigma n^{\tilde{2}})$ by $\grave{\pi}(\hat{n} + k) \le \pi(n + k) + \pi(n)$ and $\pi(4) \le \hat{\pi}(2)$ and so on due to

$$(n + k)({}_{\epsilon}(\hat{n} + k) - {}_{\epsilon}(n + k))\sigma + n({}_{\epsilon}(\hat{n} + k) - \sigma)_{\epsilon}(n + k) \ge {}_{\epsilon}(\hat{n} + k)_{\epsilon}(n + k)\sigma.\square$$

Coefficient theorem for ω-ANs: No zero of normalised irreducible polynomials and series with at least one $a_k \notin {}^{\omega}\mathbb{Z}$ is ω-AN, since these are pairwise distinct and uniquely determined.□

Bounding theorem for ω-ANs: No $z \in \mathbb{C}^*$ such that $|z| \notin [\tilde{\omega}, \omega]$ is ω-AN.

Proof: In a polynomial or series equation, put $a_m = 1$ and $a_k = -\acute{\omega}$ for $k < m$, then the real case follows from the GS formula after taking the reciprocal. The exact limit value can be found by replacing ω by $\omega(m) = \omega - \acute{\omega}/\omega(m)^m$. The complex case is solved by putting i.a. $x = \grave{y}\omega$ for $y \in {}^{\omega}\underline{\mathbb{R}}^*$.□

Conclusions: For every $z \in \mathbb{R} + \underline{\mathbb{R}}$ where $|z| \notin \mathbb{B}$ and $\eta := z^{\grave{\omega}}$, the GS is $\sum_{n=0}^{\omega} z^n = \acute{\eta}/\acute{z} \notin {}^{\omega}_{\omega}\mathbb{A}_{\mathbb{C}}$. Putting $k = \omega^2!$ implies $\Gamma(z) := k!\,k^z/(z\grave{z}...(z+k)) \notin {}^{\omega}_{\omega}\mathbb{A}_{\mathbb{R}}$ for all $z \in {}^{\omega}\mathbb{R} \setminus -{}^{\omega}\mathbb{N}$. Euler's number $\epsilon = (1 + \tilde{\omega})^{\omega}$ implies $\epsilon = (k\omega + 1)/\omega!$ for $k > \omega$ (exponential series).□

Counting theorem for ANs: For the Riemann zeta function ζ and the average number $z(m)$ of zeros of a $\grave{m}$-polynomial or $\grave{m}$-series, the ANs asymptotically have for $\check{\kappa} = n$ the number

$$\mathbb{A}(m,n) = \widetilde{\zeta(\grave{m})}\, z(m)\check{\kappa}^m \left(n + \mathcal{O}(\sigma)\right).$$

Proof: The case $m = 1$ has by ([25], p. 323) the error term $\mathcal{O}(\sigma n)$ and represents the number $4\sum_{k=1}^{n}\varphi(k) - 1$ of finite fractions by the φ-function. For $m > 1$, the divisibility conditions neither change the error term $\mathcal{O}(\sigma)$ nor the leading term. By $1/\zeta(\grave{m}) = \prod_{i=1}^{n}(1 - \tilde{p}_i^{\grave{m}})$ (GS!), which absorbs multiples of primes p_i, polynomials or series such that $\gcd(a_0, a_1, \ldots, a_m) \neq 1$ are excluded.□

Fundamental theorem of algebra: Every non-constant polynomial $p \in {}^{(\omega)}\mathbb{C}$ has at least one complex root.

Indirect proof: An affine substitution of variables reduces to the case $\widetilde{p(0)} \neq \mathcal{O}(\iota)$. Then $f(z) := \widetilde{p(z)}$ is holomorphic for all $z \in {}^{(\omega)}\mathbb{C}$ by supposing $p(z) \neq 0$. For $f(\tilde{\iota}) = \mathcal{O}(\iota)$, the mean value inequality (see Nonstandard Analysis) yields $|f(0)| \leq |f|_{\gamma}$ where $\gamma = \partial^r\dot{\mathbb{C}}$ and arbitrary $r \in {}^{(\omega)}\mathbb{R}_{>0}$, and hence $f(0) = \mathcal{O}(\iota)$.□

Examples: For $m = 1$, there are $3(n/\check{\pi})^2 + \mathcal{O}(\sigma n)$ real solutions. For $m = 2$, $\check{9}n^3/\zeta(3) + \mathcal{O}(\sigma n^2)$ real solutions arise, since a real polynomial of degree 2 has two real zeros with probability $\frac{9}{16}$ by the quadratic formula. For $a_m = 1$, there are $z(m)\check{\kappa}^{\acute{m}}(\kappa + \mathcal{O}(\sigma))$ algebraic integer solutions.

Remark: In the real case, $z(m)$ is asymptotically equal to ${}_{\epsilon}m/\check{\pi} + \mathcal{O}(1)$ according to [14].

Conclusion: For $m = n = \acute{\nu}$, it is true that $|{}^{\nu}\mathbb{A}_{\mathbb{R}}| = \tilde{\pi}\sigma\check{\kappa}^{\acute{\nu}}\left(\hat{\nu} + \mathcal{O}(\sigma)\right)$ and $|{}^{\nu}\mathbb{A}_{\mathbb{C}}| = \check{\kappa}^{\nu}\left(\check{\nu} + \mathcal{O}(\sigma)\right)$.□

Bunyakovsky theorem: If $a_n > 0$ holds for a minimal polynomial $p \in {}^{(\omega)}\mathbb{C}$ of degree n where $p(i) \perp p(j)$ for every ${}^{\omega}\mathbb{N}^* \ni i < j \in {}^{\omega}\mathbb{N}^*$, then it has infinitely many primes as values.

Proof: The Dirichlet prime number theorem yields for $m \in {}^{\omega}\mathbb{N}^*$ infinitely many primes $q_m = bm + c$ where $b \perp c \in {}^{\omega}\mathbb{N}^*$ are fix. Vice versa there are infinitely many m such that $d_m := p(m) = \widetilde{m}(q_m - c)$. By rearranging the latter formula, the claim follows by mathematical induction for the polynomial degree n.□

Theorem: The BBP series $s_k := \sum_{n=1}^{\omega} p(n)\widetilde{q(n)}\tilde{b}^n$ implies, where $b \in {}^{\omega}\mathbb{N}_{\geq 2}$ and integer polynomials resp. series p and $q \in {}^{\omega}\mathbb{Z}$ with $q(n) \neq 0$ and $\deg(p) < \deg(q)$, that $s_k \notin {}^{\omega}_{\omega}\mathbb{A}_{\mathbb{R}}$ due to $\mathrm{den}(s_k) \geq b^m > \omega$ with $m \in \mathbb{N}^*$.□

Theorem: The maximum distance between two neighbouring real ω-ANs is $\Omega/\acute{\omega}$ for the not ω-algebraic omega constant $\Omega = \tilde{\epsilon}^{\Omega} = W(1)$ (see below Lambert-W function).

Proof: The distance between two real ω-ANs is largest around the points ± 1. A real ω-algebraic x approximates 1 satisfying the polynomial or series equation $\acute{x}x^{\acute{m}}\acute{\omega} = 1$ for $x > 1$ or $x^m = -\acute{x}\acute{\omega}$ for $x < 1$.□

Approximation theorem for ω-ANs: The average asymptotic error to approximate every real ω-AN of degree $n > 1$ by a real ω-AN of degree $m < n$ is $|{}^{\omega}\mathbb{Z}|^{-m}\widetilde{{}_{\epsilon}\omega}\zeta(\grave{m})\check{\pi}$.

Proof: The number of ω-ANs approximately evenly distributed between fixed limits increases in ${}^{\omega}\mathbb{R}$ by a factor of approximately $|{}^{\omega}\mathbb{Z}|$ per degree. The error corresponds to the distance between ω-ANs. The non-real ω-ANs are less dense.□

Conclusion: Two distinct real ω-ANs have an average distance of at least $|{}^{\omega}\mathbb{Z}|^{-\acute{\omega}}\widetilde{{}_{\epsilon}\omega}\pi$. Determining this minimum distance exactly requires an infinite non-linear non-convex optimisation problem to be solved. Therefore, the ν-ANs have an approximate order of $\mathcal{O}(\nu)$. This disproves Roth's theorem, which begins with false assumptions.□

The greatest-prime criterion (GPC) for ω-ANs: If $r := \tilde{a}pb \pm \tilde{s}t \in {}^{\omega}\mathbb{R}$ is an irreducible fraction, where $a, b, s,$ and t are natural numbers, $abst \neq 0$, $a + s > 2$, and the (second-)greatest prime number $p \in {}^{\omega}\mathbb{P}, p \nmid b$ and $p \nmid s$, then $r \notin {}^{\omega}_{\omega}\mathbb{A}_{\mathbb{R}}$, since the prime number theorem implies $\mathrm{den}(\widetilde{aps}(bs \pm apt)) \geq \hat{p} \geq \hat{\omega} - \mathcal{O}({}_{\epsilon}\omega\, \omega^{\tilde{2}}) > \omega$.□

Theorem: The constants of Catalan (G), Gieseking ($\pi\,{}_{\epsilon}\beta$), Smarandache (S_1) and Taniguchi (C_T) are not ω-algebraic because of the GPC.□

Theorem: It holds $\pi \notin {}^{\omega}_{\omega}\mathbb{A}_{\mathbb{R}}$ provided that its different representations are accepted as Wallis product, or product using the gamma function with value $\tilde{2}$ (see Nonstandard Analysis), by alternatively applying the GPC to the Leibniz series, or the TS of $\arcsin(x)$ at $x = 1$.□

Theorem: The constants of (C_{Artin}), Baxter (C^2), Chaitin (Ω_F), Champernowne (C_{10}), Copeland-Erdős (C_{CE}) (valid for every base from ${}^{\nu}\mathbb{N}^*$), Erdős-Borwein (E), Feller-Tornier (C_{FT}), Flajolet and Richmond (Q), Glaisher-Kinkelin (A), Heath-Brown-Moroz (C_{HBM}), Landau-Ramanujan (K), Liouville ($£_{Li}$), Murata (C_M), Pell (P_{Pell}), Prouhet-Thue-Morse (C_{PTM}), Sarnak (C_{sa}) and Stephen (C_S) as well as the Euler resp. Landau totient constant (ET resp. LT), the twin prime constant (C_2) and the carefree constants (K_1, K_2 and K_3) are not ω-ANs, since simplifying cannot remove a certain power of a prime from a fraction.□

Theorem: The trigonometric and hyperbolic functions and their inverse functions, the digamma function ψ, the Lambert-W-function, the Ein function, the (hyperbolic) sine integral $S(h)i$, Euler's Beta function B, and, for $s, u \in {}^{\omega}\mathbb{N}^*$ and $t \in {}^{\omega}\mathbb{N}$, the generalised error function E_t, the hypergeometric function ${}_0F_t$, the Fresnel integrals C and S and the Bessel function I_t and the Bessel function of the first kind J_t, the Legendre function χ_t, the polygamma function ψ_t, the generalised Mittag-Leffler function $E_{s,t}$, the Dirichlet series $\sum_{n=1}^{\omega} \tilde{n}^s f(n)$ with maximally finite real $|f(n)|$, the prime zeta function $P(s)$, the polylogarithm Li_s and the Lerch zeta-function $\Phi(q, s, r)$ always yield no ω-ANs for real arguments and maximal finite real $|q|$ and $|r|$ at points where their TS converge.

Proof: GPC, Dirichlet prime number theorem and Wallis product prove the claim. For the digamma function, the claim follows from the missing of the ω-algebraicity of Euler's constant γ below.□

Gelfond-Schneider theorem: It holds $a^b \notin {}^{\omega}_{\omega}\mathbb{A}_{\mathbb{C}}$ where $a, c \in {}^{\omega}\mathbb{A}_{\mathbb{C}} \setminus \mathbb{B}$ and infinitesimal $\varepsilon, b \in {}^{\omega}\mathbb{A}_{\mathbb{C}} \setminus {}^{\omega}_{\omega}\mathbb{R}$.

Indirect proof: The minimal polynomials p (and q) of c^r resp. $c^{r\pm\varepsilon} = a^b$ for maximal $r \in {}^{\omega}_{\omega}\mathbb{R}_{>0}$ and $f = p\ (q)$ lead to the contradiction ${}^1f(c^{r(\pm\varepsilon)}) \neq 0 = (f(c^r) - f(c^{r\pm\varepsilon}))/(c^r - c^{r\pm\varepsilon}) = {}^1f(c^{r(\pm\varepsilon)})$.□

Theorem: Let $\gamma := Li_1(1) - {}_{\epsilon}\omega = \uparrow_1^{\omega}\left(\widetilde{\lfloor x \rfloor} - \tilde{x}\right)\downarrow x$ for $Li_s(z) := \sum_{n=1}^{\omega} \tilde{n}^s z^n, z \in {}^1\dot{\mathbb{C}}$ and $s \in {}^{\omega}\mathbb{C}$, where rearranging yields $\gamma \in]0, 1[$. Accepting ${}_{\epsilon}\omega = Li_1(\tilde{2})\ {}_2\omega$ shows $\gamma \notin {}^{\omega}_{\omega}\mathbb{A}_{\mathbb{R}}$ with a precision of $\mathcal{O}(\tilde{2}^{\omega}\tilde{\omega}\ {}_{\epsilon}\omega)$.

Proof: The GS implies $-_{\epsilon}(-\acute{x}) = Li_1(x) + \mathcal{O}(\tilde{\omega}x^{\grave{\omega}}/\acute{x}) + t(x)\downarrow x$ for $x \in [-1, 1 - \tilde{\nu}]$ and $t(x) \in {}^{\omega}\mathbb{R}$ such that $|t(x)| < \omega$. Apply Fermat's (little) theorem and GPC to $\mathrm{den}(\tilde{p}(1 - 2^{-p}\ {}_2\omega))$ for $p = \max {}^{\omega}\mathbb{P}$.□

Theorem: It holds $\sum_{n=-1}^{\omega} \widetilde{a_n b_n} \notin {}^{\omega}_{\omega}\mathbb{R}$ only for arbitrary $b_n \in {}^{\omega}\mathbb{N}^*$, if it does also for $\sum_{n=-1}^{\omega} \widetilde{a_n}$ or $\widetilde{a_{-1}} - \widetilde{a_{\omega}}$ where $a_n \in {}^{\omega}\mathbb{N}^*$, since $b_n := 1\,(+\,a_n)$ (telescoping sum) may be true (cf. [6], p. 346).□

Theorem: All $p \in {}^{\omega}\mathbb{P}$ where $q \in Q := {}^{\omega}_{\omega}\mathbb{R}_{>0} \ni q^x$ and ${}_2\omega \gg |x| \in {}^{\omega}\mathbb{R}$ imply $x \in {}^{\omega}\mathbb{Z}$ for $p^2 \nmid \mathrm{num}(q)\,\mathrm{den}(q)$.

Indirect proof: Let wlog $x > 0$. Since there is no contradiction for $x \in {}^{\omega}\mathbb{N}^*$, assume $x \in Q \setminus {}^{\omega}\mathbb{N}^*$. Since this implies $q^x \in {}^{\omega}\mathbb{A}_R \setminus Q$, assume $x := k/d \in {}^{\omega}\mathbb{R}_{>0} \setminus Q$ for $d, k \in \mathbb{N}^*$ and $d \perp k$. This implies $q^k = r^d$ for an $r \in Q$. The fundamental theorem of arithmetic yields a numerator or denominator of q or r greater than 2^{ω}.□

Remark: The previous theorem proves the Alaoglu and Erdős conjecture, which states that p^x and q^x are ν-real for distinct $p, q \in {}^{\nu}\mathbb{P}$ if and only if $x \in {}^{\nu}\mathbb{Z}$ and $|x|$ is not excessively large.

Example: Every $m_0 \in {}^{(\omega)}\mathbb{N}^*$ implies $m_{\check{\iota}} = 1$ if there are no cycles in $m_{\hat{k}} := \left\lfloor m_k^{\check{3} - \chi_{2\mathbb{N}}(m_k)} \right\rfloor$.

Collatz theorem: Every $n_0 \in {}^{(\omega)}\mathbb{N}^*$ implies $n_{\check{\iota}} \in \{1, 2, 4\}$ where $\hat{n}_{\hat{k}} := n_k + \chi_{2\mathbb{N}}(\acute{n}_k)(5n_k + 2)$.

Proof: If the trivial cycle (here by [27]) is the only one, every such oftener descending than ascending procedure ends up below every $n_0 \geq 2$, since the expected value is $3^{\tilde{2}}/2$ of the previous one.□

Definition: When two numbers $x_0, y_0 \in {}^{\omega}\mathbb{C}^*$ do not satisfy any non-trivial polynomial equation $p(x, y) = 0$, so they are called *ω-algebraically independent*.△

Theorem: The GPC, with $\epsilon = (1 + \tilde{p})^p$ for maximal $p \in {}^{\omega}\mathbb{P}$ and π as Wallis product, yields pairwise ω-algebraically independent representations of $A, C_2, \gamma, \epsilon, K$ and π.□

Remark: These conditions are not sufficient as the examples $a_n := 1, b_n := 2$ resp. $(a_n) := (12, 12, 12, 12, 12, 6, 12, 20, 30, 42, ..., \acute{\omega}\omega), b_n := 1$ show for the sums $\check{\omega} + 1$ resp. $(\omega - 2)/\acute{\omega}$. Considering $(n!)$ resp. (a_n) where $\acute{a}_{\grave{n}} = a_n \grave{a}_n$, it holds that $+_{n=1}^{\omega} \widetilde{a_n b_n} \notin {}^{\omega}_{\omega}\mathbb{A}_{\mathbb{R}}$ for $b_n := n + 2$ resp. $b_n := 1$.

Beal's theorem: Equation $a^m + b^n = c^k$ for $a, b, c, d, e, r, s \in {}^{\omega}\mathbb{N}^*$ and $k, m, n \in {}^{\omega}\mathbb{N}_{\geq 3}$ implies $d := \gcd(a, b, c) > 1$.

Indirect proof: Put necessarily $(da)^n + (db)^m = (dc)^n$ to cancel d after computing mod d. Multiply $b^s + e^m = c^s$ generalised by d^{rm} again to yield $b = 1$, $c = d$ and $s \neq m \notin 2\mathbb{N}$. The second general form $1 + d^s = e^m$ shows likewise the claim by contradiction for $m \neq s \notin 2\mathbb{N}$.□

Conclusion: Due to $m \neq s$, no $n \in {}^{\omega}\mathbb{N}_{\geq 3}$ satisfies $a^n + b^n = c^n$ for arbitrary $a, b, c \in {}^{\omega}\mathbb{N}^*$.□

Catalan's theorem: It holds that $\{(m, n, x, y) \in {}^{\omega}\mathbb{N}^4_{\geq 2} : 1 + x^m = y^n\} = \{(3, 2, 2, 3)\}$.

Indirect proof: The previous theorem gives $\min(m, n) = 2$. Squaring shows $1 + 4\check{x}^2 = (1 + \acute{y})^n$ and $(1 + 2^r z)^n \equiv 1 + 2^r \check{x}^2 \equiv 1 \mod 2^{\grave{r}}$ for every $r \in {}^{\omega}\mathbb{N}_{\geq 3}$, and $1 + x^2 \equiv 2 \neq 0 \equiv 2^n \check{y}^n \mod 8$ where $\acute{y} = 4z$ and n is odd. Odd m and $s \in {}^{\omega}\mathbb{N}^*$ imply $x^m = \acute{y}\grave{y}$ where $s^m \neq \acute{y} \in {}^{\omega}2\mathbb{N}^*$ such that $x^m = 8\check{s}\acute{s} = 8$.□

Erdős-Moser theorem: Faulhaber's formula (see Nonstandard Analysis) implies for $k, n \in {}^{\omega}\mathbb{N}^*$ that from $n\grave{n} \mid +_{m=1}^{n} m^k = \grave{n}^k$ only the solution $k = 1 = \acute{n}$ follows due to $1 < n \nmid \grave{n}$.□

Three-cube theorem: It holds $S := \{n \in \mathbb{Z} : n \neq \pm 4 \mod 9\} = \{n \in \mathbb{Z} : n = a^3 + b^3 + c^3 + 3(a + b)c(a - b + c) = (a + c)^3 + (b - c)^3 + c^3\} \subset a^3 + b^3 + c^3 + 6\mathbb{Z}$, since independent mathematical induction by equitable variables $a, b, c \in \mathbb{Z}$ first shows $\{0, \pm 1, \pm 2, \pm 3\} \subset S$, and then the claim.□

Brocard's theorem: It holds that $\{(m, n) \in {}^{\omega}\mathbb{N}^2 : n! + 1 = m^2\} = \{(5, 4), (11, 5), (71, 7)\}$.

Proof: From $n! = \acute{m}\grave{m}$, it follows that $m = \hat{r} \pm 1$ for $r \in {}^{\omega}\mathbb{N}^*$ and $n \geq 3$. Thus $n! = \hat{r}(\hat{r} \pm 2) = 8s(\hat{s} \pm 1)$ holds for $s \in {}^{\omega}\mathbb{N}^*$. Let $2^q \mid n!$ and $2^{\grave{q}} \nmid n!$ for maximal $q \in {}^{\omega}\mathbb{N}^*$. Therefore $n! = 2^q(\hat{u} + 1)$ holds for $u \in {}^{\omega}\mathbb{N}^*$ and necessarily $n! = 2^q(2^{q-2} \pm 1)$. Then the prime factorisation of $n!$ requires $n \leq 7$ giving the claim.□

Wilson's prime number theorem: Only the primes $p \in \{5, 13, 563\}$ satisfy $q := p^2 \mid (\acute{p}! + 1)$.

Indirect proof: Adding and subtracting afterwards a suitable power of two on the left-hand side, division by $\acute{p}$ or $\grave{p}$ yields in $\hat{n}q + \acute{q} = \acute{p}!$ for $n \in {}^{\omega}\mathbb{N}^*$ a higher one on the right-hand side for sufficiently big p.□

Littlewood theorem in conventional mathematics: For all $a, b \in {}^{\nu}\mathbb{R}$ and $n \in {}^{\nu}\mathbb{N}^*$, it holds that $\liminf_{n \to \infty} n \, \|na\|_d \, \|nb\|_d = 0$ where $\| \cdot \|_d$ is the *distance to the nearest integer*.

Proof: For $k, m \in {}^{\nu}\mathbb{N}^*$ as denominators of the continued fraction of a resp. b with precision $g \in {}^{\nu}\mathbb{R}_{>0}$ and n/km again and again integer, Dirichlet's approximation theorem (see [25], p. 63) yields that:

$$\liminf_{n \to \infty} n\|na\|_d\|nb\|_d = \liminf_{n \to \infty} nO(\tilde{n})^2 = \liminf_{n \to \infty} O(\tilde{n}) = 0.\square$$

Refutation by nonstandard mathematics: For $a = b := \tilde{\omega}^{\check{3}}$, it holds that $\omega \, \|\omega a\|_d \, \|\omega b\|_d = 1$.□

Example: For $s \in {}^{(\omega)}\mathbb{C}$ where $\operatorname{Re}(s) \leq 1$ and $z := \tilde{2}^{\acute{s}}$, $\zeta(s) = +_{n=1}^{\omega} \tilde{n}^s$ has definitely no analytic continuation (cf. [13], p. 4) and no zeros. This disproves the Riemann hypothesis:

$$\mp_{n=1}^{\omega} \tilde{n}^s = z +_{n=1}^{\check{\omega}} \tilde{n}^s - +_{n=1}^{\omega} \tilde{n}^s \neq \acute{z} +_{n=1}^{\omega(/2)} \tilde{n}^s.$$

Theorem: Since the Dirichlet L-function $L(s, \chi) = +_{n=1}^{\omega} \chi(n) \, \tilde{n}^s$ has only zeros for $s = 0$ and nontrivial Dirichlet characters $\chi(n)$, it disproves the generalised Riemann hypothesis.□

6 Euclidean Geometry

The following section presupposes Set Theory, Topology and Nonstandard Analysis.

Definition: Let R a Euclidean space as subspace of $\mathbb{R}^n$ with $n \in {}^{\omega}\mathbb{N}_{\geq 3}$ (see Set Theory). Two distinct *points* as elements of R are called *pair (of points)*. A one-dimensional point set in R is called a *line* if each point has at least one and at most two gaplessly neighbouring points. A one-dimensional subspace of R is called *straight line*, a two-dimensional one is said to be a *plane*. The *distance* is given by the Euclidean norm $||\cdot||$ and results summed up in the Euclidean *arc length* A.△

Definition: A *line segment* is a connected subset of a straight line, whose *starting point* and *end point* determine it precisely and have to only one neighbour finite distance. Two line segments are said to *intersect* if they have one interior point in common. Two line segments are said to be *parallel* if one line may be obtained from the other by means of a translation. All points of a plane with the same distance (called *radius*) to another point constitute a *circle*, which forms with its interior a *disk*.△

Result: By defining short straight lines, many counterexamples can be given based on the above to Pasch's axiom, the axiom of line completeness, as well as various other axioms and their equivalents. The subspace concept can prove Hilbert's incidence axioms I.6 and I.8. If a straight line uniquely defines a parallel straight line running through a given point by their shortest distance, the parallel postulate is dispensable in Euclidean Geometry. The axioms of order and congruence are redundant.

Contrary to Hilbert's incidence axiom I.4, three distinct points are not enough to uniquely determine a plane since there are different infinities. Hilbert's incidence axiom I.7 is false (cf. [11], p. 2 - 17), since planes are yet limited in infinity and can precisely intersect in one point. All (three ancient) exactly unsolvable problems may be solved with arbitrary precision by the fundamental theorem of set theory via the intercept theorem and Farey sequences ([25], p. 62 f.).

Rating two straight lines only as parallel when they lie in the same plane and do not intersect, the parallel postulate does not hold: The reciprocal of the distance between the straight line and the given point may be greater than infinity or smaller than $|{}^{\omega}\mathbb{N}|$, and then infinitely many distinct straight lines can be found that pass through the given point without intersecting the original straight line. Formulating the axiom of completeness vaguely, Hilbert renders it finally unapt in future.

The Archimedean axiom must be extended to the case where a segment is marked off an infinite natural number of times without exceeding the starting point or end point of a straight line. It must be in the finite case replaced by the Archimedean theorem (see Set Theory). Pasch's axiom is also unnecessary, since every straight line must be fully contained in the interior of some triangle due to its maximum length, and hence so must its boundary, provided that one of its points is in this interior.

Normal theorem: For $n \in {}^{\omega}\mathbb{N}_{\geq 2}$, every convex set S from ${}^{\omega}\mathbb{R}^n$ contains a point lying on $\hat{n}$ distinct normals through ∂S (also through a vertex), since this is possible every 90° in the plane (cf. [2], p. 14 f.).□

Toeplitz' conjecture: Every Jordan curve admits an inscribed square.

Counterexamples: The right-angled triangle with two sides of length ι and the obtuse triangle where a vertex of at most one inscribed square is infinitesimally moved within the limits.□

Theorem: Infinitesimally juxtaposing equichordal points within the limits leads to a Jordan domain with more than one equichordal point (cf. [2], p. 9 f.).□

Fickett's theorem: Any relative position of two overlapping congruent rectangular n-prisms P and Q (see [2], p. 25) with $n \in {}^{\omega}\mathbb{N}_{\geq 2}$ and $\grave{m} := \hat{n}$ implies for the exact standard measure μ, which is A for $n = 2$:

$$\widetilde{m} < r := \mu(\partial P \cap Q)/\mu(\partial Q \cap P) < m.$$

Proof: The underlying extremal problem has its maximum for rectangles with side lengths s and $s + \hat{\iota}$. Putting $q := 3 - \hat{\iota}\tilde{s}$ implies $\min r = \tilde{q} \leq r \leq \max r = q$. The proof for $n > 2$ works analogously.□

7 Linear programming

The following likewise presupposes Set Theory, Topology and Nonstandard Analysis.

Diameter theorem for polytopes and polyhedra: Every diameter of an n-dimensional polytope or polyhedron given by m constraints with $m, n \in {}^{\omega}\mathbb{N}_{\geq 2}$ is at most $2(m + n - 3)$.

Proof: At most $\acute{m}$ hyperplanes can be combined to form an incomplete cycle of dimension 2, and there are at most $n - 2$ options to deviate sideways in the remaining dimensions. Traversing each minimal segment requires at most two edges and thus yields the factor 2.□

Theorem on Strassen's algorithm: For sufficiently large $n := 2^{\ell}$, $\ell \in {}^{\nu}\mathbb{N}^*$, $ß := {}_27$ and $A \in {}^{\nu}\mathbb{C}^{n \times n}$, the GS proves the reduction of the running time $T(n) = O(n^{ß})$ (see [5], pp. 31 ff.) by approximately $\tilde{3}$ for the computation of

$$AA^H = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} A_{11}^H & A_{21}^H \\ A_{12}^H & A_{22}^H \end{pmatrix} = \begin{pmatrix} A_{11}A_{11}^H + A_{12}A_{12}^H & A_{11}A_{21}^H + A_{12}A_{22}^H \\ A_{21}A_{11}^H + A_{22}A_{12}^H & A_{21}A_{21}^H + A_{22}A_{22}^H \end{pmatrix}. \square$$

Theorem on fast matrix multiplication: With A and $B = (b_{ij}) \in {}^{\nu}\mathbb{C}^{n \times n}$ as above and n as before, AB can be computed for

$$s := \min\{2^k > \max_{i,j}(|a_{ij}|^2, |b_{ij}|^2)\acute{n} : k \in {}^{\nu}\mathbb{Z}\}$$

from

$$A_{11}(sB_{11} + B_{12}),\ A_{21}(sB_{11} + B_{12}),\ A_{12}(sB_{21} + B_{22}),\ A_{22}(sB_{21} + B_{22})$$

by computing modulo s in running time $O(p^2)$ with $p = \ell n$ (see Appendix).□

The hybrid Intex–Merit method for LPs

Problem statement.

Let

$$A \in {}^{\nu}\mathbb{R}^{m \times n}, \qquad b \in {}^{\nu}\mathbb{R}^{m}, \qquad c \in {}^{\nu}\mathbb{R}^{n}, \qquad e \in {}^{\nu}\mathbb{R}.$$

We first consider the primal LP

$$\max\ e + c^{\top}x \qquad \text{unter} \qquad Ax \leq b, \qquad x \geq 0.$$

The corresponding dual LP is

$$\min\ e + b^{\top}y \qquad \text{unter} \qquad A^{\top}y \geq c, \qquad y \geq 0.$$

For feasible pairs (x, y), weak duality gives

$$c^{\top}x \leq b^{\top}y.$$

At the optimum, the duality gap vanishes:

$$b^{\top}y - c^{\top}x = 0.$$

Meaning of pro.

The parameter pro $\in \{1, -1\}$ determines how the input data are interpreted.

For pro $= 1$, the input problem is read as a maximising primal problem:

$$\max\ e + c^{\top}x \qquad \text{unter} \qquad Ax \leq b, \qquad x \geq 0.$$

For pro $= -1$, the internal primal-dual orientation is exchanged. The method then works with the transposed-negated tableau form, so that the internal objective function plays the same geometric role as in the maximising case. However, the mathematical objective value must not be output with the wrong sign. Therefore the internally scaled objective value is multiplied by pro during back-transformation:

$$f_{\text{orig}} = e + \text{pro}\ \kappa f_{\text{int}}, \qquad \kappa > 0.$$

Thus the factor pro belongs to the back-transformation of the internal objective value, not to the primal and dual feasibility conditions themselves. The test quantities

$$Ax \leq b, \qquad A^{\top}y \geq c, \qquad b^{\top}y - c^{\top}x \geq 0$$

are always evaluated in the respective back-transformed original orientation.

Theorem. The hybrid Intex–Merit method determines a primal-dual solution (x^o, y^o) for every solvable LP of the above form, provided that the arithmetic precision used, the relaxation sequence, and the final crossover (CO) resolve the active structure of the problem. The Intex core works by *Int*er-/*Ex*trapolations. For a matrix density $d \in [0,1]$ and a suitable binary precision scale $_2(\tilde{\alpha}\rho)$, the geometric Intex phase has, in the unstructured case, the order

$$O\Big({}_2(\tilde{\alpha}\rho)^2\, dmn\Big).$$

If the linear algebra is carried out in structured form by fast-Toeplitz decomposition (FTD)/fast-shift-recursive (FSR), the respective structured operator cost takes the place of dmn. Let $z := m + n$.

Intex relaxation. Instead of enforcing directly

$$Ax \le b, \qquad A^\top y \ge c, \qquad b^\top y - c^\top x = 0$$

the Intex method considers, for $r \in [0, \rho]$, a family of relaxed primal-dual polytopes

$$P_r := \Big\{(x,y)^\top \in {}^v\mathbb{R}^z_{\ge 0} : b^\top y - c^\top x \le r,\ Ax - b \le r\mathbb{1}_m,\ c - A^\top y \le r\mathbb{1}_n\Big\}.$$

The initial radius is chosen so that the origin 0 is certainly contained in P_ρ. One possible choice is

$$\rho := s\,|\min\{b_1, \dots, b_m, -c_1, \dots, -c_n\}|\,, \qquad s \in]1,2],$$

possibly after prior scaling and shifting of the data.

The original LP is solved exactly when a pair $(x,y) \in P_0$ is found. By strong duality, this is equivalent to

$$Ax \le b, \qquad x \ge 0, \qquad A^\top y \ge c, \qquad y \ge 0, \qquad b^\top y - c^\top x = 0.$$

Thus the same variables simultaneously solve

$$\max\{c^\top x : x \in {}^v\mathbb{R}^n_{\ge 0},\ Ax \le b\}$$

and

$$\min\{b^\top y : y \in {}^v\mathbb{R}^m_{\ge 0},\ A^\top y \ge c\}.$$

Normalisation and scaling. Before the geometric iteration, the objective function, right-hand sides, and constraint matrix are normalised. The aim is not to change the problem, but to obtain a balanced representation of the three types of error

$$Ax - b, \qquad c - A^\top y, \qquad b^\top y - c^\top x.$$

The scaling is intended to prevent one of these three quantities from dominating numerically. To ensure that the objective function coupling in the gradient descent of the subsequent merit polishing neither disappears nor the hinge errors are numerically masked, an initial scalar linear transformation is applied. Specifically, the vector c and the constant offset e are scaled by the factor $\lambda = ||c||^{-1}$ (for $c \ne 0$).

Under this transformation, the primal solution x remains invariant, while the dual solution y and the absolute objective function values scale strictly linearly with λ. At the end of the overall procedure, an exact back-transformation to the original physical space is performed. After the computation, all variables and objective values are transformed back to the original orientation. In particular, when pro $= -1$, the internally computed objective value is output with the correct sign.

Geometric Intex step. Within a fixed polytope P_r, an interior centre of mass is approximated. For each coordinate, a feasible interval is determined. For $v \in P_r$, one approximately forms

$$v_k^* := \min \check{v}_k + \max \check{v}_k, \qquad k = 1, \dots, z,$$

This yields a geometric centre $v = (x,y)^\top$, which remains inside the current relaxation and serves as a stable starting point for the next radius reduction.

The most recent centre movements are then extrapolated. For this purpose, a direction Δv is formed, and a maximally feasible step along

$$v(w) = v + w\Delta v, \qquad w \ge 0,$$

is determined. The new radius is obtained from the smallest relaxation still required:

$$r_{\text{new}} = \min\{r \geq 0 : v(w) \in P_r\}.$$

If $r_{\text{new}} = 0$, is achieved, a primal-dual optimal pair has been reached. If $r_{\text{new}} > 0$ remains stably separated from zero, this indicates infeasibility, insufficient precision, or an unresolved active structure.

Intex as a feasible starting point. The essential advantage of the Intex phase is that it does not search blindly in the exterior space. It works with relaxed feasible regions and typically provides a point which is already well oriented in primal-dual terms:

$$Ax - b \lesssim r\mathbb{1}_m, \qquad c - A^\top y \lesssim r\mathbb{1}_n, \qquad b^\top y - c^\top x \lesssim r.$$

This point is not necessarily optimal yet, but it is often close to an active face of the optimal polytope. This is precisely where the Merit polishing phase starts.

Primal-dual Merit polishing

Residuals. For $x \geq 0$ and $y \geq 0$, let

$$r^p(x) := \max(Ax - b, 0) \in {}^v\mathbb{R}^m,$$

$$r^d(y) := \max(c - A^\top y, 0) \in {}^v\mathbb{R}^n$$

denote the primal and dual violations. The signed duality gap is

$$g(x, y) := b^\top y - c^\top x.$$

For a maximising primal problem, only

$$g_+(x, y) := \max(g(x, y), 0)$$

is relevant as a positive gap violation. At the exact optimum,

$$r^p(x) = 0, \qquad r^d(y) = 0, \qquad g(x, y) = 0.$$

The unscaled control quantities are

$$p_\infty = \|r^p(x)\|_\infty, \qquad d_\infty = \|r^d(y)\|_\infty, \qquad g_{\text{abs}} = |b^\top y - c^\top x|.$$

The primal and dual objective values are

$$p^* = c^\top x, \qquad d^* = b^\top y.$$

The constant offset e is added only at output.

Merit function. The Merit function combines primal violation, dual violation, and duality gap:

$$\widehat{\Phi}(x, y) = \|r^p(x)\|_2^2 + \|r^d(y)\|_2^2 + \gamma^2\, g_+(x, y)^2, \qquad \gamma > 0.$$

In scaled variables

$$x = S_x\check{x}, \qquad y = S_y\check{y}$$

and with diagonal residual weights

$$D_p = \operatorname{diag}(d_p), \qquad D_d = \operatorname{diag}(d_d)$$

the scaled form is

$$\widehat{\Phi}(\check{x}, \check{y}) = \|D_p \max(AS_x\check{x} - b, 0)\|_2^2 + \|D_d \max(c - A^\top S_y\check{y}, 0)\|_2^2 + d_g^2 \max(b^\top S_y\check{y} - c^\top S_x\check{x}, 0)^2.$$

Optionally, a small regularisation term

$$\varepsilon \left(\|\check{x}\|_2^2 + \|\check{y}\|_2^2\right)$$

may be added.

Gradient. With

$$w_i^p = d_{p,i}^2 \, r_i^p(\breve{x}), \qquad w_j^d = d_{d,j}^2 \, r_j^d(\breve{y}), \qquad w_g = d_g^2 \, g_+(\breve{x}, \breve{y})$$

the gradients are

$$\frac{\downarrow \Phi}{\downarrow \breve{x}_j} = s_{x,j} \left(\mathop{+}_{i=1}^{m} A_{ij} w_i^p - c_j w_g \right), \qquad j = 1, \dots, n,$$

$$\frac{\downarrow \Phi}{\downarrow \breve{y}_i} = s_{y,i} \left(b_i w_g - \mathop{+}_{j=1}^{n} A_{ij} w_j^d \right), \qquad i = 1, \dots, m.$$

The term $b_i w_g$ is essential: the gradient with respect to y must contain the contribution of $b^\top y$ in the duality gap. Without this contribution, the dual side is steered incorrectly.

Feasible step control. Starting from the Intex point, a descent step is computed. In the simple projected variant,

$$\breve{x}' = \Pi_{\mathbb{R}_+^n} \left(\breve{x} - \alpha \nabla_{\breve{x}} \widehat{\Phi} \right), \qquad \breve{y}' = \Pi_{\mathbb{R}_+^m} \left(\breve{y} - \alpha \nabla_{\breve{y}} \widehat{\Phi} \right).$$

In the feasible hybrid variant, a hard orthant projection is avoided during the Merit phase. Instead, a step is rejected or shortened if it would leave the feasible relaxation or violate non-negativity. This preserves the character of the Intex relaxation.

A step is accepted only if an Armijo-type condition is satisfied:

$$\widehat{\Phi}(\breve{x}', \breve{y}') \le \widehat{\Phi}(\breve{x}, \breve{y}) - c_1 \alpha \left(\|\breve{x}' - \breve{x}\|_2^2 + \|\breve{y}' - \breve{y}\|_2^2 \right).$$

Otherwise, α is reduced.

Barzilai–Borwein step size. After an accepted step,

$$s_k = \begin{pmatrix} \breve{x}^{\grave{k}} - \breve{x}^k \\ \breve{y}^{\grave{k}} - \breve{y}^k \end{pmatrix}, \qquad q_k = \begin{pmatrix} \nabla_{\breve{x}} \widehat{\Phi}(\breve{x}^{\grave{k}}, \breve{y}^{\grave{k}}) - \nabla_{\breve{x}} \widehat{\Phi}(\breve{x}^k, \breve{y}^k) \\ \nabla_{\breve{y}} \widehat{\Phi}(\breve{x}^{\grave{k}}, \breve{y}^{\grave{k}}) - \nabla_{\breve{y}} \widehat{\Phi}(\breve{x}^k, \breve{y}^k) \end{pmatrix}$$

are formed. For $\eta_k := s_k^\top q_k > 0$, one sets $\alpha_{\grave{k}} = \tilde{\eta}_k \, s_k^\top s_k$, and subsequently restricts it to an interval $[\alpha_{\min}, \alpha_{\max}]$. If $\eta_k \le 0$, the previous usable step size is retained, or a conservative initial step size is used.

Algebraic CO

Purpose. The Merit phase is a first-order method. It reduces residuals and the duality gap, but does not always deliver the active basis exactly. Therefore, if required, an algebraic CO follows. It attempts to reconstruct an exactly feasible primal-dual solution from the approximation point (x, y).

Active candidates. Primal active constraints have those indices i for which $a_i^\top x \approx b_i$ holds. Active non-negativity conditions have those indices j for which $x_j \approx 0$ holds. From these candidates, linear systems of size n are formed. A primal candidate is accepted if the reconstructed x satisfies

$$x \ge 0, \qquad Ax \le b.$$

A dual candidate y is then sought with

$$y \ge 0, \qquad A^\top y \ge c.$$

Only if, in addition, $|b^\top y - c^\top x|$ is sufficiently small is the pair accepted as an optimal solution.

Certified CO. An isolated good primal candidate is not sufficient. Likewise, an isolated good dual candidate is not sufficient. Only a primal-dual pair satisfying simultaneously

$$p_\infty = 0, \qquad d_\infty = 0, \qquad g_{\text{abs}} = 0$$

up to the required tolerance is accepted. This prevents a merely formally feasible dual candidate with an incorrect objective value from replacing the solution.

Staged CO. The staged CO substantially limits the combinatorial cost.

In Stage 1, the probably active constraints are obtained from small slacks $b_i - a_i^\top x$ and variables x_j. This stage is fast, but may fail if the active sets are poorly separated.

In Stage 2, the candidate set is enlarged. Additional constraints with medium-sized slacks are admitted. This increases the probability of success, but also the number of tested bases.

In Stage 3, a complete or nearly complete fallback is admitted. This stage is expensive, but robust. It is used only if the previous stages do not produce a certified primal-dual solution.

The CO can also be switched off completely. In that case, the method outputs the best Intex–Merit point found so far, without algebraic basis reconstruction. This algorithmic limitation is accompanied by a strictly memory-efficient implementation at the software level. The inner loop of the combinatorial candidate search operates entirely without allocation.

All vectors and index arrays are pre-allocated, and only overwritten during each iteration. This prevents any dynamic memory pressure and ensures that the garbage collection of the runtime environment is not triggered during the critical phase. Due to this mechanical efficiency, the theoretical worst-case complexity of the basis search loses its impact in practice, as traversing the permutations only requires minimal CPU cycles.

Timeout. If a time limit is reached, the method terminates in a controlled way and returns the best state known up to that point:

$$x_{\text{best}}, \qquad y_{\text{best}}, \qquad p_\infty, \qquad d_\infty, \qquad g_{\text{abs}}, \qquad p^*, \qquad d^*.$$

If no certified solution has been found up to that point, this is explicitly indicated. A timeout is therefore not a mathematical false conclusion, but an incomplete computational state.

Behaviour for unsolvable LPs

For a primal infeasible LP, there is no $x \geq 0$ with $Ax \leq b$: neither can Intex lower the radius to $r = 0$, nor can Merit enforce $p_\infty = 0$. Typically, $p_\infty > 0$ remains, or the radius reduction stagnates.

If the dual LP is infeasible, then there is no $y \geq 0$ with $A^\top y \geq c$. Usually $d_\infty > 0$. In primal terms, this often corresponds to an unbounded primal problem.

If the problem is formally solvable but numerically badly scaled or degenerate, Intex and Merit may stagnate even though a solution exists. The CO then decides whether the active structure can nevertheless be reconstructed algebraically. If this fails, the method does not provide a false proof of optimality, but a non-certified approximation point.

Thus the method distinguishes, in practice, four cases:

$$\begin{array}{ll}
p_\infty = d_\infty = g_{\text{abs}} = 0 & \text{certified as solved,} \\[1em]
p_\infty > 0 & \text{primal feasibility not reached,} \\[1em]
d_\infty > 0 & \text{dual feasibility not reached,} \\[1em]
p_\infty, d_\infty \approx 0, \; g_{\text{abs}} > 0 & \text{feasible, but optimality not reached.}
\end{array}$$

A complete infeasibility certificate in the sense of a Farkas witness is therefore not automatically included. However, it may be added as an additional diagnostic stage.

Structured linear algebra by FTD and FSR

Motivation. Intex, Merit, and CO repeatedly require operations of the form

$$Ax, \qquad A^\top y, \qquad b^\top y, \qquad c^\top x.$$

For a general sparse matrix this costs $O(\mathrm{nnz}(A))$. For a dense unstructured matrix, the cost is $O(mn)$. If A has structure, this need not be the case. For Toeplitz, Hankel, block-circulant, or convolution-like matrices, the operator evaluations can be replaced by FTD and FSR.

FTD. The FTD uses the fact that many structured matrices vary only slowly along their diagonals or can be described by convolution operators. A matrix $A \in {}^{v}\mathbb{R}^{m\times n}$ is then not treated as an arbitrary collection of mn entries, but described by a small number of generators. Formally, one seeks a representation of the type $A \approx T_1 T_2$, where T_1 and T_2 are Toeplitz, Hankel, or related convolution operators.

The entries of these operators are obtained from local averages, diagonal sums, and recursive compensation steps. Instead of the full matrix, only generator, diagonal, or convolution data are stored. Matrix-vector products, transposed products, and selected projections can therefore be carried out in structured form.

If the effective structure width is ℓ and $p = \ell n$, the cost of structured operations is typically of the form $O(p^2)$, instead of dense unstructured cost.

FSR. The FSR complements the FTD arithmetically. While the FTD provides a structured representation of the operator, FSR organises the actual computation recursively and with controlled word length. A matrix or operator is decomposed into high- and low-weight parts:

$$A = (A_{\mathrm{hi}} \ll s) + A_{\mathrm{lo}},$$

where $\ll s$ denotes a shift by s positions. Products, Schur complements, inverse subsystems, and reconstructions are recursively built from smaller parts.

The advantage is that intermediate values remain controlled and BigInt or BigFloat arithmetic does not grow unnecessarily. The recursion is also well suited to parallelisation. Thus FSR is particularly suitable for high precision, exact integer arithmetic, and large structured linear systems.

Interaction. The FTD provides the structured decomposition $A \approx T_1 T_2$, while the FSR performs the resulting products, inversions, Schur-complement steps, and back-transformations recursively. Here the combination replaces the general operator evaluation $O(dmn)$ by $O(p^2)$.

Embedding in Intex–Merit. The FTD/FSR structure enters at three points.

First, it accelerates the Intex phase, because bounds, cuts, and centre movements are repeatedly computed there by operations with A and $A^\top$.

Second, it accelerates the Merit polishing phase, because the function value and gradient consist only of

$$Ax - b, \qquad c - A^\top y, \qquad b^\top y - c^\top x$$

Third, it can accelerate the CO. Solving a dense active system of size n without structure costs $O(n^3)$. However, if the active subsystem has the same Toeplitz, Hankel, or convolution structure, FTD/FSR reduce this step to $O(p^2)$ per structured subsystem.

Thus Intex, Merit, CO, FTD, and FSR do not form separate building blocks, but one coherent architecture: Intex = geometric reduction, Merit = numerical polishing, CO = algebraic certification FTD/FSR = structured linear algebra.

Complexity

Let $\mathcal{M}_A$ denote the cost of a complete operator evaluation with A and $A^\top$. Then

$$\mathcal{M}_A = \begin{cases} O(\mathrm{nnz}(A)), & \text{for sparse matrices,} \\ O(mn), & \text{for dense unstructured matrices,} \\ O(p^2), & \text{for matrices structured by FTD/FSR.} \end{cases}$$

The Intex phase therefore has the order $O\Big(_2(\tilde{\alpha}\rho)^2\,\mathcal{M}_A\Big)$. For K accepted iterations and B backtracking evaluations, the Merit polishing additionally requires $O((K+B)\mathcal{M}_A)$. For the CO without structure, $C = O(Tn^3)$, if T active basis candidates are actually tested. Since $T \le \binom{Q}{n}$ may hold for Q candidate equations, an unrestricted CO is potentially combinatorial. The staged model initially reduces Q substantially and activates larger candidate sets only if required.

With structured linear algebra, the approximate cost instead becomes $C = O(Tp^2)$. Altogether, the order of the hybrid method is

$$O\Big(_2(\tilde{\alpha}\rho)^2\mathcal{M}_A + (K+B)\mathcal{M}_A + C\Big).$$

The memory requirement in addition to the matrix or generator representation is linear, $O(m+n)$, in the Intex–Merit core: with dense matrix storage this gives altogether $O(mn+m+n) = O(mn)$. With structured storage, the memory requirement depends on the generator width and is closer to $O(p+m+n)$.

Hardware vs. Software Arithmetic.

In addition to the algorithmic complexity, the practical runtime is primarily dictated by the machine level. The solver gains a substantial speed advantage from avoiding software-emulated high-precision arithmetic (such as multi-precision data types) where it is not strictly necessary.

Since the system is excellently balanced through preconditioning, hardware-native 64-bit floating-point arithmetic is sufficient for almost all practical problems. This consistent avoidance of emulated overhead allows matrix-vector multiplications to be performed vectorially and directly in hardware, reducing the effective solution time by several orders of magnitude.

Consequences and applications

A second solution on the same objective face.

If x^o is an optimal solution, the LP

$$\max\left\{\|x - x^o\|_1 : c^\top x = c^\top x^o,\ Ax \le b,\ x - x^o \in [-1,1]^n,\ x \in {}^v\mathbb{R}^n_{\ge 0}\right\}$$

can determine a second optimal solution, provided that the optimal face contains more than one point. The dual vector y^o can be treated analogously.

Linear systems of equations.

A linear system $Ax = b$ can be treated by means of the Merit function $\widehat{\Phi}(x) = \|Ax - b\|_2^2$. In the regular case, this leads to the solution of the system; in the singular or inconsistent case, only to the least-squares solution. The gradient is $\nabla\Phi(x) = A^\top(Ax-b)$. Additional constraints such as $x_j \ge 0$ can be added by projection or by an LP formulation.

Regularity.

The matrix A is regular if and only if the homogeneous system $Ax = 0$ has only the zero solution. This can be tested by an LP of the form

$$\max\{\|x\|_1 : Ax = 0,\ x \in [-1,1]^n\}$$

If the optimal value is zero, the kernel is trivial.

Inverse columns.

For regular $A \in {}^v\mathbb{R}^{n\times n}$, each column α_j of A^{-1} can be determined by $A\alpha_j = e_j$, where e_j is the j-th unit vector. Without structure, this is a classical linear system; with FTD/FSR, the structure of A can be exploited in the solution.

Eigenvalue problems.

Eigenvalue equations $Ax = \lambda x$ can be transferred into constraints of an extended optimisation problem. In practice, this is of interest when additional constraints, normalisations, or structural conditions are imposed on x and λ. For general dense matrices, classical spectral methods usually remain more suitable; however, for structured or constrained formulations, the Merit-LP viewpoint can be advantageous.

Polynomial and Padé approximation. Least-squares problems of the form

$$y = r + Xc, \qquad X^\top r = 0$$

can also be formulated as linear or quadratic Merit problems. Additional constraints on the coefficients c can be inserted directly. For Padé approximations, the discrete Fourier transform can provide structured convolution components, which are then processed by FTD/FSR.

Convex programmes. For convex programmes $\min\{f_1(x) : x \in {}^v\mathbb{R}^n,\ (f_2(x), \ldots, f_m(x))^\top \leq 0\}$ the same idea can be transferred to nonlinear residuals. The linear terms $Ax - b$ are then replaced by the violations $f_k(x)$. The basic structure remains:

$$\text{feasibility + duality or optimality condition} \quad \longrightarrow \quad \text{Merit function.}$$

A polynomial runtime statement, however, additionally depends on the evaluability, smoothness, conditioning, and structure of the functions f_k.

Assessment

The Intex method works geometrically. It generates interior points from relaxed polytopes and reduces the radius of the relaxation. The Merit method works analytically. It minimises a primal-dual residual function and thereby improves feasibility and the duality gap. The CO works algebraically. It reconstructs a certified active basis from a good approximation point. The FTD and FSR methods, by contrast, work structurally and arithmetically. They replace general matrix operations by structured generator, convolution, and recursion computations.

Thus the hybrid method combines three different principles and, in the general case, works as a geometrically supported primal-dual first-order method which can become considerably faster when structured linear algebra is available. Its strength is not that it defeats classical simplex or interior-point methods on every small problem. Its strength lies in the combination of feasible geometric approximation, robust residual polishing, algebraic certification, and exploitable structure.

8 Scientific Computing

The following section presupposes Set Theory and Nonstandard Analysis.

Definition: Let $\varepsilon \in [\tilde{\nu}, 1-\tilde{\nu}]$. Let according to the *trapezoidal rule* ${}^{T}\!\uparrow_{z\in A} f(z)\downarrow z := \mathop{+}_{z\in A} \tilde{2}(f(z)+f(\vec{z}))(\vec{z}-z)$ and according to the *midpoint rule* ${}^{M}\!\uparrow_{z\in A} f(z)\downarrow z := \mathop{+}_{z\in A} f(\tilde{2}(z+\vec{z}))(\vec{z}-z)$. For $j,k \in {}^{\omega}\mathbb{N}, f \in C_{\pi}^{j+2}$ and *Fourier coefficients* $c_k := \tilde{\hat{\pi}}\uparrow_{-\pi}^{\pi} f(t)\tilde{\epsilon}^{k\underline{t}}\downarrow t$, the series $\mathop{+}_{k=-\omega}^{\omega} c_k \epsilon^{k\underline{t}}$ is called *Fourier series*, with possibly other *period lengths* than $\hat{\pi}$ (see [30], p. 358 - 364). Let $f_n^*(z) = f(\eta_n z)$ *sisters* of the TS $f(z) \in \mathcal{O}(\mathbb{D})$ centred on 0 on the domain $\mathbb{D} \subseteq {}^{\omega}\mathbb{C}$ where $m,n \in {}^{\omega}\mathbb{N}^*$ and $\eta_n^m := \underline{1}^{2^{\lceil m/n \rceil}}$. Then let $\delta_n^* f = \check{f} - \check{f}_n^*$ the *halved sister distances* of f. Let $s \in {}^{\omega}\mathbb{Z}^*, \mu_s^m := z^{-s}m!/(m-s)!$ and $\widetilde{|s|\text{-}}! := 0$. On the level of TS, μ and η form a *μ-η-calculus*.△

Remark: The latter allows representing integrals and derivatives easily and finitely closed (cf. [22], p. 165 f.). It holds for the TS $f(x), g(x) \in {}^{\omega}\mathbb{R}$ e. g. that ${}^{2}f(x)\uparrow_0^x\uparrow_0^y g(v)\downarrow v\downarrow y = f(x\mu_2)g(x\mu_{-2})$.

Example: Re $c \in [\tilde{\nu}, 1+\tilde{\nu}[, c \in {}^{\omega}\mathbb{C}$ and $\varepsilon := \tilde{2}^j, j \in {}^{\omega}\mathbb{N}^*$ imply $\zeta(n+c) = a_n \mathop{+}_{k=1}^{n} \delta_n^* v_c(\varepsilon u^k)$ for $z \in {}^{1-\tilde{\nu}}\dot{\mathbb{C}} \subset D$ (cf. [23], p. 42) and $v_c(z) := \mathop{+}_{m=1}^{\omega} \zeta(m+c)z^m = z\mathop{+}_{m=1}^{\omega} \tilde{m}^c \widetilde{z-m}$.□

Example: A Richardson extrapolation determines the digamma function ψ and

$$\zeta(\grave{n}) = \widetilde{\varepsilon^n \hat{n}} \mathop{+}_{k=1}^{n} (\psi(\varepsilon u^k \underline{1}^{2\tilde{n}}) - \psi(\varepsilon u^k)) + \mathcal{O}(\varepsilon^{\hat{n}})$$

for $n = 2$ and $\varepsilon = \tilde{2}^{13}$ as $\zeta(3) = 2^{24} \mathop{+}_{k=1}^{2} (\psi(\underline{\varepsilon} u^k) - \psi(\varepsilon u^k)) + \mathcal{O}(10^{-16})$.

Representation theorem for integrals: The TS (see below) $f(z) \in \mathcal{O}(\mathbb{D})$ centred on 0 for $\mathbb{D} \subseteq {}^{\omega}\mathbb{C}$ gives for $\grave{m}, n \in {}^{\omega}\mathbb{N}^*$

$$\uparrow_0^z \ldots \uparrow_0^{\zeta_2} f(\zeta_1)\downarrow\zeta_1 \ldots \downarrow\zeta_n = f(z\mu_{-n}).\square$$

Representation theorem for derivatives: For ${}^{\tilde{\nu}}\dot{\mathbb{C}} \subset \mathbb{D} \subseteq {}^{\omega}\mathbb{C}$, n-th unit roots, TS

$$f(z) := f(0) + \mathop{+}_{m=1}^{\omega} \widetilde{m!}\, {}^{m}f(0)z^m,$$

$\varepsilon := \tilde{2}^j\tilde{r}, j \in {}^{\omega}\mathbb{Z}, n = \epsilon^{\sigma} \in {}^{\omega}\mathbb{N}^*, u := \epsilon^{\tilde{n}\underline{\hat{\pi}}}$ and f's radius of convergence $r \in {}^{\nu}\mathbb{R}_{>0}$ imply

$${}^{n}f(0) = 2^{jn}\acute{n}! \mathop{+}_{k=1}^{n} \delta_n^* f(\tilde{2}^j u^k).\square$$

Universal multistep theorem: For $n \in {}^{\nu}\mathbb{N}_{\leq p}, k, m, p \in {}^{\nu}\mathbb{N}^*, \downarrow\vec{x} \in]0,1[, x \in [a,b] \subseteq {}^{\omega}\mathbb{R}, y : [a,b] \to {}^{\omega}\mathbb{R}^q, f : [a,b] \times {}^{\omega}\mathbb{R}^{q\times n} \to {}^{\omega}\mathbb{R}^q, g_k(\vec{x}) := g_{\acute{k}}(x)$, and $g_0(a) = f(\bar{a}, y_0, \ldots, y_{\acute{n}})$, the TS of the initial value problem ${}^{1}y(x) = f(x, y((\rightharpoonup)^0 x), \ldots, y((\rightharpoonup)^{\acute{n}} x))$ of order n implies

$$y(\vec{x})_p = y(x)_p + \downarrow\vec{x} \mathop{\pm}_{k=1}^{p} \Big(g_{p-k}(\vec{x}) \mathop{+}_{m=k}^{p} \widetilde{m!\tbinom{\acute{m}}{\acute{k}}}\Big) + \mathcal{O}((\downarrow\vec{x})^{\grave{p}}).\square$$

Theorem for (anti-) derivatives of crown series (CS): For $m \in {}^{\omega}\mathbb{Z}$, $q = \tilde{\varepsilon}(z-a)$, $a \in \mathbb{D}$ and $j,k \in \mathbb{N}_{<n}$, modular arithmetic (cf. [16], p. 302 - 311) and n-th unit roots result in the analogous TS:

$${}^{m}f_{\tilde{n}}(z) := \tilde{n}(q^j)^{\top}(\delta_{jk}\widetilde{\varepsilon q}^{j} j!/(j-m)!)(\tilde{u}^{jk})(f(\varepsilon u^k + a)) + R_n^m(z).\square$$

Reversion theorem of CS: For $y \in f(\mathbb{D}), y(a) = b$ and ${}^{1}y(a) \neq 0$, Bürmann's theorem yields ([32], p. 129 ff.):

$$f_{\tilde{n}}^{-1}(y) := a + \tilde{n} \mathop{+}_{m=1}^{n} \tilde{m}\tilde{\varepsilon}^{\acute{m}}(y-b)^m(\tilde{u}^{\acute{m}k})^{\top}(f(\varepsilon u^k + a)^{-m}) + R_n^-(y).\square$$

Remark: Put $y = 0$ for a sufficiently small $\tilde{n}|a - z_0|$ such that $z_0 \in {}^{\omega}\mathbb{C}$ is a zero of $f(z) \in {}^{\omega}\mathbb{C}$. The previous theorem allows to easily transform implicit differential equations for CS into explicit ones.

Theorem for derivatives of Fourier series: For $f \in C_{\pi}^{m+2}$ (cf. [30], p. 358 ff.), $m \in {}^{\omega}\mathbb{N}, j \in [-\acute{n}, n] \cap \mathbb{Z}, t \in [-\pi, \pi]$ and $k \in \mathbb{N}_{<\hat{n}}$, the following CS analogue is implied allowing integration of individual terms:

$${}^{m}\check{f}_{\tilde{n}}(t) := \tilde{n}(u^{\tilde{\pi}jnt})^{\top}(\delta_{(j+\acute{n})k}\underline{j}^m)(\tilde{u}^{jk})(\check{f}(\tilde{n}\pi k - \pi)) + \mathcal{O}(\tilde{n}).\square$$

Conclusion: Supporting points $kr := \tilde{n}\pi k$ of the smooth $f(kr)$ yield for j like k interpolating in $\mathcal{O}(\sigma n)$:

$${}^{m}\check{f}_{\tilde{n}}(t) := \tilde{n}(u^{\tilde{r}jt})^{\top}(\delta_{jk}\underline{j}^m)(\tilde{u}^{jk})(\check{f}(kr)).\square$$

Remark: The identity instead of δ_n^* provides arbitrarily precise approximations for the ${}^n f$. The error is comparable for analogously defined m-dimensional CS with $\binom{m+n}{n}$ derivatives.

Theorem (fixed-point method for CS): Every $z \in {}^{\omega}\mathbb{C}$ of an arbitrary m-polynomial $p(z) = 0$ with $m \in [2, n] \cap \mathbb{N}$ for $n := 2^r, r \in {}^{\nu}\mathbb{N}^*$ and coefficients from ${}^{\nu}\mathbb{C}$ can be (initiating) determined also in $\mathcal{O}(\sigma n)$.

Proof and algorithm: Let $U = (\tilde{u}^{jk})$ for $j, k \in \mathbb{N}_{<n}, u := \epsilon^{\check{n}\hat{\pi}}, q := \hat{z}$ and $s_k := p(\tilde{2}u^k)$. A simple transform achieves $|q| < \tilde{2}$ for all zeros ζ of $p(z)$ and $p(0) = 1$. It follows from $p(z) = \tilde{n}(q^j)^{\top}Us = \tilde{n}\mu^{\top}s = 0$ the simplified iteration $\mu^* = U_1^{-T}\mu U((\delta_{jk}\tilde{u}^j)U^{-1}\mu - (U_{\check{n}}^{-T}\mu + \beta s^{\top}\mu, \beta s^{\top}\mu, ..., \beta s^{\top}\mu)^{\top})$. Here let δ_{jk} the Kronecker delta, starting point $q := \tilde{2}$ and $\beta \in {}^{\nu}\mathbb{C}^*$ such that each time $||\mu^* - \mu||$ is roughly halved and $\mu^{\top}s = 0$ holds. Finish by polynomial division where $m > 2$.□

Approximation theorem: The ${}^s f(x) \in {}^{\omega}\mathbb{R}$ for $x \in {}^{\omega}\mathbb{R}$ as above allow computing the interpolant

$$g(x) := \sum_{r=0}^{\acute{m}} \chi_{]x_r, x_{\acute{r}}[}(x)((x_{\acute{r}} - x)p_r(x) + (x - x_r)p_{\acute{r}}(x))/(x_{\acute{r}} - x_r) + \sum_{r=0}^{m} \chi_{\{x_r\}}(x)p_r(x)$$

for $m, n \in {}^{\nu}\mathbb{N}$ and $p_r(x) := \sum_{s=0}^{n} {}^s f(x_r)(x - x_r)^s/s!$ in $\mathcal{O}(\sigma mn)$ where ${}^s f(x_r) = {}^s g(x_r)$ holds for every $x_r \in {}^{\omega}\mathbb{R}$. Replace in the complex case ${}^{\omega}\mathbb{R}$ by ${}^{\omega}\mathbb{C}$ and put $x = \gamma(t) \in {}^{\omega}\mathbb{C}$ for the path $\gamma(t)$ where $t \in {}^{\omega}\mathbb{R}$.□

Theorem (differential equations): Let $d(z) \in {}^{\omega}\mathbb{C}^{n+2}$ vector of CS for $n \in {}^{\omega}\mathbb{N}$. The solution y_0 of the system ${}^{\acute{n}}y = \left(\complement_0^n {}^m y, 1\right) d(z) \in {}^{\omega}\mathbb{C}$ where (not) $\left(\complement_0^n {}^m y(a)\right)^{\top} \in {}^{\omega}\mathbb{C}^{\acute{n}}$ and $a, z \in {}^{\omega}\mathbb{C}$ can be determined alike in $\mathcal{O}(n^3)$.□

Remark: If function values of *crown points* $\varepsilon u^m + a$ around the *crown centre* a are evaluated, where $\varepsilon = \tilde{2}$ and $n = 64$ are well-chosen, and terminating CS (possibly by using Laurent and Fourier series) are used, both a polynomial division may be executed and (at least in sections) analytic, nonlinear or partial (integro-) differential equation systems can be solved.

Example: For $b = (f(\varepsilon u^0 + a), ..., f(\varepsilon u^{\check{n}} + a), 1)^{\top} \in {}^{\omega}\mathbb{C}^{\acute{n}}$ and an upper triangle matrix $T(z) \in {}^{\omega}\mathbb{C}^{\acute{n}\times\acute{n}}$, ordinary differential equations (e. g. in [31]) have only solutions if $T(z)b = 0$ holds.

Series theorem for integrals: For $c \in [0, a], f \in {}^{\nu}C^n$ and $n, \hat{p}, r (= 2^p), t (= 2^v), \hat{v} \in {}^{\nu}2\mathbb{N}^*$, CS show

$$\uparrow_0^a f(w)\downarrow w = \uparrow_0^{\tilde{t}a} \sum_{s=1}^{\top} f(x + \acute{s}\tilde{t}a)\downarrow x = \uparrow_0^1 g(y)\downarrow y = \sum_{\check{q}=1}^{\check{r}} \sum_{\check{m}=0}^{\check{n}-1} \widetilde{\acute{m}!}\, {}^m\hat{g}(\tilde{r}\acute{q})\, \tilde{r}^{\acute{m}} + \mathcal{O}\left(\widetilde{\acute{n}!}\, {}^n g(\tilde{a}c)\tilde{r}^n\right).□$$

Remarks: CS equivalents may replace any ${}^m g$ for $g(y) := af(ay)$ and $a > 0$. Transforming $z := \epsilon^{\pm x}$ makes here infinite bounds of integration finite. A finite integral decreases the remainder's modulus sufficiently. The midpoint rule is here more advantageous than the trapezoidal rule.

Example: For ${}_{\mp}g(\vartheta) := \check{\pi}(1 - \varepsilon^2\sin^2(\check{\pi}\vartheta))^{\mp\tilde{2}}$, complete elliptic integrals of the first and second kind are

$$\uparrow_0^1 {}_{\mp}g(\vartheta)\downarrow\vartheta = {}_{\mp}g(\check{1}) + \widetilde{24}\, {}^2_{\mp}g(\check{1}) + \widetilde{1920}\, {}^4_{\mp}g(\check{1}) + \mathcal{O}\left(\widetilde{9!}\, {}^6_{\mp}g(\check{1})\right).$$

Second Euler-Maclaurin formula: For $f(\check{q}) = g(\tilde{r}\acute{q})$ and $k = \grave{a} = \check{r}$, the preceding theorem yields in $\mathcal{O}(\sigma n)$

$$\sum_{\check{q}=1}^{\check{r}} f(\check{q}) = \uparrow_{\check{1}}^{\check{k}} f(x)\downarrow x + \sum_{\check{m}=1}^{\check{n}-1} H_m\left({}^{\acute{m}}f(\check{k}) - {}^{\acute{m}}f(\check{1})\right) + \mathcal{O}\left(H_n\left({}^{\acute{n}}f(\check{k}) - {}^{\acute{n}}f(\check{1})\right)\right).$$

Proof: For $h(x) = x/\sin x$ and $H_m := \underline{1}^m\widetilde{m!}\, {}^m h(0) = \widetilde{m!}B_m(2 - 2^m) \rightarrow 2\underline{\tilde{\pi}}^{-m}$,

$$\sum_{\check{q}=1}^{\check{r}} g(\tilde{r}\acute{q}) = \check{r}\uparrow_0^1 g(y)\downarrow y - \sum_{\check{q}=1}^{\check{r}} \sum_{\check{m}=1}^{\check{n}-1} \widetilde{\acute{m}!}\, {}^m g(\tilde{r}\acute{q}) + \mathcal{O}\left(\widetilde{\acute{n}!}\, {}^n g(\tilde{a}c)\tilde{r}^n\right)$$

results in the claim by inserting and combining the factorials that are counted depending on partitions where B_m (cf. Nonstandard Analysis) are Bernoulli numbers and $B_{\check{m}} = 0, B_0 = 1$ as well as $B_1 = -\tilde{2}$.□

First Euler-Maclaurin formula: Mathematical induction for n shows also that (cf. [20], p. 193 f.)

$$\sum_{\check{q}=0}^{\check{r}} f(\check{q}) = \uparrow_0^{\check{r}} f(x)\downarrow x + \check{f}(\check{r}) + \check{f}(0) + \sum_{\check{m}=1}^{\check{n}-1} \widetilde{m!}B_m\left({}^{\acute{m}}f(\check{r}) - {}^{\acute{m}}f(0)\right) + \mathcal{O}\left(\widetilde{n!}B_n\left({}^{\acute{n}}f(\check{r}) - {}^{\acute{n}}f(0)\right)\right).□$$

Remark: Replace derivatives of the Euler-Maclaurin formulas by their CS equivalents, too.

Algorithmic Refoundation: Extraction of Linear Factors Instead of Blind Evaluation

Standard approximation algorithms often operate in a structurally blind manner. They evaluate the target function at discrete nodes or calculate local derivatives without questioning their global topology. If a classical algorithm encounters a pole, it numerically interprets it simply as an extremely steep ascent. The result is a procedural breakdown: linear systems become singular, rounding errors escalate uncontrollably into `NaN` (Not a Number) values, and the resulting polynomial loses its mathematical validity.

The approach pursued here abandons this purely reactive path and implements a **proactive structural extraction** (into linear factors). The computational process is thereby divided into three strictly separated phases:

1. **Structural Scan:** Prior to any series expansion, a preliminary algorithm actively searches for true singularities and roots. Through the rigorous evaluation of the modulus of complex numbers, singularities on the imaginary axis are also identified flawlessly.

2. **Analytical Extraction of Linear Factors:** The recognised algebraic structures are extracted directly from the original function. Poles are multiplied out, and roots are divided out. What remains is a mathematically smoothed, analytical residual function that avoids numerical escalation.

3. **Approximation and Reconstruction:** The actual approximation tools are applied only to this stable residual function. Since the behaviour of the residual function is generally well-behaved, significantly lower polynomial degrees are sufficient for the approximation. In the final reconstruction step, the previously isolated algebraic poles and roots are exactly and losslessly reintegrated into the numerator and denominator polynomials of the final model.

The Blind Spot of the Real Axis: Complex Near-Singularities

An one-dimensional scan along the real axis is highly efficient, yet it possesses a fundamental blind spot: **complex near-singularities**. Let the rational function $f(z) = \frac{z^2+\epsilon^2}{z^2+\delta^2}$ be considered for very small, real values of ϵ and δ.

On the real axis, this function is continuous, smooth, strictly positive, and completely regular. A purely real structural search algorithm finds no anomalies. Nevertheless, purely complex roots at $\pm i\epsilon$ and poles at $\pm i\delta$ lie hidden in immediate proximity to the real axis.

These hidden poles act as analytical barriers. They limit the radius of convergence of classical power series to $R = \delta$, which dooms any global approximation beyond this narrow interval to failure. To ensure truly robust extraction of linear factors, structural analysis must therefore not be restricted to the mathematical ideal line of the real axis. It is imperative that it is expanded into a two-dimensional *tolerance tube* (strip) in order to capture those complex singularities whose sphere of influence reaches deep into the real domain.

This paradigm shift – the methodological separation of the defining algebraic structure and the analytical remainder – protects the system at a fundamental level from numerical overfitting. The computation avoids simulating singularities through error-prone oscillations; instead, the model adapts from the ground up to the true, internal structure of the function.

ODE example: Crown series and convolution vs. Runge–Kutta

As a deliberately simple, but nonlinear, reference problem, consider

$$ {}^{1}y(x) = y(x)^2, \qquad y(0) = 1. $$

The exact solution is

$$ y(x) = \widetilde{\acute{x}}\text{-}, $$

with a pole at $x = 1$. Hence, the radius of convergence of the TS about x_0 equals $R(x_0) = \acute{x}_0\text{-}$.

Crown series as a Taylor expansion

For a step from x_0 to $x_0 + h$, the local series expansion

$$y(x_0 + s) = +^{n}_{m=0} a_m\, s^m \; + \; O(s^{\hat{n}})$$

is used (here $s = x - x_0$). Then

$$^1y(x_0 + s) = +^{\hat{n}}_{m=0} \grave{m}\, a_{\grave{m}}\, s^m \; + \; O(s^n).$$

For the square, the Cauchy product (convolution)

$$y(x_0 + s)^2 = \left(+^{n}_{m=0} a_m s^m\right)\left(+^{n}_{m=0} a_m s^m\right) = +^{\hat{n}}_{m=0} c_m s^m$$

is obtained with

$$c_m := +^{m}_{j=0} a_j\, a_{m-j}.$$

Matching coefficients in $^1y = y^2$ yields, for $m = 0, \dots, \hat{n}$, the recursion

$$\grave{m}\, a_{\grave{m}} = c_m = +^{m}_{j=0} a_j\, a_{m-j}, \qquad a_0 = y(x_0).$$

Thus, the a_m are determined entirely by convolutions and contain no binomial coefficients.

Convolution via fast Fourier transform (FFT) (core idea)

For the discrete convolution, vectors are used. Let $A = (a_0, \dots, a_n)$. For a fast convolution, use zero-padding to length $\acute{m} \geq \hat{n}$ (typically $m = 2^{\lceil 2(\hat{n}+1) \rceil}$) and the DFT (see [26], p. 155 - 161):

$$\breve{A}_k := +^{\acute{m}}_{j=0} A_j\, u^{jk}, \qquad u := \exp(\underline{\hat{\pi}}\tilde{m}), \qquad k = 0, \dots, \acute{m}.$$

Then, for the convolution coefficients $C = A * A$, componentwise,

$$\check{C}_k = \breve{A}_k^2, \qquad C_j = \tilde{m}\, +^{\acute{m}}_{k=0} \check{C}_k\, \tilde{u}^{jk}.$$

The FFT realises these transforms in $O(m\ {}_2 m)$. Important: this is precisely where the CS becomes practical, since derivatives and products can be handled vectorially – and hence in an FFT-ready manner.

Evaluating the step

Once $A = (a_0, \dots, a_n)$ has been determined, the step value follows by polynomial evaluation

$$y(x_0 + h) \approx +^{n}_{k=0} a_k\, h^k,$$

e.g. via Horner's scheme (multiplications/additions only).

Concrete computation for $x_0 = 0$, $h = 0.1$, $n = 16$

Here the exact solution is $y(x) = \widetilde{x}\text{-}$, hence

$$a_k = 1 \quad (k \geq 0),$$

because

$$\widetilde{s}\text{-} = +^{\omega}_{k=0} s^k.$$

Thus, the n-th partial sum is

$$y(0.1) \approx +^{16}_{k=0} (0.1)^k = \frac{1 - (0.1)^{17}}{1 - 0.1}.$$

The exact value is

$$y(0.1) = \frac{1}{0.9} = 1.\overline{1}$$

and the remainder term equals

$$\left| y(0.1) - +^{16}_{k=0} (0.1)^k \right| = \frac{(0.1)^{17}}{0.9} \approx 1.11111111111111111 \cdot 10^{-17}.$$

Comparison: classical fourth-order Runge–Kutta

For $^{1}y = y^2$, one RK4 step (see [26], p. 358) with $y_0 = y(x_0)$, step size h is

$$\begin{aligned}
u_1 &= y_0^2, \\
u_2 &= \left(y_0 + h\check{u}_1\right)^2, \\
u_3 &= \left(y_0 + h\check{u}_2\right)^2, \\
u_4 &= \left(y_0 + hu_3\right)^2, \\
y_1 &= y_0 + \tilde{6}h\,(u_1 + \hat{u}_2 + \hat{u}_3 + u_4).
\end{aligned}$$

With $x_0 = 0$, $y_0 = 1$, $h = 0.1$ numerically

$$y_{\mathrm{RK4}}(0.1) \approx 1.1111104900521944,$$

is obtained and therefore

$$\left|y(0.1) - y_{\mathrm{RK4}}(0.1)\right| \approx 6.21 \cdot 10^{-7}.$$

Fair comparison (principle)

- Accuracy target: For high accuracies (e.g. 10^{-16}, 10^{-34}), RK4 must reduce the step size substantially or increase the order; otherwise the local truncation term dominates.
- Crown/convolution: One step can be chosen much larger (for an analytic solution) as long as $h < R(x_0)$, with accuracy controlled primarily by n (series depth) and working precision.
- Runtime: RK4 costs a few function evaluations per step, contrary to crown/convolution needing the computation of many coefficients (products/convolutions) plus one polynomial evaluation. The advantage arises when one crown step replaces many RK steps and when the convolutions are organised via FFT – especially for large n. Furthermore, a complete function is represented.

Remark on "continued differentiability"

For analytic right-hand sides, the series method systematically leads to high derivative orders, since

$$^{k}y(x_0) = k!\,a_k$$

and a_k are generated by convolutions/compositions. Conceptually, this is closer to "differentiating by computation" (vector arithmetic/FFT) than to quadrature with a fixed node scheme.

9 Theoretical Informatics

The following assumes set theory and answers for $k, m, n \in {}^{\nu}\mathbb{N}^*$ the questions: How can $n = \epsilon^{\sigma} \leq k^m$ values encoded in a positional numeral system to the base k be sorted in $O(1)$? How quickly does deterministic parallelisation more practically sort arbitrary-precision numbers? How can two matrices $\in {}^{\nu}\mathbb{R}^{n \times n}$ be multiplied in $O(\sigma)$? How can the minimum or maximum spanning tree of a graph with n nodes be calculated in $O(\sigma)$? What is the answer to the halting problem? Why does $P = NP$ hold?

Let the memory be partially organised in the form of a binomial tree of order k with $k^{\grave{m}}$ memory locations, where $k^{\grave{m}} = 2^h$ with $h \in {}^{\nu}\mathbb{N}$. Let each memory location be capable of fully accommodating a value to be sorted, and let each edge correspond to a data line. An index sequentially appended to multiple identical values forms the new value together with the old one. Each j-th intermediate node corresponds hierarchically to the j-th digit to the base k of the value to be sorted.

If the leaves store their addresses instead of the values (in one step), a further step allows them to be retrieved sorted into the root and pushed back down the tree at most until they separate. If a value passes through an intermediate node, the latter sets its bit to 1, but to 0 if the intermediate node has no successor to be read. The sorting order determines the first value.

The subsequent architecture with autonomous memory units possesses simple computational capabilities. Interconnected distributed computers are able to simulate it. The complexity increases by $O(\sigma)$ when inserting an additional value, as this corresponds to a binary search. It is inversely proportional to the number of tree nodes in the memory. Using $d \in {}^{\nu}\mathbb{N}^*$ data lines reduces it by the factor $\tilde{d}$; occupying g times a memory location increases it by the factor g.

If the memory does not contain a binomial tree of order k, it increases by the factor ${}_kn$, because each intermediate node must be read for a value. Using the first k^p memory locations otherwise, with $p \in {}^{\nu}\mathbb{N}$, requires searching all intermediate nodes of the p-th level. The complexity increases marginally if the memory is well-filled, otherwise by the worst-case factor $O(k^p)$.

The matrices to be multiplied are to be stored square-wise n times each in a memory cube such that each memory unit multiplies two values and each entry of the product matrix is determined as a sum with n addends in $O(\sigma)$ (vector processor). Matrix multiplication can also be realised via addressed swarming entries. Tensor products can be efficiently mapped by n-memory cubes, which remain in three-dimensional space via corresponding data lines.

The optimal sorting algorithm *Bitsort* sorts stably with a time complexity of $O(n)$ and requires $O(n)$ additional memory space. Prerequisites: Any two values exist as a bit sequence, can be compared in $O(1)$, and all their bits can be processed likewise. Furthermore, let the memory be partially realised as a binary tree, and the path from root to leaf can also be traversed in $O(1)$.

If $p \in {}^{\nu}\mathbb{N}^*$ is the number of parallel steps for parallel Bitsort, the time complexity is $O(\tilde{p}n)$. Each insertion into the constantly sorted tree or output of a value has a time complexity of $O(1)$. Let the memory be equipped with the following intelligence: Each tree node independently stores the value reaching it and can pass it on to the next tree node if it has already been visited, or otherwise return its address to the central processing unit. No bit value corresponds to the root.

It begins with the most significant bit of a value. Each of its sequentially processed bits determines the next node within the binary tree. If all bits have been processed or a difference to the compared bit arises and the corresponding node has not yet been visited, the subsequent node stores, if applicable, the address of the remaining bit chain of the value not located in the tree, or the number of values for this node if the order of identical values is immaterial.

The time complexity of all methods sorting values of bit length $b = qa \in {}^{\nu}\mathbb{N}^*$ with address length $a \in {}^{\nu}\mathbb{N}^*$ increases per value from $O(1)$ to $O(q)$. If r is the average value of q, at least one $O(n)$ in the time complexity of the sorting method must be replaced by $O(rn)$. If $s \in {}^{\nu}\mathbb{N}^*$ is the average number of distinct (!) values that land in a bucket after sorting up to bit length a and whose bit length is greater than a, the time complexity of the method is $O(({}_{\epsilon}s + r)n)$.

Thereby, the mentioned values of a bucket are stored per bucket in an AVL tree, or in a B-tree if stored in an external memory that is relatively slow compared to the main memory. There, only the significant sub-value of the address length of a value is processed at a time. Foregoing fully sorted values after processing each value reduces the overall sorting time to $O(rn)$ by using two main memory areas.

The second main memory area is cleared after each pass to be available for the next pass, until a bucket has been processed. Only the addresses in the current sorting order are consistently written back into the first main memory area before the values with the same recently processed sub-value are further sorted. For rapid access to each sub-value, every value must be stored contiguously in the memory.

Using the $O(rn)$ method for master data and the $O(({}_{\epsilon}s + r)n)$ method for transactional data combines both well. This allows, in particular, table indices to be created quickly. For extensive tables, it makes a significant difference whether an index is created in $O(rn)$ instead of $O(r\sigma n)$ and a value can then be searched for in $O({}_{\epsilon}s + r)$ on average instead of $O(\sigma + r)$ or worse. Sorting by merging sorted lists is also possible in $O(rn)$.

Discussion of Bitsort: It is optimal under the specified prerequisites and well-suited for parallel programming, the rapid ongoing construction of table indices in databases, and searching within them. Its hardware implementation is rather unconventional. While it remains more theory than practice, the reverse is generally true for the following method.

Square Sort: In computer-assisted nonstandard mathematics, traditional sorting paradigms encounter physical boundaries. The processing of highly complex, pointer-based data types with extreme precision is not primarily limited by the number of mathematical comparisons, but by memory bandwidth, latencies during pointer dereferencing, and the overhead of dynamic memory management. Its time complexity is $O(N\ {}_2N)$, and is space complexity is $O(N)$.

It combines static memory management with asymmetric work-stealing scheduling and significantly outperforms classic generic methods (such as Timsort or conventional parallel MergeSort implementations) for complex data types, particularly with pseudo-random numbers. The exact runtimes and performance proofs can be found below. An internal porting to C++ yielded comparable results.

Classic merge approaches scale excellently in the lower tree but force the hardware into a single-thread bottleneck during the final merging step. Square Sort instead uses a parallel divide-and-conquer merge. Through the strategic search for medians and binary localisation within the sub-vectors, the remaining solution space is continuously partitioned. Each asymmetric sub-task is dynamically distributed to free processor cores.

Overcoming the limitations described by Amdahl's law permits hybrid chunking for radical cache locality. To minimise expensive main memory accesses when evaluating heap references, Square Sort delegates the base level to an in-place algorithm. The data blocks are dimensioned such that they remain in the CPU's core-exclusive L1/L2 cache. This eliminates the "cache thrashing" that inevitably occurs with traditional bottom-up methods.

While other algorithms place the garbage collector (GC) under massive pressure with millions of arbitrary-precision numbers, Square Sort permits high resource efficiency and zero GC overhead through static memory management. The temporal working buffer is $O(N)$. Alternating swaps of memory pointers at each recursion level save additional memory requirements and guarantee a linear, deterministic reliability without allocation spikes.

Discussion of Square Sort: It is not a generic all-rounder for primitive data types, but a precise mathematical instrument. It sacrifices the SIMD vectorisability of native machine numbers in favour of superior, fail-safe, and massively parallel scalability for the compute-intensive, one-dimensional structures of nonstandard mathematics, which has yet to establish itself. For this purpose, Square Sort provides a specialised, highly parallel sorting architecture.

Table 1: Runtime and allocations on AMD EPYC™ 9645 (24 cores, 128 GB DDR5): T_t (Timsort) compared to T_s (Square Sort) for sorted data with window 32 and a perturbation rate of 10 % (256-bit BigInt, 24 threads)

$n \times 10^5$	T_t (ms)	T_s (ms)	**Ratio**	A_t (MB)	A_s (MB)	**Ratio**
1	4.29	0.95	**4.51**	1.26	0.06	**20.08**
4	16.36	2.25	**7.27**	5.05	0.10	**50.89**
16	64.51	11.56	**5.58**	20.23	0.15	**139.34**
64	271.74	46.83	**5.80**	80.90	0.30	**267.73**
256	4250.16	377.37	**11.26**	323.60	0.87	**371.40**
1024	18984.41	1438.08	**13.20**	1294.67	3.08	**419.73**

The table dismantles the idealised assumption of Landau notation $O(N\ {}_2N)$. Its supposedly invariant proportionality factor C increases significantly in Timsort beyond the L3 cache threshold ($N = 12.8 \times 10^6$). Unpredictable pointer accesses oversaturate the L3 cache, causing C to escalate abruptly due to massive RAM latencies. Square Sort keeps C stable through strict locality: real-world scalability is thus determined not by asymptotic mathematics, but by physical memory management.

Let $n \in {}^{\nu}\mathbb{N}^*$ be the number of nodes in a finite, simple, connected graph. (If it is not simple, sorting out in $O(n)$ allows only the smallest parallel edge and no loops. If it is not connected, all isolated nodes are removed in $O(n)$.) Excluding negative weights, Dijkstra's algorithm requires $O(n)$ steps for each minimum path in the graph, since sorting as well as parallel addition and comparison are possible in $O(1)$.

Its minimum or maximum spanning tree can be calculated in $O(\sigma n)$ or even in $O(\sigma)$: The edges of a node can be sorted in parallel in $O(n)$ or $O(1)$, respectively. Let all edges be coloured white initially. Each node appends the smaller and then the larger number of the involved nodes of the edge with the smallest weight among all white-coloured edges originating from it to this, bit by bit, as a value. It stores the smallest value that reaches it. This is possible in $O(n)$ or $O(1)$, respectively.

Colouring the edges black between any two nodes of the same minimum spanning tree requires at most the same time. Each node then has a nearest neighbour node. Continuing with the white-coloured edges in $O(\sigma)$ steps can determine the smallest value of a minimum spanning tree in $O(n)$ or $O(1)$, respectively, until the entire minimum spanning tree is formed in $O(\sigma n)$ or $O(\sigma)$. For a maximum spanning tree, all edge weights are subtracted from a sufficiently large constant.

Halting Theorem: To reach a fixed r_m of the number $0.r_1r_2...r_m$, a programme halts for arbitrary (pseudo-)random numbers $r_j \in {}^{\nu}\mathbb{N}$ with $k \in {}^{\nu}\mathbb{N}^*$ digits and base $p := 2^k$ in $m \in \mathbb{N}^*$ steps with the probability $\tilde{p}$.

Proof: This is an optimal result of probability theory (cf. [16], pp. 10 - 40). □

Remark: Randomness is undecidable. Thus, programmes exist that halt with arbitrary probability, and Gödel's incompleteness theorems follow. Gödel's statement "I am not provable." is sophistic insofar as it exhibits the fallacious argument *petitio principii ex negativo*. For solving the halting problem, a self-reference that simultaneously commands and forbids halting is nonsensical. Self-reference renders the diagonal argument for the halting problem invalid.

Theorem: If $\mathcal{L}$ is the set of all (Type-0) languages L_k composed of finite words with a maximum length ν for mid-finite $|\mathcal{L}| = 2^{(\nu^{\grave{\nu}}-1)/\acute{\nu}}$ (applying the GS) and $k \in \{1, \ldots, |\mathcal{L}|\}$, then $P = NP = \mathcal{L}$ holds.

Proof: Since every deterministic Turing machine can also be conceived as non-deterministic (cf. [12], pp. 162, 318 and 361 f.), $P \subseteq NP$ holds. P contains all linear languages L_k, and NP contains only such languages. □

10 Theoretical Physics

The following assumes Scientific Computing.

Casimir effect

The corrected derivation of the Casimir effect (the physical reality)

The physical setup consists of two idealised conducting plates at a distance d. What is sought is not the absolute vacuum energy, but the energy difference between the vacuum with plates and the corresponding reference without plates. In the interior, the longitudinal modes are discretised by the boundary conditions; outside, or in the reference geometry, the corresponding continuous modes occur.

The hyperreal vacuum energy and the plasma frequency

In the discretised nonstandard space, the eigenfrequencies do not run into an indefinite infinity, but up to an exactly specified midfinite maximum frequency ω.

Real plates, however, are not perfect mirrors up to arbitrarily high frequencies. Above their characteristic material frequency, or plasma frequency, they become increasingly transparent. The physically appropriate target function therefore contains a damping factor $g(x)$, which is approximately 1 for small frequencies and vanishes at the upper midfinite boundary:

$$g(0) = 1, \qquad g(\omega) = 0.$$

For a smooth cut-off function, it is additionally assumed that the relevant derivatives also vanish at the upper boundary. Schematically, the longitudinal summation kernel is then

$$E_{\text{inside}} \propto +_{n=1}^{\omega} n^3 \, g(n).$$

In hyperreal space, this sum is not a mysterious negative quantity, but an exactly defined hyperfinite sum. The finite Casimir contribution arises only through the structured separation of the common background terms.

Structural deflation by Euler-Maclaurin

For the analysis, the first Euler-Maclaurin formula is used in its hyperreal extension. It reads:

$$+_{\check{q}=0}^{\check{r}} f(\check{q}) = \uparrow_0^{\check{r}} f(x) \downarrow x + \check{f}(\check{r}) + \check{f}(0) + +_{\check{m}=1}^{\check{n}-1} \widetilde{m!} B_m \left({}^{\acute{m}} f(\check{r}) - {}^{\acute{m}} f(0) \right) + O\left(\widetilde{n!} B_n \left({}^{\acute{n}} f(\check{r}) - {}^{\acute{n}} f(0) \right) \right).$$

Substituting

$$f(x) = x^3 \, g(x)$$

separates the relevant contributions into the integral term, the upper boundary, and the lower boundary.

Integral: The integral term

$$\uparrow_0^{\omega} x^3 \, g(x) \downarrow x$$

yields the dominant hyperreal background contribution. It belongs to the common vacuum core and is not, by itself, the measurable Casimir energy.

Upper boundary at ω: Since real plates become transparent at sufficiently high frequencies, the cut-off function is chosen so that, at the upper boundary,

$$f(\omega) = 0, \qquad {}^{1}f(\omega) = 0, \qquad {}^{2}f(\omega) = 0, \qquad {}^{3}f(\omega) = 0$$

holds. Hence the relevant Euler-Maclaurin boundary terms at the upper boundary vanish.

Lower boundary at 0: At the lower boundary, $g(0) = 1$ implies

$$f(0) = 0, \qquad {}^{1}f(0) = 0, \qquad {}^{2}f(0) = 0.$$

The third derivative, however, remains:

$${}^{3}f(0) = 6.$$

The finite regular contribution therefore arises from the fourth-order Bernoulli term:

$$R_{\text{constant}} = \widetilde{4!}\, B_4 \left({}^{3}f(\omega) - {}^{3}f(0)\right).$$

With

$$B_4 = -\widetilde{30}, \qquad {}^{3}f(\omega) = 0, \qquad {}^{3}f(0) = 6$$

it follows that

$$R_{\text{constant}} = \widetilde{4!} \cdot \left(-\widetilde{30}\right) \cdot (0 - 6).$$

Thus the longitudinal sum deflates into a dominant hyperreal background contribution and a positive, singularity-free remainder:

$$R_{\text{constant}} = \widetilde{120}.$$

The decisive point is this: the hyperfinite sum itself is not equal to $\widetilde{120}$. Rather, $\widetilde{120}$ is the regular remainder that remains after the common background structure has been separated off.

The exterior-space integral as reference

Outside the plates, or in the reference geometry, the mode spectrum is continuous. The energy density there is described by the corresponding nonstandard integral.

Since the same vacuum and the same physical high-frequency cut-off are used, the same dominant hyperreal pole is generated as in the interior. This contribution describes the common background energy and is not directly measurable.

Schematically, therefore,

$$E_{\text{inside}} = P_{\text{inside}}(\omega, d) + R_{\text{constant}},$$

$$E_{\text{outside}} = P_{\text{outside}}(\omega, d).$$

Under consistent regularisation, the common divergent contributions agree:

$$P_{\text{inside}}(\omega, d) = P_{\text{outside}}(\omega, d).$$

Transverse integration and the physical sign

The longitudinal constant

$$R_{\text{constant}} = \widetilde{120}$$

is positive. This, however, does not yet determine the sign of the physical Casimir energy.

In the real three-dimensional system, the continuous transverse momenta $k_\perp$ parallel to the plates must also be integrated. Only this phase-space integration supplies the physical prefactor of the mode difference.

Schematically, the transverse integration generates a negative geometrical factor in front of the regular longitudinal remainder. In the usual normalisation this gives

$$\frac{\Delta E}{A} = -\frac{\hbar c \pi^2}{720 d^3}.$$

The negative sign therefore does not arise from the longitudinal sum being negative. It arises from the complete three-dimensional mode calculation, in particular from the transverse phase-space integration and the subsequent subtraction of the reference geometry.

Exact algebraic cancellation of the background contributions

The physically measurable Casimir energy is the difference between the interior space and the reference space:

$$\Delta E = E_{\text{inside}} - E_{\text{outside}}.$$

Inserting the deflated expressions gives, schematically,

$$\Delta E \propto (P_{\text{inside}} + R_{\text{constant}}) - P_{\text{outside}}.$$

Since both calculations use the same vacuum, the same cut-off, and the same regularisation, the common hyperreal poles cancel exactly:

$$P_{\text{inside}} - P_{\text{outside}} = 0.$$

Only the finite regular remainder remains, multiplied by the physical geometrical prefactor:

$$\Delta E \propto -\widetilde{120}.$$

In full physical normalisation, the result per unit area is

$$\frac{\Delta E}{A} = -\frac{\hbar c \pi^2}{720 d^3}.$$

The associated force per unit area is

$$\frac{F}{A} = -\frac{\partial}{\partial d}\left(\frac{\Delta E}{A}\right) = -\frac{\hbar c \pi^2}{240 d^4}.$$

The negative sign means: the plates attract one another.

Conclusion of the derivation

The Casimir effect does not arise because

$$\mathop{\Large +}_{n=1}^{\omega} n^3 = -\widetilde{120}$$

were true. This statement is false in nonstandard space.

The hyperfinite vacuum sum remains positive and exactly determined. The decisive point is the structured deflation:

$$\text{hyperreal background} + \text{regular remainder}.$$

For the longitudinal cubic summation kernel, the lower boundary of the Euler-Maclaurin formula, together with the physical high-frequency cut-off, yields the positive regular remainder

$$R_{\text{constant}} = \widetilde{120}.$$

The physically negative sign of the Casimir energy then does not arise from a negative sum, but from the complete three-dimensional mode difference:

$$\frac{\Delta E}{A} = -\frac{\hbar c \pi^2}{720 d^3}.$$

Real materials additionally enforce this clean view, because they become transparent at very high frequencies. Nonstandard analysis thereby makes visible what is often hidden in the usual regularisation: the enormous hyperreal background contributions are not arbitrarily discarded, but are exactly balanced against one another under the same physical regularisation. What remains measurable is only the finite, distance-dependent remainder of the altered mode structure.

Electron self-energy

A nonstandard derivation of the electron self-energy (the physical interpretation)

In quantum electrodynamics (QED), the interaction of an electron with its own virtual photons contributes to its self-energy. In perturbative QED this contribution contains divergent terms. Established physics handles these terms by renormalisation: the bare parameters of the theory are related to the finite, experimentally measured quantities by a regularisation and renormalisation prescription.

In the present nonstandard formulation, the aim is not to introduce an undefined infinity. Instead, the divergent contribution is represented as an exactly specified hyperreal pole, which is then separated from the finite regular remainder.

The hyperreal momentum sum and the physical cut-off

The electron is not treated here as living in an unrestricted continuum, but in a discretised nonstandard space whose momenta extend up to an exactly specified midfinite maximum momentum ω. Such a scale may be interpreted as a physical or effective cut-off, for example near the Planck scale, although this interpretation is an additional modelling assumption.

The interaction of the electron with vacuum fluctuations is described by a sum over virtual momenta p. To express that the model is only trusted up to its effective resolution scale, the summand is multiplied by a smooth cut-off function or damping factor $g(p)$. For low momenta this factor is approximately 1; at the upper midfinite boundary it vanishes, together with the relevant derivatives:

$$g(0) = 1, \qquad g(\omega) = 0.$$

The dominant mass-renormalisation kernel behaves schematically like $\tilde{p}$ at large momentum. Thus the hyperfinite model sum has the form

$$E_{\mathrm{self}} \propto \boldsymbol{+}_{p=1}^{\omega} \tilde{p} g(p).$$

In the standard continuum limit this expression has logarithmic growth. In the hyperreal setting, however, it is an exactly defined hyperfinite quantity: very large, but not indeterminate.

Structural deflation by Euler-Maclaurin

To isolate the finite regular part, the target function

$$f(x) = \tilde{x} g(x)$$

is deflated by the Euler-Maclaurin formula. Since the kernel has a pole at $x = 0$, the sum is taken from 1 rather than from 0. This gives exactly

$$\boldsymbol{+}_{\check{q}=1}^{\check{r}} f(\check{q}) = \uparrow_1^{\check{r}} f(x) \downarrow x + \check{f}(\check{r}) + \check{f}(1) + \boldsymbol{+}_{\check{m}=1}^{\check{n}-1} \widetilde{m!} B_m \left({}^{\acute{m}} f(\check{r}) - {}^{\acute{m}} f(1) \right) + O\left(\widetilde{n!} B_n \left({}^{\acute{n}} f(\check{r}) - {}^{\acute{n}} f(1) \right) \right).$$

Here B_m are the Bernoulli numbers, and $\mathcal{R}$ denotes the corresponding remainder term.

Integral: the hyperreal pole The integral term

$$P_{\mathrm{self}} = \uparrow_1^{\omega} \tilde{x} g(x) \downarrow x$$

contains the dominant hyperreal background contribution. In the simplest undamped approximation it is proportional to ${}_e\omega$, and is therefore the nonstandard representative of the usual logarithmic divergence. By itself, however, this pole is not a directly measurable mass.

Boundary values at ω Because the cut-off function suppresses the spectrum at the upper boundary, the relevant Euler-Maclaurin boundary terms vanish there:

$$f(\omega) = 0, \qquad {}^{1}f(\omega) = 0, \qquad {}^{2}f(\omega) = 0, \qquad \dots$$

Boundary values at 1: the regular remainder The lower boundary at $x = 1$ avoids the singularity at $x = 0$ and supplies the finite boundary contributions. These terms depend on the chosen cut-off function and on the chosen renormalisation condition. They form a well-defined, finite, singularity-free remainder:

$$R_{\mathrm{mass}} = \check{f}(1) - +_{\check{m}\geq 1}\widetilde{m!}B_m{}^{\acute{m}} f(1) + \mathcal{R}_{\mathrm{finite}}.$$

Thus the self-energy is separated algebraically into

$$E_{\mathrm{self}} = P_{\mathrm{self}} + R_{\mathrm{mass}}.$$

The first term is the hyperreal pole; the second term is the finite regular contribution after the common divergent structure has been separated off.

The bare mass as a counterterm

In standard QED, the bare mass m_0 is not an independently observable mass. It is a parameter in the regularised theory. Its task is to absorb the divergent part of the self-energy so that the measured mass remains finite.

In the present nonstandard interpretation, this bare mass is represented by a hyperreal counterterm. It is not an arbitrary mystical infinity, but a precisely specified pole contribution fixed by the same cut-off structure:

$$m_0 = P_{\mathrm{bare}} + R_{\mathrm{bare}}.$$

The renormalisation condition requires that the pole in the bare mass cancel the pole in the self-energy:

$$P_{\mathrm{bare}} = -P_{\mathrm{self}}.$$

This equation is not an additional physical force law. It is the algebraic statement that the regularised theory is calibrated to the observed electron mass.

Exact algebraic cancellation and the physical mass

The physical electron mass is obtained from the bare mass plus the self-energy correction:

$$m_{\mathrm{phys}} = m_0 + E_{\mathrm{self}}.$$

Substituting the deflated quantities gives

$$m_{\mathrm{phys}} = (P_{\mathrm{bare}} + R_{\mathrm{bare}}) + (P_{\mathrm{self}} + R_{\mathrm{mass}})\,.$$

Using the cancellation condition

$$P_{\mathrm{bare}} = -P_{\mathrm{self}},$$

the hyperreal poles cancel exactly:

$$m_{\mathrm{phys}} = (-P_{\mathrm{self}}) + P_{\mathrm{self}} + R_{\mathrm{bare}} + R_{\mathrm{mass}}.$$

Therefore

$$m_{\mathrm{phys}} = R_{\mathrm{bare}} + R_{\mathrm{mass}}.$$

The measurable electron mass is thus not the hyperreal pole itself, but the finite remainder fixed by the renormalisation condition.

Interpretation

This nonstandard formulation replaces the informal language of divergent infinities by an exact hyperreal juxtaposition:

$$\text{bare mass} + \text{self-energy} = \text{pole} + \text{counter-pole} + \text{finite remainder.}$$

The pole and counter-pole cancel algebraically:

$$P_{\text{bare}} + P_{\text{self}} = 0.$$

What remains is finite and measurable:

$$m_{\text{phys}} = R_{\text{bare}} + R_{\text{mass}}.$$

The finite remainder is not obtained from the self-energy sum alone. It also depends on the renormalisation condition, that is, on how the theoretical parameters are matched to the experimentally observed electron mass.

Conclusion of the derivation

The electron self-energy need not be read as an undefined infinity. In the nonstandard formulation it becomes an exactly specified hyperreal pole plus a finite regular remainder.

The essential structural decomposition is

$$E_{\text{self}} = P_{\text{self}} + R_{\text{mass}}.$$

The bare mass supplies the corresponding counter-pole,

$$P_{\text{bare}} = -P_{\text{self}},$$

so that the physical mass remains finite:

$$m_{\text{phys}} = R_{\text{bare}} + R_{\text{mass}}.$$

Thus the renormalisation of QED can be interpreted as an exact algebraic cancellation of hyperreal pole terms under a common cut-off. The measurable electron mass is not a divergent quantity, but the finite remainder left after this cancellation and after the theory has been matched to experiment.

Feynman's path integrals

Feynman's path integrals as strictly hyperfinite sums

The path-integral formulation of quantum mechanics by Richard Feynman is physically very intuitive: a particle is not described by one single distinguished path, but by a superposition of possible paths through space and time. Mathematically, however, the continuous path integral is subtle. The formal expression contains a measure over an infinite-dimensional function space; an ordinary translation-invariant measure of the Lebesgue type does not exist there in naive form.

Established physics therefore works with limiting constructions, regularisations, discrete time slices, or with Wick rotation to the Euclidean theory. In the Euclidean formulation, many models can be treated rigorously by means of the Wiener measure and the Feynman-Kac formula.

In the discretised nonstandard space, this measure problem can be organised differently. The path integral is not regarded as a direct integral over an already given infinite-dimensional measure space, but as a hyperfinite sum, or as a hyperfinite product, over an extremely fine lattice. The formal expression thereby first becomes an exactly defined algebraic lattice quantity.

The macroscopically measurable part does not then arise by subsequently discarding hyperreal quantities. Rather, the hyperfinite quantity is normalised, expanded in the infinitesimal scales $\tilde{\omega}$ and $\tilde{\nu}$, and structurally deflated. The physical measured value is the regular scale-free part of this expansion.

The measure problem in the continuum

In the standard continuum, summation over all paths requires an integral over infinitely many degrees of freedom. The formal expression for the propagator is

$$K(b,a) \propto\uparrow \mathcal{D}x \exp(\tilde{\hbar}\underline{S}).$$

Here $K(b,a)$ denotes the propagator, that is, the complex transition amplitude from the initial point $a = (x_a, t_a)$ to the final point $b = (x_b, t_b)$. It is the kernel of quantum-mechanical time evolution. Here S is the classical action, $\hbar$ is the Planck constant, and $\mathcal{D}x$ formally denotes the measure over all paths.

In this form, the expression is not an ordinary Lebesgue integration. In particular, there is no naive translation-invariant Lebesgue measure on the infinite-dimensional path space that possesses all the desired properties at once. For this reason, actual calculations must be made precise by means of discrete approximations, limiting procedures, oscillatory integrals, distributions, operator methods, or Euclidean methods.

The hyperfinite discretisation of space and time

In nonstandard space, the time interval between the initial and final points is subdivided into an exactly specified midfinite number ω of time steps of infinitesimal duration Δt. Space is likewise modelled by a hyperfine, but strictly discrete, lattice.

A single path $x(t)$ is then no longer an arbitrary continuous functional object, but an exactly specified vector consisting of ω spatial coordinates. The space of all lattice paths is extremely large, but hyperfinite. In this lattice formulation, the formal path measure $\mathcal{D}x$ is replaced by a hyperfinite product of the individual lattice contributions:

$$\mathcal{D}x \to \bigtimes_{k=1}^{\omega} \Delta x_k.$$

It is important here that, in physical applications, this product usually also has to be supplemented by normalisation factors depending on mass, time step, dimension, and the chosen discretisation. The product therefore describes the structural measure kernel, not yet the fully normalised propagator formula.

The exact algebraic sum

Because of the lattice structure, there is no immediate need first to construct an infinite-dimensional Lebesgue measure. The path integral becomes a hyperfinite sum over the discrete set of all lattice paths P. The amplitude then has the schematic form

$$K(b,a) \propto +_P \exp(\tilde{\hbar}\underline{S}[P]).$$

Since each summand is a complex phase factor of modulus 1, each individual contribution is exactly defined. The hyperfinite sum is therefore algebraically well-defined. Its physical evaluation nevertheless still requires a suitable normalisation and the determination of the regular part.

It is therefore more accurate to say: the hyperfinite path integral avoids the naive infinite-dimensional measure problem, but it does not automatically replace all analytical questions. Rather, these are translated into questions of normalisation, expansion in the scales $\tilde{\omega}$ and $\tilde{\nu}$, structural deflation, and stability of the macroscopic limiting transition.

A typical normalised hyperfinite quantity has the form

$$X_{\omega,\nu} = P_{\omega,\nu} + R + I_{\omega,\nu}.$$

Here $P_{\omega,\nu}$ denotes the hyperreal pole or background part, R the regular scale-free part, and $I_{\omega,\nu}$ the infinitesimal remainder, consisting of positive powers of $\tilde{\omega}$ or $\tilde{\nu}$. The macroscopic measured value is then not an externally formed standard part, but the regular part R isolated by deflation.

Stationary phase and structural cancellation

How does classical physics arise from this enormous hyperfinite sum? The decisive mechanism is stationary phase. For macroscopic systems one typically has $S \gg \hbar$. The phase factor $\exp(\tilde{\hbar}\underline{S})$ then oscillates very rapidly for most neighbouring paths. In the hyperfinite sum, these contributions largely cancel by destructive interference.

In the neighbourhood of those paths for which the action is stationary, by contrast, the phase changes only slowly between neighbouring paths. There the contributions add constructively. These stationary paths satisfy the principle of least action, or more precisely the principle of stationary action.

In this view, the sum deflates structurally: the strongly oscillating non-stationary paths form a background whose contributions cancel one another in the macroscopic approximation. The leading contribution comes from the neighbourhood of the stationary paths. It does not necessarily mark exactly one single classical path, but rather the classical trajectory together with its local fluctuations.

Quantum fluctuations as the regular remainder

The neighbouring paths around the stationary trajectory do not cancel completely. They form the regular remainder of the summation. This remainder contains the quantum corrections to classical motion, the spread around the classical path, and contributions that become essential, for example, in the tunnel effect.

On a hyperreal lattice, such remainder contributions can be formulated as hyperfinite sums over discrete paths. In suitable Euclidean formulations they are closely related to random-walk models, Brownian motion, and the Feynman-Kac formula.

The Schrödinger equation is not necessarily the starting point of this representation. Rather, it can be understood as a macroscopic equation equivalent to the correctly normalised path integral. In this sense, the path-integral formulation connects the local differential equation with a global summation over paths.

Conclusion of the derivation

Feynman's path integrals need not be understood as naive integrals over an infinite-dimensional Lebesgue measure. Precisely this naive reading is mathematically problematic.

In nonstandard space, the path integral can be formulated as a hyperfinite algebraic summation over discrete lattice paths. This gives precision to Feynman's original intuition: the quantum amplitude arises from the superposition of many path contributions, while the classical trajectory is singled out by stationary phase and destructive interference of the non-stationary paths.

Nonstandard analysis does not solve the measure problem by producing an ordinary infinite-dimensional Lebesgue measure. It shifts the basis of the calculation to a hyperfinite lattice, on which sums and products are first algebraically exactly defined.

The physical theory then arises through normalisation, structural deflation, and the proof that the macroscopic limiting transition yields the known equations of quantum mechanics. The macroscopically measurable quantity is not obtained by subsequently discarding hyperreal parts, but by expansion in the infinitesimal scales $\tilde{\omega}$ and $\tilde{\nu}$. The regular part is the scale-free coefficient of this expansion, after the hyperreal poles have been structurally separated off.

Hawking radiation

Hawking radiation and the trans-Planckian problem

The theoretical discovery of Hawking radiation is regarded as one of the important milestones of modern theoretical physics, since it brings thermodynamics, quantum mechanics, and general relativity into contact within a common framework. Stephen Hawking's derivation shows that, in the semiclassical picture, black holes possess thermal radiation. At the same time, this derivation contains a well-known conceptual difficulty: the so-called trans-Planckian problem.

In the standard continuum, tracing later observable radiation modes backwards leads to extremely high frequencies near the event horizon. This is not a directly experimentally accessible statement, but a consequence of the idealised mathematical backward evolution of the modes. Nonstandard analysis can serve here as a precise language for treating such contributions not as indeterminate infinities, but as exactly specified hyperreal quantities, and for separating them structurally from finite remainder terms.

The trans-Planckian problem in the continuum

If one formally traces a thermal radiation quantum, registered far away from a black hole, backwards in time, its origin approaches the event horizon. Owing to the gravitational frequency shift, that is, the gravitational redshift, the same mode is described near the horizon by ever higher frequencies.

In standard analysis, this frequency may tend to infinity under idealised backward evolution. Formally, this means that the corresponding mode enters a regime in which its wavelength lies below the Planck scale. Yet precisely there it is no longer guaranteed that the semiclassical approximation consisting of classical spacetime and quantum field theory remains valid without modification.

The trans-Planckian problem is therefore not an immediate experimental contradiction, but an indication that the derivation depends sensitively on how the extremely short-wavelength degrees of freedom are treated.

The hyperfinite lattice and the constant ω

In the nonstandard analytical modelling considered here, there is no unbounded continuum in which frequencies can escalate without limit. The spacetime structure at the event horizon is described by a fine, but strictly hyperfinite, lattice. The maximum frequency is exactly specified and bounded by the midfinite constant ω.

In addition, it is assumed that vacuum modes at the Planck boundary no longer couple unchanged to the macroscopic curvature of spacetime. For this purpose, a smooth damping function $g(x)$ is introduced. For macroscopic frequencies it is approximately 1; at the upper boundary ω it falls smoothly to zero. Thus $g(0) = 1$ and $g(\omega) = 0$.

In this model, the energy density of the vacuum state at the horizon is written as a hyperfinite sum over discrete frequencies n:

$$E_{\text{horizon}} \propto +_{n=0}^{\omega} n \cdot g(n).$$

This representation is to be understood as a model kernel. It does not replace the full derivation of Hawking radiation via field modes, Bogoliubov coefficients, and the comparison of ingoing and outgoing vacuum states.

Structural deflation at the horizon

To separate the finite, physically relevant contribution from the background structure at the event horizon, the hyperreal Euler-Maclaurin formula is applied to the target function $f(x) = x \cdot g(x)$:

$$+_{\check{q}=0}^{\check{r}} f(\check{q}) = \uparrow_0^{\check{r}} f\,(x) \downarrow x + \check{f}(\check{r}) + \check{f}(0) + +_{\check{m}=1}^{\check{n}-1} \widetilde{m!} B_m \left({}^{\acute{m}} f(\check{r}) - {}^{\acute{m}} f(0) \right) + O\left(\widetilde{n!} B_n \left({}^{\acute{n}} f(\check{r}) - {}^{\acute{n}} f(0) \right) \right).$$

Here too, the problem separates into three essential structural components:

The integral: the hyperreal pole The integral term $P_{\text{vacuum}} = \uparrow_0^{\omega} x \cdot g(x) \downarrow x$ represents the dominant background contribution of the vacuum model directly at the event horizon. It is the part associated with the trans-Planckian problem in the standard continuum. In hyperreal space, however, this term is not an undefined infinity, but an exactly specified pole. By itself, it does not describe emitted Hawking radiation, but rather the not directly measurable background structure of the model.

The upper boundary at ω Since the spacetime lattice is cut off at the constant ω, and since $g(\omega) = 0$ is assumed together with the relevant derivatives, the corresponding boundary terms at the maximum frequency vanish. Thus $f(\omega) = 0$, ${}^{1}f(\omega) = 0$, ${}^{2}f(\omega) = 0$, and so on. The added term $\check{f}(\omega)$ is also zero.

In this modelling, the trans-Planckian problem is thereby regularised: the highest-energy boundary modes are not treated as physically freely extendable modes, but are removed from the measurable remainder by the cut-off.

The lower boundary at 0: the regular remainder Since $f(0) = 0$, the isolated term $\check{f}(0)$ also vanishes. The low-frequency, macroscopic modes at the lower boundary determine the remaining term through the derivative terms:

$$R_{\text{thermal}} = +_{\check{m}=1}^{\check{n}-1} \widetilde{m!} B_m \left(0 - {}^{\acute{m}}f(0)\right).$$

This boundary term is to be understood as the regular, macroscopic remainder of the deflated mode structure. In a complete physical derivation it must be combined with the usual mode analysis of Hawking radiation, in particular with the thermal occupation number arising from the mixing of positive and negative frequencies.

Conclusion of the derivation

The trans-Planckian problem arises when the semiclassical calculation is traced back into frequency ranges in which the underlying continuum description itself becomes questionable.

The nonstandard analytical perspective offers a precise way of reorganising this situation: the extreme frequencies at the horizon are collected in a dominant hyperreal integral P_{vacuum}. This is not interpreted as a physically directly radiating component, but as a background structure of the model.

The actual Hawking radiation is then not obtained from undefined trans-Planckian infinities, but from a finite, regular remainder term that remains after the algebraic separation of the hyperreal pole. Hawking's result is therefore not replaced, but reorganised in a nonstandard analytical language: the thermal radiation remains the macroscopic result, while the trans-Planckian contribution appears as an exactly specified background pole.

String theory

The critical dimension of string theory and the regularised zero-point energy

String theory describes fundamental particles not as point-like objects, but as one-dimensional vibrating strings. In the simple bosonic string theory, the requirement of quantum-mechanical consistency, in particular the vanishing of the conformal anomaly, leads to the critical dimension

$$D = 26.$$

In superstring theory, the corresponding result is

$$D = 10.$$

A central component of these calculations is the regularised zero-point energy of the infinitely many string modes. Formally, this involves the divergent sum

$$1 + 2 + 3 + \dots$$

which, in the usual treatment, is assigned the value

$$\zeta(-1) = \widetilde{12\text{-}}$$

by analytic continuation, zeta regularisation, or related procedures. From a nonstandard analytical point of view, however, this value is not the sum of all modes itself, but the regular remainder after a dominant background contribution has been separated off.

The zero-point energy of strings

A string has infinitely many oscillation modes. Each mode formally behaves like a quantum-mechanical harmonic oscillator. The zero-point energy therefore contains a summation kernel of the form

$$E \propto +_{n=1}^{\omega} n.$$

In the ordinary sense this sum diverges. In the standard formulation it is therefore not evaluated as an ordinary sum, but regularised. In abbreviated form one often writes

$$1 + 2 + 3 + 4 + \cdots = \widetilde{12\text{-}}.$$

This notation is useful, but potentially misleading: it does not say that the ordinary or hyperfinite sum is actually equal to $\widetilde{12\text{-}}$. It denotes only the regularised remainder value.

In bosonic string theory, this remainder occurs in the normal-ordering contribution, or intercept, of the string modes. Schematically, the dimensional condition can be written as

$$(\check{D} - 1) \cdot \widetilde{12\text{-}} + 1 = 0$$

from which

$$D = 26$$

follows. This short form, however, summarises only part of the full consistency condition. The usual derivation also involves the Virasoro algebra, Lorentz invariance, and the cancellation of the conformal anomaly.

Hyperreal deflation of the oscillation modes

In discretised nonstandard space, physical frequencies do not run into an indefinite infinity. They extend only up to a fundamental midfinite constant ω. With a physical damping factor $g(n)$ at the Planck scale,

$$g(0) = 1, \qquad g(\omega) = 0,$$

the hyperfinite mode sum is

$$E \propto +_{n=1}^{\omega} n \cdot g(n).$$

Applying the hyperreal Euler-Maclaurin formula to

$$f(x) = x \cdot g(x)$$

splits the sum into a dominant integral term and boundary contributions.

If g is chosen so that the relevant derivatives vanish at the upper boundary,

$$f(\omega) = 0, \qquad {}^{1}f(\omega) = 0, \qquad {}^{2}f(\omega) = 0, \qquad \text{and so on,}$$

then the upper boundary terms make no regular contribution. At the lower boundary, for low frequencies one has approximately $g(x) = 1$, hence

$${}^{1}f(0) = 1.$$

The first regular remainder term then comes from the second-order Bernoulli term with

$$B_2 = \widetilde{6}.$$

Thus

$$R_{\text{string}} = \widetilde{2!}B_2\left(0 - {}^{1}f(0)\right) = \widetilde{12\text{-}}.$$

The hyperfinite zero-point structure is therefore not

$$E \propto \widetilde{12\text{-}},$$

but rather

$$E \propto \uparrow_0^{\omega} x \cdot g(x) \downarrow x - \widetilde{12}.$$

The value $\widetilde{12\text{-}}$ is therefore the regular remainder of the deflated summation structure, not the entire hyperfinite energy.

The critical point

The nonstandard analytical view shows where the value

$$\widetilde{12\text{-}}$$

comes from structurally: it is the finite remainder term at the lower boundary of the mode spectrum after the dominant hyperreal integral contribution has been separated off.

The problem is therefore not regularisation itself, but an overly literal reading of the short form

$$1 + 2 + 3 + \cdots = \widetilde{12\text{-}}.$$

From the nonstandard analytical point of view, it also has to be clarified what happens to the hyperreal background contribution

$$\uparrow_0^{\omega} x \cdot g(x) \downarrow x.$$

In the Casimir effect, there is a difference formation between two geometries, by which common background contributions are balanced against one another. In the isolated treatment of a string, a corresponding cancellation or calibration mechanism must be stated explicitly.

The critical dimension

$$D = 26$$

is therefore not simply an arbitrary consequence of a forbidden summation. It belongs to a broader consistency condition of quantised string theory. The nonstandard analytical objection is more precise: the regularised value $\widetilde{12\text{-}}$ may enter the dimensional condition by itself only if the hyperreal background contribution is removed or compensated from the measurable balance by symmetry, a gauge condition, normal ordering, reference subtraction, or some other exactly specified mechanism.

Conclusion of the derivation

The equation

$$1 + 2 + 3 + \cdots = \widetilde{12}\text{-}$$

is not a statement about the ordinary sum of all natural numbers, nor about the full hyperfinite mode sum. It describes the regularised remainder value

$$\zeta(-1) = \widetilde{12}\text{-}.$$

Nonstandard analysis makes this distinction explicit: the hyperfinite sum of the string modes deflates into a dominant background contribution and a regular remainder. The value $\widetilde{12}$- belongs to the remainder, not to the entire summation kernel.

The critical dimension of string theory should therefore not be read as an isolated identification of the full zero-point energy with this remainder value. From the nonstandard analytical point of view, one must explicitly state by which structural mechanism the hyperreal background contribution disappears from the physically measurable balance.

Boris Haase

Curriculum Vitae

1964 born in Kiel as the first son of LLD Udo Haase and his wife Angela,

1966 brother Nicolai born,

1969 moved to Detmold followed by elementary and grammar school,

1983 participated in the state competition "Jugend forscht" in Lower Saxony,

1984 A level in Göttingen, then basic military service with the 5th/12th in Osterode am Harz,

1985 studies in mathematics started with minor subject computer science (initially physics),

1990 trained as a management assistant in data processing,

1993 employment started as IT specialist at the county hospital Hameln,

1997 moved to the Hanau Clinic in the same position (until today),

1998 studies in philosophy started with minor subjects political science and sociology (until 2004),

2000 work on the homepages started and mother died (2006),

2011 book "Relil - Religion und Lebensweg" published,

2014 first edition published of the book "Nonstandard Mathematics",

2020 studies in informatics started at the Fernakademie Hamburg with ending 2021.

Address: Cranachstr. 4, D-63452 Hanau, mail@boris-haase.de

Appendix: FastShift Algorithm for Matrix Multiplication in C++

Modelling the nonlinear runtime $T_n^b = O((\ell^2))n^a$ checked by Mathematica

The following C++ programme yields especially the runtime $T_c^b = O(1)n^3$ resp. $T_f^b = O(\ell^2)n^2$ for every $n \times n$ matrix with *loop number* $a \in \{2, 3\}$, *bit length* b and integer (!) *depth of recursion* $\ell = {}_2n$.

```cpp
// FastShift algorithm as fastintopt.cpp by Boris Haase (01.01.2026)
// g++ -O3 -std=gnu++17 -march=native -fopenmp fastintopt.cpp -o fastintopt -lgmpxx -lgmp
#include <algorithm>
#include <chrono>
#include <iostream>
#include <vector>
#include <gmpxx.h>
#include <omp.h>
using namespace std;
static inline void mpz_addmul_cpp(mpz_class &rop, const mpz_class &op1, const mpz_class &op2) {
    mpz_addmul(rop.get_mpz_t(), op1.get_mpz_t(), op2.get_mpz_t());
}
static inline void fdiv_qr(mpz_class &q, mpz_class &r, const mpz_class &a, const mpz_class &b) {
    mpz_fdiv_qr(q.get_mpz_t(), r.get_mpz_t(), a.get_mpz_t(), b.get_mpz_t());
}
struct FlatM {
    mpz_class* ptr = nullptr;
    std::size_t row_stride = 0;
    std::vector<mpz_class> data;
    FlatM() = default;
    explicit FlatM(int n) : row_stride((std::size_t)n), data((std::size_t)n * n) { ptr = data.data(); }
    FlatM(mpz_class* raw_ptr, std::size_t stride) : ptr(raw_ptr), row_stride(stride) {}
    inline mpz_class& operator()(int i, int j) { return *(ptr + i * row_stride + j); }
    inline const mpz_class& operator()(int i, int j) const { return *(ptr + i * row_stride + j); }
    inline FlatM submatrix(int ro, int co) const { return FlatM(ptr + ro * row_stride + co, row_stride); }
};
static FlatM multiply_blocked(const FlatM& A, const FlatM& B, int n) {
    FlatM C(n);
    const int bs = 64;
    if (!omp_in_parallel()) {
        #pragma omp parallel for collapse(2) schedule(static)
        for (int i0 = 0; i0 < n; i0 += bs)
            for (int j0 = 0; j0 < n; j0 += bs) {
                int ie = min(i0 + bs, n), je = min(j0 + bs, n);
                for (int k0 = 0; k0 < n; k0 += bs) {
                    int ke = min(k0 + bs, n);
                    for (int i = i0; i < ie; ++i)
                        for (int k = k0; k < ke; ++k) {
                            const mpz_class& av = A(i, k);
                            for (int j = j0; j < je; ++j) mpz_addmul_cpp(C(i, j), av, B(k, j));
                        }
                }
            }
        return C;
    }
    for (int i0 = 0; i0 < n; i0 += bs)
        for (int j0 = 0; j0 < n; j0 += bs) {
            int ie = min(i0 + bs, n), je = min(j0 + bs, n);
            for (int k0 = 0; k0 < n; k0 += bs) {
                int ke = min(k0 + bs, n);
                for (int i = i0; i < ie; ++i)
                    for (int k = k0; k < ke; ++k) {
                        const mpz_class& av = A(i, k);
                        for (int j = j0; j < je; ++j) mpz_addmul_cpp(C(i, j), av, B(k, j));
                    }
            }
        }
    return C;
}
static FlatM add_shiR(const FlatM& C, const FlatM& D, int n, int r_bits) {
    FlatM R(n);
    if (!omp_in_parallel() && n >= 256) {
        #pragma omp parallel for collapse(2) schedule(static)
        for (int i = 0; i < n; ++i)
            for (int j = 0; j < n; ++j) {
                mpz_mul_2exp(R(i, j).get_mpz_t(), C(i, j).get_mpz_t(), r_bits);
                mpz_add(R(i, j).get_mpz_t(), R(i, j).get_mpz_t(), D(i, j).get_mpz_t());
            }
        return R;
    }
    for (int i = 0; i < n; ++i)
        for (int j = 0; j < n; ++j) {
            mpz_mul_2exp(R(i, j).get_mpz_t(), C(i, j).get_mpz_t(), r_bits);
            mpz_add(R(i, j).get_mpz_t(), R(i, j).get_mpz_t(), D(i, j).get_mpz_t());
        }
    return R;
}
static FlatM div_modR(const FlatM& S11, const FlatM& S12, const FlatM& S21, const FlatM& S22, int n, int s_bits) {
    int m = n / 2; FlatM R(n); mpz_class B = mpz_class(1) << s_bits, H = B >> 1;
    if (!omp_in_parallel() && m >= 256) {
        #pragma omp parallel for collapse(2) schedule(static)
        for (int i = 0; i < m; ++i)
            for (int j = 0; j < m; ++j) {
                mpz_class q, r, St = S11(i, j) + S12(i, j);
                fdiv_qr(q, r, St, B);
                if (r > H) { ++q; r -= B; } else if (r <= -H) { --q; r += B; }
                R(i, j) = q; R(i, j + m) = r;
                mpz_class Sb = S21(i, j) + S22(i, j);
                fdiv_qr(q, r, Sb, B);
                if (r > H) { ++q; r -= B; } else if (r <= -H) { --q; r += B; }
                R(i + m, j) = q; R(i + m, j + m) = r;
            }
```

```cpp
        return R;
    }
    for (int i = 0; i < m; ++i)
        for (int j = 0; j < m; ++j) {
            mpz_class q, r, St = S11(i, j) + S12(i, j);
            fdiv_qr(q, r, St, B);
            if (r > H) { ++q; r -= B; } else if (r <= -H) { --q; r += B; }
            R(i, j) = q; R(i, j + m) = r;
            mpz_class Sb = S21(i, j) + S22(i, j);
            fdiv_qr(q, r, Sb, B);
            if (r > H) { ++q; r -= B; } else if (r <= -H) { --q; r += B; }
            R(i + m, j) = q; R(i + m, j + m) = r;
        }
    return R;
}
static FlatM fastshift_rec(const FlatM& Y, const FlatM& Z, int n, int s_bits, int leaf_cutoff, int depth) {
    if (n <= leaf_cutoff) return multiply_blocked(Y, Z, n); else depth++;
    int m = n / 2, t_bits = s_bits * 2;
    FlatM A = add_shiR(Z.submatrix(0, 0), Z.submatrix(0, m), m, s_bits);
    FlatM B = add_shiR(Z.submatrix(m, 0), Z.submatrix(m, m), m, s_bits);
    FlatM U, V, W, X;
    if (depth < 3) {
        #pragma omp taskgroup
        {
            #pragma omp task shared(U)
            U = fastshift_rec(Y.submatrix(0, 0), A, m, t_bits, leaf_cutoff, depth);
            #pragma omp task shared(V)
            V = fastshift_rec(Y.submatrix(0, m), B, m, t_bits, leaf_cutoff, depth);
            #pragma omp task shared(W)
            W = fastshift_rec(Y.submatrix(m, 0), A, m, t_bits, leaf_cutoff, depth);
            #pragma omp task shared(X)
            X = fastshift_rec(Y.submatrix(m, m), B, m, t_bits, leaf_cutoff, depth);
        }
    } else {
        U = fastshift_rec(Y.submatrix(0, 0), A, m, t_bits, leaf_cutoff, depth);
        V = fastshift_rec(Y.submatrix(0, m), B, m, t_bits, leaf_cutoff, depth);
        W = fastshift_rec(Y.submatrix(m, 0), A, m, t_bits, leaf_cutoff, depth);
        X = fastshift_rec(Y.submatrix(m, m), B, m, t_bits, leaf_cutoff, depth);
    }
    return div_modR(U, V, W, X, n, s_bits);
}
int main(int argc, char** argv) {
    int threads = 8, bits  = 256, m_min = 8, m_max = 13, leaf_cutoff = 2048;
    if (argc > 1) threads  = max(1, atoi(argv[1]));
    if (argc > 2) bits     = max(1, atoi(argv[2]));
    if (argc > 3) m_min    = max(1, atoi(argv[3]));
    if (argc > 4) m_max    = max(m_min, atoi(argv[4]));
    if (argc > 5) leaf_cutoff = max(16, atoi(argv[5]));
    omp_set_num_threads(threads);
    for (int m = m_min; m <= m_max; ++m) {
        int n = 1 << m, s_bits = bits * 2 + m;
        cout << "FastShift " << bits << " Bits | n=2^" << m << " | leaf_cutoff=" << leaf_cutoff << endl;
        gmp_randclass rng(gmp_randinit_default); rng.seed(42u + m);
        FlatM A(n), B(n);
        for (int i = 0; i < n; ++i) for (int j = 0; j < n; ++j) { A(i, j) = rng.get_z_bits(bits);
            B(i, j) = rng.get_z_bits(bits); }
        auto t1 = chrono::steady_clock::now();
        FlatM Cf = fastshift_rec(A, B, n, s_bits, leaf_cutoff, 0);
        auto t2 = chrono::steady_clock::now();
        FlatM Cs = multiply_blocked(A, B, n);
        auto t3 = chrono::steady_clock::now();
        auto ms_fs = chrono::duration_cast<chrono::milliseconds>(t2 - t1).count();
        auto ms_std = chrono::duration_cast<chrono::milliseconds>(t3 - t2).count();
        cout << "n = " << n << " | FS: " << ms_fs << " ms | Std: " << ms_std << " ms | "
             << (Cf(0,0)==Cs(0,0) ? "OK" : "Error") << endl;
    }
    return 0;
}
```

Listing 1: Integer FastShift Algorithm in C++

Table 2: Mean execution times on an AMD EPYC™ 9645 (24 cores, 128 GB DDR5)
Comparison of T_c (Standard) and T_f (FastShift) in ms for $k = 3$ measurements:

ℓ	$\bar{T}_c^{64}$	$\bar{T}_f^{64}$	**Ratio**	$\bar{T}_c^{128}$	$\bar{T}_f^{128}$	**Ratio**
10	938	881	1.06	1096	1151	0.95
11	6157	6170	1.00	8295	8440	0.98
12	41 525	27 483	**1.51**	62 070	42 443	**1.46**
13	322 066	131 767	**2.44**	484 108	257 820	**1.88**

Note: At 256 bits, the advantages almost level out on this architecture.

List of Symbols

Symbol	Usage	Interpretation
$\sim$	$\tilde{a}$	Reciprocal of a: $1/a$ resp. a^{-1} for $a \neq 0$ (read as "turn")
$\acute{}$	$\acute{a}$	Decrement of a: $a - 1$ (read as "dec")
$\grave{}$	$\grave{a}$	Increment of a: $a + 1$ (read as "inc")
$\hat{}$	$\hat{a}$	Double of a: $2a$ (read as "hat")
$\check{}$	$\check{a}$	Half of a: $a/2$ (read as "half")
-	a-	a negated: a- (read as "neg")
$\_$	$z = a + \underline{b}$	Complex part of z: $\underline{1}b$ with imaginary unit $\underline{1}$ (read as "im")
ν	${}^{\nu}A$	Greatest finite number: intersection of the complex or real set A with ${}^{\nu}\mathbb{C} := [-\nu, \nu] + \underline{1}[-\nu, \nu]$
ω	${}^{\omega}A$	Greatest mid-finite number: intersection of the complex or real set A with ${}^{\omega}\mathbb{C} := [-\omega, \omega] + \underline{1}[-\omega, \omega]$
ι	$\iota = \min \mathbb{R}_{>0}$	Smallest positive real number
^{n}	${}^{n}a = a^{(n)}$	n-th derivative of a (read as "n of a")
${}_{b}$	${}_{b}a = \log_b a$	Logarithm to base b for $a \in \mathbb{C} \setminus \mathbb{R}_{\leq 0}$ (read as "b log a")
${}_{1}$	${}_{1}x = x/\|\|x\|\|$	Unit vector for $x \neq 0$
∞	$\infty \gg \tilde{\iota}^2$	Replacing ± 0 by $\pm\widetilde{\infty}$ as well as $A_\infty := A \cup \{\pm\infty\}$ for the set $A \subseteq \mathbb{R}$
$\mathbb{M}$	$\mathbb{M}_{\mathbb{R}} = {}^{\omega}\mathbb{R} \setminus {}^{\nu}\mathbb{R}$	Mid-finite numbers: $\mathbb{M}_{\mathbb{C}} := \mathbb{M}_{\mathbb{R}} + \underline{\mathbb{M}}_{\mathbb{R}}$
$\dot{}$	$\dot{A}$	Point-symmetric set A (read as "point")
$\complement$	$\complement_{(m=)1}^{n;s} a_m$	Concatenation (read as "con") of the a_m to $a_1, \ldots, a_n$ with step width s (read as "step") - analogies exist for $+$, $\pm$, $\mp$ and $\times$ instead of $\complement$
$\leftharpoonup$	$\overleftharpoon{a}$	Predecessor of a (read as "pre")
$\rightharpoonup$	$\overrightharpoon{a}$	Successor of a (read as "post")
$\upharpoonleft$	$a\upharpoonleft_n$	n-fold repetition of a in the form $(a, \ldots, a)^\top$ (read as "rep")
$\upharpoonright$	$a\upharpoonright_{f(,w)}$	f-mean of $a_1, \ldots, a_n$ for $a \in \mathbb{R}^n_{\geq 0}$ with function f (maybe with weight vector $w \in [0, 1]^n$) or Hölder mean for level $r \in {}^{\nu}\mathbb{R}^*$ instead of f (read as "mean")
$\downarrow$	$\downarrow x$	Differential of x (read as "down")
$\uparrow$	$\uparrow f(x) \downarrow x$	Integral of $f(x)$ (read as "up")
$\square$		End of proof
$\triangle$		End of definition

Bibliography

[1] D. Borwein, J. M. Borwein, P. B. Borwein und R. Girgensohn: *Giuga's Conjecture on Primality*; Amer. Math. Monthly 103:40-50; 1996.

[2] Croft, Hallard T.; Falconer, Kenneth J.; Guy, Richard K.: *Unsolved Problems in Geometry*; Reprint of 1st Ed.; 2013; Springer; New York.

[3] Deiser, Oliver: *Einführung in die Mengenlehre*; 1. Aufl.; 2002; Springer; Berlin.

[4] Gelbaum, Bernard R.; Olmsted, John M. H.: *Counterexamples in Analysis*; Republ., unabr., slightly corr.; 2003; Dover Publications; Mineola, New York.

[5] Golub, Gene H.; van Loan, Charles F.: *Matrix Computations*; 3rd Ed.; 1996; Johns Hopkins University Press; Baltimore.

[6] Guy, Richard K.: *Unsolved Problems in Number Theory*; 3rd Ed.; 2004; Springer; New York.

[7] Hämmerlin, Günther; Hoffmann, Karl-Heinz: *Numerische Mathematik*; 2. Aufl.; 1991; Springer; Berlin.

[8] Hermann, Martin: *Numerische Mathematik*; 3., überarb. u. erw. Aufl.; 2011; Oldenbourg; München.

[9] Heuser, Harro: *Lehrbuch der Analysis Teil 1*; 17., akt. Aufl.; 2009; Vieweg + Teubner; Wiesbaden.

[10] Heuser, Harro: *Lehrbuch der Analysis Teil 2*; 14., überarb. Aufl.; 2008; Vieweg + Teubner; Wiesbaden.

[11] Hilbert, David: *Grundlagen der Geometrie*; 2. Aufl.; 1903; Teubner; Leipzig.

[12] Hoffmann, Dirk W.: *Theoretische Informatik*; 4., akt. Aufl.; 2018; Carl Hanser; München.

[13] Ivic, Aleksandar: *The Riemann Zeta-Function*; Reprint; 2003; Dover Publications; Mineola.

[14] Kac, Mark: *On the Average Number of Real Roots of a Random Algebraic Equation*; Bull. Amer. Math. Soc. 49 (4); 1943; 314 - 320.

[15] Knabner, Peter; Barth, Wolf: *Lineare Algebra*; 2., überarb. und erw. Aufl.; 2018; Springer; Berlin.

[16] Knuth, Donald Ervin: *The Art of Computer Programming Volume 2*; 3rd Ed.; 1997; Addison Wesley; Reading.

[17] Köhler, Günter: *Analysis*; 1. Aufl.; 2006; Heldermann; Lemgo.

[18] Kowalsky, Hans-Joachim: *Topologische Räume*; 1. Aufl.; 1961; Birkhäuser; Basel.

[19] Landers, Dieter; Rogge, Lothar: *Nichtstandard Analysis*; 1. Aufl.; 1994; Springer; Berlin.

[20] Mollin, Richard A.: *Advanced Number Theory with Applications*; 1st Ed.; 2017; Routledge; Milton Park.

[21] Querenburg, Boto von: *Mengentheoretische Topologie*; 3., neu bearb. u. erw. Aufl.; 2001; Springer; Berlin.

[22] Remmert, Reinhold: *Funktionentheorie 1*; 3., verb. Aufl.; 1992; Springer; Berlin.

[23] Remmert, Reinhold: *Funktionentheorie 2*; 1. unveränd. Nachdruck der 1. Aufl.; 1992; Springer; Berlin.

[24] Robbins, Herbert, *A Remark on Stirling's Formula*, The American Mathematical Monthly, 62 (1); 1955; 26 - 29.

[25] Scheid, Harald: *Zahlentheorie*; 1. Aufl.; 1991; Bibliographisches Institut; Mannheim.

[26] Schwarz, Hans Rudolf; Köckler, Norbert: *Numerische Mathematik*; 7., überarb. Aufl.; 2009; Vieweg + Teubner; Wiesbaden.

[27] Slapničar, Ivan: *There are no cycles in the $3n + 1$ sequence*; arXiv: 1706.08399v1.

[28] Vanderbei, Robert J.: *Linear Programming*; 3rd Ed.; 2008; Springer; New York.

[29] Walter, Wolfgang: *Analysis 1*; 3., verb. Aufl.; 1992; Springer; Berlin.

[30] Walter, Wolfgang: *Analysis 2*; 5., erw. Aufl.; 2002; Springer; Berlin.

[31] Walter, Wolfgang: *Gewöhnliche Differentialgleichungen*; 7. Aufl.; 2000; Springer; Berlin

[32] Whittaker, Edmund T.; Watson, George N.: *A Course of Modern Analysis*; 5th Ed.; 2021; Cambridge University Press; Cambridge.

Notes: