BEYOND POPULAR SCIENCE

DAVID H. SILVER

BEYOND POPULAR SCIENCE

David H. Silver

https://www.openbookpublishers.com

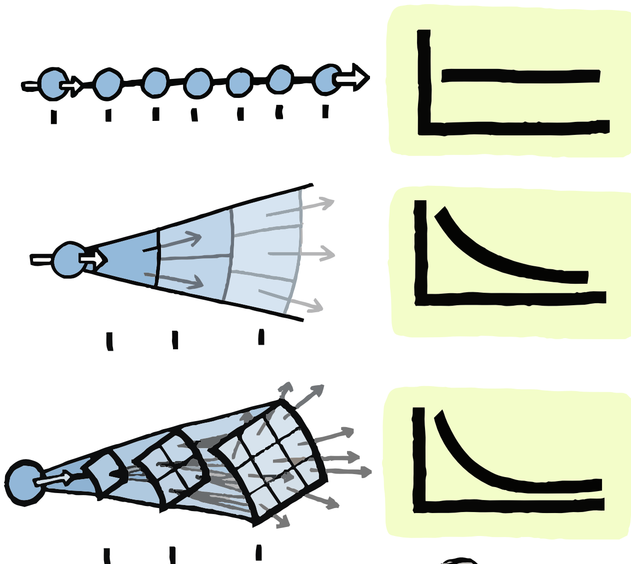
Digital material and resources associated with this volume are available at https://doi.org/10.11647/OBP.0526#resources

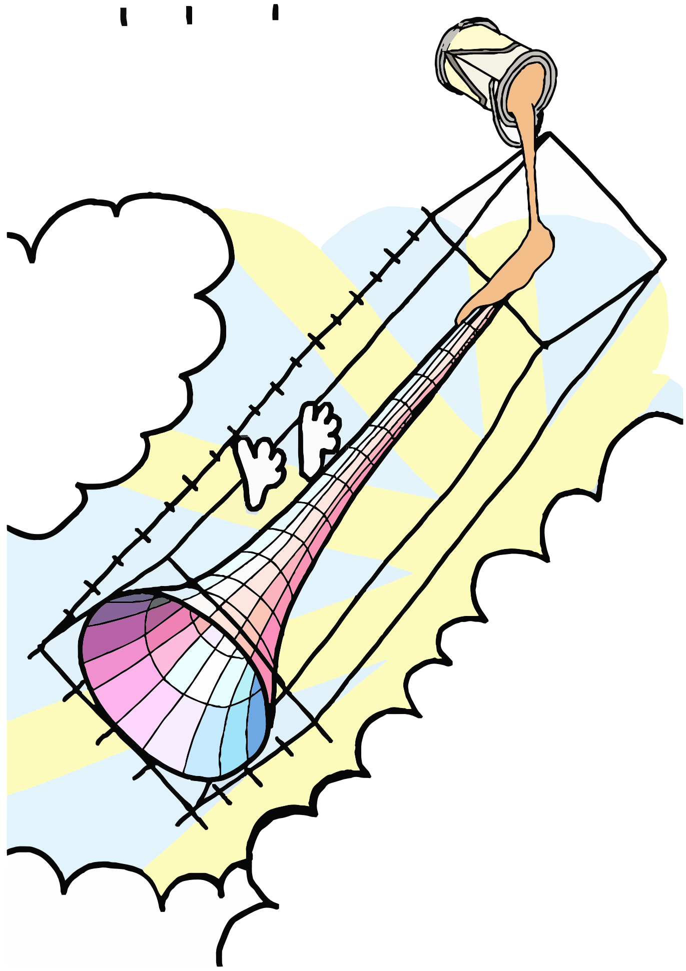
ISBN Paperback:	978-1-80511-877-0
ISBN Hardback:	978-1-80511-878-7
ISBN Digital (PDF):	978-1-80511-879-4
ISBN HTML:	978-1-80511-881-7
ISBN Digital ebook (epub):	978-1-80511-880-0
DOI:	10.11647/OBP.0526

Cover image by Enny Silver and David H. Silver

Cover design by Jeevanjot Kaur Nagpal

Put on Your 4D Glasses

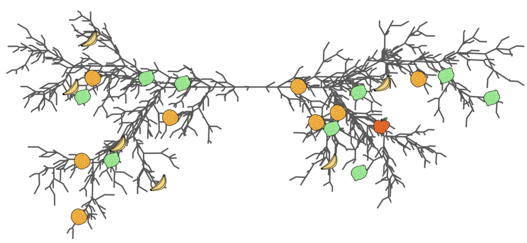




Top (Dimensional Scaling of Inverse Laws): The same point-source strength spreads differently depending on spatial dimension. In 1D, the influence remains constant along a line. In 2D, the effect dilutes like $1/r$ as it spreads over a circle. In 3D, it falls off as $1/r^2$, spreading over the surface of a sphere. This explains why gravitational and electrostatic forces scale as inverse-square laws in 3D space.
Bottom (Gabriel's Horn and the Painter's Paradox): A surface of revolution formed by rotating $y = 1/x$ around the x-axis for $x > 1$. Though the horn extends infinitely, it encloses a finite volume ($\int_1^\infty \pi(1/x)^2\,dx = \pi$) but has infinite surface area ($2\pi \int_1^\infty (1/x)\,\sqrt{1 + 1/x^4}\,dx = \infty$). Paradoxically, one could 'fill' it with a finite amount of paint, but never coat its inner surface.

Put on Your 4D Glasses

Why does our universe have exactly three spatial dimensions plus time? Multiple independent constraints converge on $D = 4$: only in 3D space do inverse-square laws produce stable planetary orbits and bound atoms; only in 4D spacetime are fundamental forces renormalizable in quantum field theory; only in 4D do waves propagate cleanly without trailing echoes (Huygens's principle). The arithmetic fact that $4 = 2+2$ creates unique mathematical properties—from quaternion algebra to self-dual gauge fields—that cascade through physics. Lower dimensions cannot support complex chemistry, while higher dimensions destabilise matter and causality.



Four-Dimensional Spacetime ∘ Inverse-Square Force Law ∘ Stable Orbits Only $n = 3$ ∘ Huygens's Principle ∘ QFT Renormalizability ∘ Exotic $\mathbb{R}^4$ Topology ∘ Division Algebras ∘ $4 = 2 + 2$ Identity ∘ Black Hole No-Hair ∘ Atomic Stability ∘ Chemical Bonding

"The miracle of the appropriateness of the language of mathematics
for the formulation of the laws of physics
is a wonderful gift which we neither understand nor deserve."

— Eugene Wigner, 1960

Put on Your 4D Glasses

The idea that the dimensionality of physical space might be constrained by necessity predates the formal development of modern physics. Gottfried Wilhelm Leibniz suggested in the *Discourse on Metaphysics* (1686) that the actual world should be understood as the one ‘simplest in hypotheses and richest in phenomena,’ implicitly framing dimensionality as subject to selection principles. In the eighteenth century, Immanuel Kant proposed that Newton's inverse-square laws implied the three-dimensionality of space, although this causal inference would later be reversed—the inverse-square law follows from spatial geometry via Gauss's theorem, not the reverse.

A rigorous analytical approach began with Paul Ehrenfest's seminal 1917 paper ‘In what way does it become manifest in the fundamental laws of physics that space has three dimensions?’ He demonstrated that classical orbit stability requires exactly three spatial dimensions. In higher spatial dimensions ($N > 3$), the effective gravitational potential falls off too rapidly to maintain closed, bounded orbits; in lower dimensions, the dynamics become pathologically confined. Separately, the Huygens principle for the wave equation holds only in odd spatial dimensions $n \geq 3$, so $(3,1)$ spacetime is the lowest-dimensional case with sharp wavefronts and no interior tails.

Gerald Whitrow's influential 1955 paper “Why Physical Space has Three Dimensions” marked a turning point by explicitly connecting dimensional constraints to the possibility of life and observers. He argued that intelligent life capable of formulating physics could only arise in three spatial dimensions—an example of anthropic reasoning. Whitrow systematically examined how communication, neural networks, and information processing would fail in spaces of different dimensionality, establishing that the question ‘why three dimensions?’ might be answered by ‘because otherwise we wouldn't be here to ask.’

Freeman Dyson and Andrew Lenard's 1967 theorems on the stability of matter established that systems of electrons and nuclei interacting via Coulomb forces are stable of the second kind in three spatial dimensions. In higher spatial dimensions, Coulomb interactions scale differently and can lead to instabilities; in two or one spatial dimensions the Coulomb law changes and the stability analysis requires separate arguments.

Four-dimensional spacetime with a metric of signature $(3,1)$ satisfies multiple independent physical constraints that fail in other dimensionalities. The signature means the metric has three positive and one negative eigenvalue, but this does not single out a canonical ‘time direction’—timelike vectors form an open cone, and one can choose bases in which all four basis vectors are lightlike. The physical content lies in the signature, not in any preferred decomposition. Gravitational orbits destabilise, many familiar interactions become non-renormalizable for $D > 4$ (and super-renormalizable for $D < 4$), wave propagation develops trailing echoes, and atomic structures can fail to be stable when $D \neq 4$. Multiple time-like dimensions destroy the well-posedness of the initial value problem for hyperbolic differential equations such as the wave equation, rendering physics unpredictable, and generically spoil causality and energy positivity.

 https://doi.org/10.11647/OBP.0526.24

In classical potential theory, the spatial decay of fields from a point source—or any spherically symmetric mass, regardless of its internal dimensionality—follows a general scaling law determined by Gauss's theorem: the flux through a sphere in n spatial dimensions scales with its surface area, yielding a radial dependence of $1/r^{n-1}$ for the field and $1/r^{n-2}$ for the associated potential. In $n = 3$ this produces the inverse-square law that governs Newtonian gravity (Newton, 1687) and electrostatics. This particular falloff enables stable bound orbits under central forces, since it balances centripetal acceleration with potential curvature. In $n > 3$, the force falls too quickly to support closed, radially stable Kepler orbits; in $n < 3$, the dynamics cease to admit Kepler-type closed, radially stable motion.

Wave propagation obeys Huygens's principle (Huygens, 1690) only in odd spatial dimensions $n \geq 3$. In $(3, 1)$ spacetime, a localised disturbance generates a sharp spherical wavefront without trailing components. In even spatial dimensions, persistent field residuals remain after the main wave passes, blurring temporal boundaries between cause and effect as the Green's function of the wave equation has tails inside the light cone.

Quantum field theory imposes stringent dimensional restrictions on interaction consistency. Renormalizability—the ability to absorb divergences into a finite set of physical parameters—depends on the dimensional scaling of coupling constants. In $D = 4$, key interactions such as ϕ^4 theory, quantum electrodynamics (QED), and non-abelian gauge theories feature dimensionless couplings, rendering loop corrections manageable via renormalization group techniques. In $D > 4$, the same interactions become non-renormalizable, requiring an infinite tower of counterterms. In $D < 4$, they become super-renormalizable.

The manifold $\mathbb{R}^4$ exhibits an anomaly in differential topology: it admits uncountably many smooth structures that are pairwise non-diffeomorphic yet topologically equivalent. For $\mathbb{R}^n$ with $n \neq 4$, the smooth structure is unique—smoothness and topology coincide. (For general manifolds, distinct smooth structures on the same topological manifold, and topological manifolds admitting no smooth structure at all, were already known before exotic $\mathbb{R}^4$.) The exotic $\mathbb{R}^4$ phenomenon is intertwined with deep four-dimensional phenomena revealed by gauge theory, notably Donaldson invariants and Seiberg–Witten theory; the smooth 4D Poincaré conjecture itself remains open.

In the algebraic classification of normed division algebras over $\mathbb{R}$, there exist only four: $\mathbb{R}$ (dimension 1), $\mathbb{C}$ (dimension 2), $\mathbb{H}$ (dimension 4), and $\mathbb{O}$ (dimension 8). Of these, the quaternions $\mathbb{H}$ preserve associativity while the octonions $\mathbb{O}$ do not. They form the algebraic underpinning of spinor representations and enable the group isomorphism $\mathrm{SU}(2) \cong \mathrm{Spin}(3)$, which double-covers the rotation group $\mathrm{SO}(3)$. This supports the representation theory of spin-$\frac{1}{2}$ particles and the construction of Dirac spinors. No higher-dimensional associative division algebra exists, and the non-associativity of octonions complicates their use in comparable representation frameworks, though they underlie exceptional Lie groups (e.g., G_2, F_4, E_6–E_8).

The necessity of spinors in four dimensions emerges from a tension between quantum mechanics and relativity. Schrödinger's equation is linear in time derivatives, while relativistic energy obeys $E^2 = p^2c^2 + m^2c^4$, quadratic in energy. Dirac sought to linearize the wave operator (Dirac, 1928)—to extract a 'square root' of the d'Alembertian $\Box = \partial_t^2 - c^2\nabla^2$. Just as $x^2 + y^2$ cannot be factored into $(ax + by)$ using real numbers, this operator resists scalar factorization. The solution requires anticommuting coefficients, matrices γ^μ satisfying $\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}$. In four-dimensional spacetime, the minimal representation uses 4×4 matrices, forcing the wavefunction to be a four-component spinor rather than a scalar. The four components do not represent spatial directions, but two particle states and two antiparticle states, each with two spin orientations. Antimatter is a result of this mathematical necessity from the requirement to linearize energy in $D = 4$ spacetime. The restriction to three spatial dimensions is important: the rotation group SO(3) is unique in admitting a double cover Spin(3) $\cong$ SU(2) that links vector and spinor representations through this square-root relationship.

In general relativity, the uniqueness of black hole solutions—encapsulated by the no-hair theorems—holds in four-dimensional, asymptotically flat spacetime under suitable regularity and symmetry assumptions. Theorems by Israel, Carter, and Robinson prove that stationary black holes in $D = 4$ are characterised entirely by mass, charge, and angular momentum. In higher dimensions, this rigidity fails. New solutions emerge with toroidal or ring-like horizons, including black rings and black strings, and the solution space displays richer phases.

The quantum mechanical stability of atomic matter depends sensitively on the spatial dimension n. For hydrogen-like atoms with a $1/r^{n-2}$ potential when $n \geq 3$, the familiar Coulombic spectrum arises for $n = 3$. In $n > 3$, the potential decays too rapidly to maintain binding; in $n = 2$, the potential becomes logarithmic. Chemical bonding patterns also require three dimensions. Tetrahedral carbon and chiral centres depend on SO(3) symmetry. In $n = 2$, bonding is planar and chirality is lost; in $n > 3$, additional rotational degrees of freedom would alter biochemical recognition.

I recommend watching Mikhail Gromov's lecture on the topic (minute 19:30 in the video titled "What is a Manifold? - Mikhail Gromov") where Gromov traces the exceptional nature of four dimensions to the arithmetic identity $4 = 2+2$. A four-element set partitions into two pairs in exactly three ways: $\{\{1,2\},\{3,4\}\}$, $\{\{1,3\},\{2,4\}\}$, and $\{\{1,4\},\{2,3\}\}$. The number of such partitions for a set of size $2n$ is $(2n-1)!! = (2n-1)(2n-3)\cdots 3 \cdot 1$. For $n = 2$: three partitions. For $n = 3$: fifteen partitions. For $n = 4$: one hundred and five partitions. Only when $n = 2$ does the partition count (3) remain smaller than the set size (4), enabling the symmetric group S_4 to map onto the smaller group S_3.

For odd-sized sets, no symmetric pair partitions exist. Only the four-element set achieves both symmetry (all parts equal) and economy (partition count smaller than set size).

This homomorphism $\varphi : S_4 \to S_3$ tracks how permutations of four elements permute the three partitions. Its kernel contains precisely those permutations preserving all partitions: the Klein four-group $V_4 = \{e, (12)(34), (13)(24), (14)(23)\}$. This renders A_4 non-simple; indeed A_n is simple for all $n \geq 5$, making A_4 the unique non-abelian, non-simple alternating group.

This manifests in Lie theory through the decomposition SO(4) $\cong$ (SU(2) $\times$ SU(2))$/\mathbb{Z}_2$, splitting the Lie algebra $\mathfrak{so}(4) \cong \mathfrak{so}(3) \oplus \mathfrak{so}(3)$. This corresponds to decomposing 2-forms into self-dual and anti-self-dual components: $\Lambda^2(\mathbb{R}^4) = \Lambda^2_+ \oplus \Lambda^2_-$, where the Hodge star operator satisfies $\star^2 = 1$ on oriented Riemannian 4-manifolds (and $\star^2 = -1$ on 2-forms in Lorentzian signature), yielding eigenspaces with eigenvalues ± 1 in the Euclidean case.

In gauge theory, this splitting transforms second-order Yang–Mills equations into first-order conditions (Yang & Mills, 1954). A connection with curvature F satisfying $F = \star F$ (self-dual) or $F = -\star F$ (anti-self-dual) automatically solves $D \star F = 0$ since the Bianchi identity guarantees $DF = 0$. This dimensional coincidence—that the electromagnetic field strength is a 2-form and its Hodge dual has the same rank—makes Maxwell's equations acquire their most natural geometric expression in four dimensions. Here the equation $F = \pm \star F$ for 2-forms is specific to four dimensions; higher-dimensional analogues (e.g., G_2-instantons and Spin(7)-instantons) exist but differ in structure.

Donaldson's theorem, a hallmark of four-dimensionality (Donaldson, 1983), uses this. The moduli space of anti-self-dual connections on a 4-manifold yields polynomial invariants distinguishing smooth structures. Two homeomorphic 4-manifolds may have different Donaldson invariants, proving they are not diffeomorphic—a phenomenon occurring in no other dimension. The identity $4 = 2 + 2$ enables the entire apparatus.

So the identity $4 = 2 + 2$ creates the alternating group exception, the Lie algebra splitting, the self-duality decomposition, and the instanton solutions that distinguish four-dimensional gauge theory. Stable orbits, renormalizable interactions, exotic smooth structures—these may indeed stem from this single combinatorial fact. The most sophisticated features of our universe follow from the simplest patterns in the integers.

For more details on the underlying mathematics, I recommend the book *The Wild World of 4-manifolds* by Alexandru Scorpan (2005).

Why Four Might Be 'Special'

The convergence of independent constraints—orbit stability, renormalizability, Huygens principle, division algebras, gauge theory, the arithmetic identity $4 = 2 + 2$—all pointing to four dimensions invites three interpretations. First, four may encode a fundamental truth of geometry, where mathematical coherence uniquely selects this dimensionality as the only one supporting complex, predictable structures. Second, the apparent necessity may reflect selection bias: we observe four dimensions because observers can only arise where physics permits stable atoms and chemistry, rendering our conclusion inevitable yet uninformative about whether other dimensionalities 'exist' in some broader sense. Third, the entire exercise may be backfitting logic to a random parameter—finding post-hoc explanations for an arbitrary feature of our universe, mistaking coincidence for profundity.

The proliferation of independent mathematical arguments favouring four suggests the first interpretation, yet the arguments themselves presuppose frameworks (differential geometry, quantum field theory, group representation theory) constructed within and calibrated to a four-dimensional universe. Whether these constraints reveal something deeper or merely echo the assumptions embedded in our theories remains unresolved.

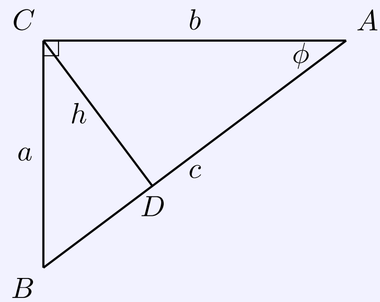
Exercise: Pythagorean Theorem via Dimensional Analysis

Derive $a^2 + b^2 = c^2$ using a dimensional argument involving area scaling.

1. **Scaling Argument**
   A right triangle is uniquely determined by its hypotenuse c and one acute angle ϕ. The only dimensionally valid expression for its area is

$$\Delta = c^2 f(\phi),$$

   where $f(\phi)$ is a dimensionless function. *(Why c rather than a or b? By similarity, the shape depends on ϕ and a squared length. Any side could be used, but expressing the result in terms of c emphasises the relation to the hypotenuse.)*



2. **Similar Triangles and Decomposition**
   The altitude CD from C to AB divides the original right triangle $\triangle ABC$ into two smaller right triangles, $\triangle CBD$ and $\triangle ACD$. By angle-angle (AA) similarity:

$$\triangle ABC \sim \triangle CBD \sim \triangle ACD.$$

   Because area scales as the square of a characteristic length, the sub-areas (corresponding to sides a or b) take the form $\Delta_1 = a^2 f(\phi)$ and $\Delta_2 = b^2 f(\phi)$. Since the total area is additive,

$$c^2 f(\phi) = a^2 f(\phi) + b^2 f(\phi).$$

   Dividing by $f(\phi)$ (assumed nonzero) gives $c^2 = a^2 + b^2$.

Dimensional analysis can be rigorously formalised using a graded algebraic structure where physical quantities lie in one-dimensional vector spaces V^d indexed by a group G of exponents (e.g., $\mathbb{R}^n$ for base units). Multiplication corresponds to the tensor product $V^a \otimes V^b \cong V^{a+b}$, and one only adds elements within the same space V^d. For example, a quantity in kilograms lies in $V^{(1,0,0,\dots)}$ and one in metres per second lies in $V^{(0,1,-1,\dots)}$, reflecting mass and velocity dimensions respectively. Physical laws hold under any consistent rescaling of units, leading to constraints like Buckingham's π-theorem. For details, see Tao (2012): https://terrytao.wordpress.com/2012/12/29/a-mathematical-formalisation-of-dimensional-analysis/

Dimensional Force Laws and Renormalizable Interactions

Notation: n denotes spatial dimensions; D denotes spacetime dimension ($D = n+1$ unless stated).

Flux Argument and $r^{-(n-1)}$ Fields. In n spatial dimensions, a spherical surface at radius r has 'area' scaling as r^{n-1}. For a source at the origin emitting flux uniformly, Gauss's law implies that flux per unit area decreases in proportion to $1/r^{n-1}$. Classical gravitational or electrostatic fields thus follow $F \sim \frac{1}{r^{n-1}}$. For $n = 3$, this becomes the familiar inverse-square relation. The corresponding potential $V(r)$ integrates (away from $n = 2$) as.

$$V(r) \sim \int \frac{dr}{r^{n-1}} \approx r^{2-n}.$$

When $n = 3$, $V(r) \sim 1/r$. For $n = 2$, $V(r) \sim \log r$.

Stable Orbits in Three Dimensions. A $1/r$ potential in $n = 3$ produces near-circular orbits that are stable under perturbations. Small changes in velocity cause bounded oscillations rather than catastrophic collapse or unbounded escape. In $n < 3$, forces decay more slowly (logarithmically at $n = 2$), creating strong long-range effects that disrupt stable Keplerian orbits. In $n > 3$, forces diminish rapidly, so small perturbations can disorder the trajectories.

From Classical to Quantum. This dimensional dependence also occurs in quantum physics. Atomic stability relies partly on the $1/r$ Coulomb potential in $n = 3$. In $n = 2$, the potential becomes logarithmic with qualitatively different bound states; for $n > 3$, the faster falloff reduces binding and can eliminate it at comparable scales.

Renormalizable Couplings. Quantum field theories (QFTs) further illustrate how dimensionality restricts allowed interactions. Consider a scalar field ϕ in D-dimensional *spacetime*. The ϕ^4 interaction $\mathcal{L}_{\text{int}} = \lambda\,\phi^4$ requires λ to be dimensionless or of non-negative mass dimension to avoid an infinite series of divergences. The mass dimension of ϕ is $[\phi] = (D-2)/2$, so

$$[\lambda] = D - 4[\phi] = 4 - D.$$

In $D = 4$ spacetime dimensions (signature $(3,1)$), λ is marginal (dimensionless). At $D > 4$, λ becomes irrelevant at high energy: the theory is non-renormalizable, demanding new terms for each new order in perturbation theory. For $D < 4$, the interaction is super-renormalizable with strong infrared effects.

Gauge Fields and Anomalies. In signature $(3,1)$, Yang–Mills gauge couplings are dimensionless, and the CP-odd topological density

$$\tfrac{1}{8\pi^2}\,\mathrm{tr}\,F \wedge F = \tfrac{1}{32\pi^2}\,\epsilon^{\mu\nu\rho\sigma}\,\mathrm{tr}\big(F_{\mu\nu}F_{\rho\sigma}\big)\,d^4x$$

integrates to an integer (the second Chern number) on compact 4-manifolds, independent of matter content. Gauge and mixed anomalies arise from chiral fermions; cancellation imposes constraints on representations, e.g., $\sum_{\text{fermions}} \mathrm{Tr}_R\big(T^a\{T^b, T^c\}\big) = 0$, where $\{T^b, T^c\} := T^bT^c + T^cT^b$ is the anticommutator (Jordan bracket) of the generators.

Analogous anomaly phenomena exist in other even spacetime dimensions, but renormalizable chiral gauge theories with dimensionless couplings occur naturally in signature $(3,1)$.

References:

Peskin, M. E., and Schroeder, D. V. (1995). *An Introduction to Quantum Field Theory.* Addison-Wesley.

Weinberg, S. (1995). *The Quantum Theory of Fields, Vol. I.* Cambridge University Press.

Ehrenfest, P. (1917). *In what way does it become manifest in the fundamental laws of physics that space has three dimensions?* Proceedings of the Amsterdam Academy.

Lenard, A., and Dyson, F. J. (1967). Stability of matter I, II. *Journal of Mathematical Physics.*