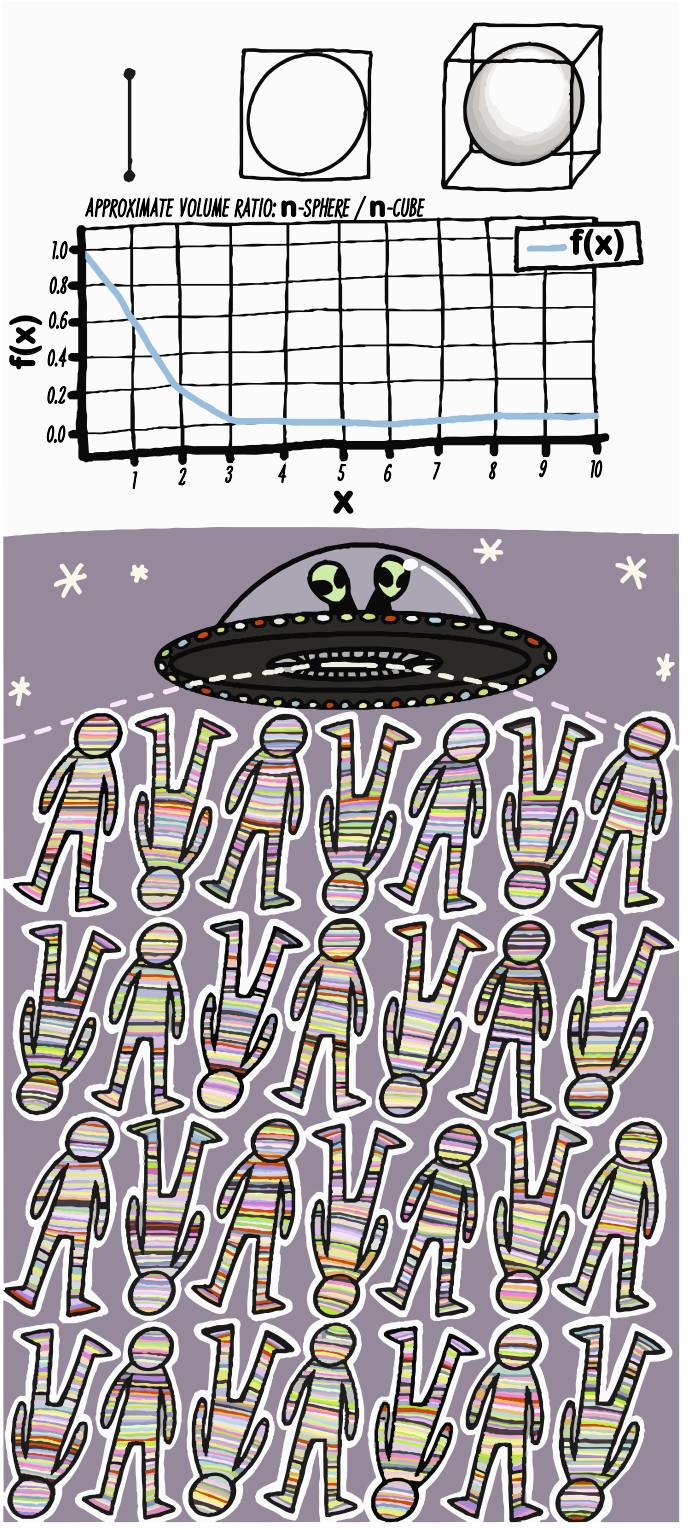
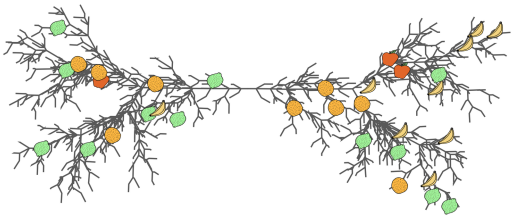
The Center Holds



Top (Curse of Dimensionality – Vanishing Volume of the Sphere): As dimensionality increases, the volume of an n-sphere relative to the surrounding n-cube rapidly shrinks toward zero. In low dimensions, the sphere fills most of the cube; by around ten dimensions, it's practically gone. This shows that intuition from 2D or 3D fails in high dimensions — most volume concentrates in the corners.
Bottom (Blessing of Dimensionality – Human Uniqueness in High-D Spaces): If each person is described by even fifty independent traits (drawn from uniform or Gaussian distributions), then the "average" human lies in a vanishingly small region of space. Almost everyone is in the high-dimensional fringes — radically unique combinations of attributes. High dimensionality ensures that individuality is not rare but inevitable.

The Center Holds

A geometric puzzle about Gaussian probability stumped mathematicians for over sixty years: prove that convex sets that are symmetric around the origin have enhanced overlap under Gaussian measure — that $P(A \cap B) \geq P(A) \cdot P(B)$. Despite partial results for boxes, ellipsoids, and slabs, the general case resisted all attempts. In 2014, Thomas Royen, a retired pharmaceutical statistician from a small German university, solved it using textbook methods: transforming to squared variables, applying Laplace transforms, and checking matrix determinants. His proof, published in an obscure journal, went unnoticed for years.



Gaussian Correlation Inequality ◦ Thomas Royen 2014 ◦ Convex Symmetric Sets ◦ High-Dimensional Geometry ◦ Concentration of Measure ◦ Squared Variables Method ◦ Laplace Transform Proof ◦ 60-Year Conjecture ◦ Multivariate Gaussian ◦ Elementary Solution ◦ Outsider Discovery

« Tout le monde y croit cependant, me disait un jour M. Lippmann, car les expérimentateurs s'imaginent que c'est un théorème de mathématiques, et les mathématiciens que c'est un fait expérimental. »

("Everyone believes in it, Mr. Lippmann told me one day, because experimentalists imagine it is a mathematical theorem, and mathematicians that it is an experimental fact.")
— Henri Poincaré on the Gaussian Distribution, 1912

The Center Holds

Interest in how Gaussian measures behave under geometric constraints emerged in the mid-twentieth century, particularly in multivariate statistics and convex geometry. By the 1950s, researchers studying elliptical distributions began formulating conjectures about the probability content of intersections between symmetric convex regions.

The modern form of the Gaussian Correlation Inequality (GCI) was solidified in the 1970s through work by Das Gupta, Olkin, Pitt, and others, who framed it in terms of standard Gaussian measures over $\mathbb{R}^n$. They asked whether Gaussian probability favors overlap: specifically, whether the measure of the intersection of two symmetric convex sets is always at least as large as the product of their individual measures. The conjecture attracted attention because it combined natural geometric symmetry with the most analytically tractable probability distribution.

Over the following decades, progress was made in restricted settings. The inequality was proven for two-dimensional cases, for coordinate-aligned boxes, and for ellipsoids. The partial results relied on tools from real analysis, measure theory, and convex optimization. The general case resisted all attempts, despite appearing elementary in formulation.

In 2014, a breakthrough came from Thomas Royen, a retired statistician with a background in pharmaceutical applications. Royen published a short paper that resolved the inequality in full generality. His approach was elementary in the technical sense: it used standard tools, required no heavy machinery, and invoked only modest linear algebra and probability. Nonetheless, it connected several overlooked identities in a way that previous attempts had not. Although initially unnoticed, Royen's proof was soon verified and reformulated in expository papers by Latała, Matlak, and others, and has since been accepted as the definitive solution to the GCI.

The Gaussian distribution, also called the normal distribution (Gauss, 1809), is defined by the density function

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2},$$

which describes the standard normal distribution, denoted $\mathcal{N}(0,1)$ — zero mean, unit variance. More generally, a Gaussian with mean μ and variance σ^2, denoted $\mathcal{N}(\mu, \sigma^2)$, has density $\frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/(2\sigma^2)}$; the parameter σ controls the width of the bell. The standard form above corresponds to $\mu = 0, \sigma = 1$. The function is symmetric about the mean, with values decreasing smoothly as $|x - \mu|$ increases. The rate of decay is exponential in the square of the distance, causing values far from zero to be exponentially rare. The total area under the curve is normalized to one, making it a valid probability distribution. Its characteristic bell shape is recognized as the canonical model for random variation in natural and statistical systems.

The bell curve emerges as the limiting form of many sums of random variables. Consider a process of repeatedly rolling a fair die and averaging the results. Although each individual roll yields a uniformly distributed outcome on a discrete set, the distribution of the

©2026 David H. Silver, CC BY-NC 4.0

average becomes increasingly smooth and Gaussian as the number of trials grows. The Central Limit Theorem states (Laplace, 1812) that the sum (or average) of independent, identically distributed variables with finite variance converges in distribution to the Gaussian, regardless of the original distribution.

This universality explains the Gaussian's omnipresence in nature. Heights in a population result from countless genetic and environmental factors — each contributing a small push up or down. Measurement errors accumulate from vibrations, temperature fluctuations, and quantum uncertainties. Stock prices reflect millions of independent trading decisions. In each case, myriad small influences combine additively, and their sum inevitably forms a bell curve. The Gaussian is not imposed by theory but emerges from the arithmetic of aggregation. The theorem explains why Gaussian distributions appear ubiquitously in statistical mechanics, measurement theory, and signal processing — wherever many small effects compound.

The multivariate Gaussian generalizes this form to $\mathbb{R}^n$. The standard form has density

$$\varphi_n(x) = \frac{1}{(2\pi)^{n/2}} e^{-\|x\|^2/2},$$

where $\|x\|$ is the Euclidean norm of the vector $x \in \mathbb{R}^n$. The distribution is spherically symmetric: it assigns equal probability density to all points equidistant from the origin. Its contours are concentric spheres, and its value depends only on the radial distance. Every linear projection of the distribution onto a one-dimensional axis yields a standard univariate Gaussian. In the standard case, the coordinates are independent and identically distributed $\mathcal{N}(0, 1)$; by rotational invariance, this remains true in any orthonormal basis. Rotational invariance and marginal stability make the multivariate Gaussian a tractable object in high-dimensional probability.

In high dimensions, the geometry of Gaussian measure becomes profoundly unintuitive. Although the density is highest at the origin, the bulk of the probability mass concentrates near a thin spherical shell of radius approximately $\sqrt{n}$. This defies our three-dimensional experience: you might expect that since the Gaussian density peaks at the center, most random points would be found there. This is completely wrong.

This is a result of the explosive growth of the number of points at a given radius. Consider an orange inside a cubic box. In three dimensions, the sphere fills a decent portion of the box. But as dimensions increase, the hypercube's corners dominate overwhelmingly. In 1000 dimensions, over 99.999% of the hypercube's volume lurks in its corners, not near the center. The surface area of a sphere of radius r in $\mathbb{R}^n$ grows proportionally to r^{n-1}; in the radial density for $\|X\|$, this factor competes with the exponential term and pushes mass toward a thin shell.

The result is concentration of measure, which transforms probabilistic problems into geometric ones. A Gaussian random vector lies in a given region when that region intersects this nearly fixed-radius shell, rather than when it captures values near the origin.

The Gaussian Correlation Inequality concerns the probability that a standard Gaussian random vector $X \in \mathbb{R}^n$ simultaneously falls into two geometric regions. Let $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^n$ be closed, convex sets that are symmetric about the origin. Then the inequality

states:

$$\mathbb{P}(X \in A \cap B) \geq \mathbb{P}(X \in A) \cdot \mathbb{P}(X \in B).$$

The left-hand side is the probability that a single Gaussian sample lies in both sets, while the right-hand side is the product of the probabilities of lying in each separately. No notion of parametric correlation appears in this formulation — no Pearson coefficient, no covariance matrix interaction. The term "correlation" here is geometric: it measures the extent to which the spatial configurations of the sets align so that overlap under the Gaussian measure is enhanced. Symmetric convex sets interact positively under Gaussian sampling.

Imagine a dartboard in high-dimensional space. Two target zones — each convex and mirror-symmetric about the center — are drawn on the board. The dart is thrown not with uniform probability, but according to a Gaussian distribution. In our familiar world, you would expect the dart to land near the bullseye where the density is highest. But in high dimensions, the dart almost surely lands on a distant shell at radius $\sqrt{n}$. The magic of the GCI is that despite this shell phenomenon, symmetric convex sets still manage to overlap more than independence would predict. Their enforced central fatness — they cannot be hollow or lopsided — creates enough overlap at the origin's high-density region to overcome the dilution effect of the shell.

Both symmetry and convexity are essential to the validity of the inequality. If either condition is relaxed, the result can fail. For example, consider two non-convex shapes such as disconnected spherical caps placed symmetrically on opposite sides of the origin. Each may individually capture moderate Gaussian mass, but their intersection can be empty, rendering the left-hand side of the inequality zero while the right-hand side remains positive. Alternatively, take two convex balls shifted away from the origin in opposite directions: each maintains convexity, but the loss of symmetry means their overlap under Gaussian measure can be arbitrarily small, violating the inequality.

The unusual difficulty of proving the GCI arose from a geometric tug-of-war in high-dimensional space. The concentration of measure pushes probability mass outward to a distant shell, suggesting that intersection should be difficult — sets must somehow coordinate their overlap on this fragile, specific radius. But convexity and symmetry pull in the opposite direction: these shapes must be "fattest" at the center, they cannot be hollow or have their mass pushed outward. The conjecture, now proven, asserts that the central pull always wins.

Several equivalent formulations exist. The indicator function $\mathbf{1}_A(X)$ equals 1 when $X \in A$ and 0 otherwise, so its expectation is exactly the probability: $\mathbb{E}[\mathbf{1}_A(X)] = \mathbb{P}(X \in A)$. Rewriting the inequality in this notation,

$$\mathbb{E}[\mathbf{1}_A(X) \cdot \mathbf{1}_B(X)] \geq \mathbb{E}[\mathbf{1}_A(X)] \cdot \mathbb{E}[\mathbf{1}_B(X)],$$

the content is identical to the probability statement above, but the language shifts from set membership to random variables. This rephrasing is useful because it reveals the GCI as a *correlation* inequality in the literal statistical sense: the covariance $\mathrm{Cov}(\mathbf{1}_A(X), \mathbf{1}_B(X)) \geq 0$. In other words, the two indicator random variables are nonnegatively correlated under

the Gaussian measure — knowing that X lies in one symmetric convex set can only increase (never decrease) the likelihood that it lies in the other.

The Gaussian Correlation Inequality was conjectured in the 1950s and resisted proof for over six decades. During this time, it was confirmed in numerous special cases. For axis-aligned rectangles (boxes), the result was established by Šidák (1967). Other special cases — such as slabs and certain families of ellipsoids — were also resolved. Yet no general method succeeded. Classical techniques — log-concavity of Gaussian measure, the Brascamp–Lieb inequality, and concentration of measure (Brascamp & Lieb, 1976) phenomena — yielded related inequalities but stopped short of establishing the required correlation bound for arbitrary convex symmetric sets.

The proof came not from a well-known probabilist or a high-profile research program, but from Thomas Royen, a retired statistician at a university of applied sciences in Bingen, Germany. Royen had worked for decades in applied statistics, particularly in pharmaceutical research. His academic career was spent outside the core research institutions of probability theory, and his publication record was modest by conventional standards. The outsider status provided the freedom to pursue classical problems without disciplinary constraint. Royen's mathematical training was solid but practical, shaped by applications and experience. He approached the problem of Gaussian correlation not as a convex analyst but as a statistician with an eye for transformations and distributions.

The central move in Royen's proof was to reframe the inequality in terms of squared Gaussian variables. By passing to variables of the form X_i^2, he translated the problem into one involving sums of independent gamma-distributed variables. The transformation allowed the introduction of Laplace transforms (LTs) — a standard tool in distributional analysis — and reduced the problem to showing monotonicity of a certain function defined by determinants of parameter-dependent covariance matrices. Royen employed an identity involving the determinant of a positive semi-definite matrix perturbed by diagonal terms, and used it to establish the required inequality via monotonicity in a parameter. The argument was elementary in the sense that it involved no modern theorems, but subtle in its reconfiguration of the problem into a tractable analytic form.

Despite the proof's correctness, Royen's paper initially went unnoticed. It appeared in a minor journal and lacked the formal polish typically expected of breakthroughs in high-dimensional analysis. The paper did not announce its significance, and the style — direct and sparse — obscured its novelty. For a time, the result was known only to a small circle of readers, many of whom were unsure whether the argument was valid. Eventually, experts in probability and convex geometry began to scrutinize the proof, rephrasing and streamlining its components. Within a few years, the result was confirmed, disseminated, and reformulated in the language of convex analysis and Gaussian processes. Royen's name entered the canonical history of the problem, and the Gaussian Correlation Inequality was marked solved. What remained was not only a resolution of the inequality itself, but a reminder that the landscape of mathematical solutions includes not only new theories, but new configurations of old tools — found sometimes at the margins of the research world.

Unexpected Solvers with Familiar Tools

The story joins others in this book where longstanding open problems were resolved not by new machinery, but by the careful use of classical methods in unfamiliar configurations — often by researchers outside elite institutions. Like Yitang Zhang's breakthrough on bounded prime gaps, or the amateur discovery of the monotile known as the "hat," Thomas Royen's proof of the Gaussian Correlation Inequality relied on known identities and transforms applied with unusual directness. Another case is the anonymous 4chan poster who, responding to a question about anime episode orderings, derived a new lower bound on superpermutation length using standard graph-theoretic arguments — a result later verified and cited in formal literature as belonging to an unidentified author (Ch. 39). The cases share a common story: problems that resisted decades of expert attention gave way once the right pathway — already present in the mathematical landscape — was followed with formal rigor.

This is unusual. Almost always, when someone claims to have solved a famous open problem, it is crankery. The phenomenon spans the entire spectrum of mathematical sophistication. At one end: amateurs on Quora insisting they have disproven momentum conservation or constructed a perpetual motion machine, unaware of basic definitions. In the middle: professors at respectable institutions who become obsessed with problems adjacent to their expertise, producing hundreds of pages of arguments that experts dismiss within minutes. At the high end: world-renowned experts who announce breakthroughs in areas outside their domain — claiming, for instance, to have proven the Riemann hypothesis — only to have fatal errors exposed during peer review.

The most contentious cases occur when the claimant's reputation and technical sophistication make dismissal difficult. Shinichi Mochizuki's claimed proof of the abc conjecture (Masser, 1985; Oesterlé, 1988) (which posits that when a, b, and c are coprime and satisfy $a + b = c$, then c is rarely much larger than the product of the distinct primes dividing abc), spanning over 500 pages of novel theory he calls "inter-universal Teichmüller theory," has divided the mathematical community for over a decade. Leading number theorists have declared the proof fatally flawed, while Mochizuki and a small circle of collaborators maintain its validity. The dispute remains unresolved — not for lack of expertise on either side, but because the proposed framework is so idiosyncratic that consensus on its correctness may be unattainable. What separates legitimate breakthroughs from crankery is not the solver's credentials, but whether the proof can be verified, communicated, and integrated into the broader body of mathematical knowledge.

Monotonicity via Covariance Interpolation

Let $X \sim \mathcal{N}(0, C)$ be an n-dimensional Gaussian vector with zero mean and covariance matrix $C \succcurlyeq 0$. The Gaussian Correlation Inequality asserts that for any symmetric convex sets $A, B \subset \mathbb{R}^n$,

$$\mathbb{P}(X \in A \cap B) \geq \mathbb{P}(X \in A)\,\mathbb{P}(X \in B).$$

We consider the axis-aligned box case; the full GCI for all symmetric convex sets follows from Royen:

$$A = \{x \in \mathbb{R}^n : |x_i| \leq 1 \text{ for } 1 \leq i \leq k\},$$
$$B = \{x \in \mathbb{R}^n : |x_j| \leq 1 \text{ for } k < j \leq n\}.$$

Let $X = (X_1, \dots, X_n)$, and define

$$f(t) := \mathbb{P}_t\left(\max_{1\leq i\leq n} |X_i| \leq 1\right),$$

where $\mathbb{P}_t$ denotes a Gaussian measure with interpolated covariance

$$C(t) = \begin{pmatrix} C_1 & tQ \\ tQ^\top & C_2 \end{pmatrix}, \quad t \in [0,1].$$

Here, $C_1 \in \mathbb{R}^{k\times k}$, $C_2 \in \mathbb{R}^{(n-k)\times(n-k)}$, and $Q \in \mathbb{R}^{k\times(n-k)}$. Write the original covariance in the same block form, $C = \begin{pmatrix} C_1 & Q \\ Q^\top & C_2 \end{pmatrix}$, so that $C(1) = C$ and $C(0) = \operatorname{diag}(C_1, C_2)$. At $t = 0$, the covariance is block-diagonal with independent blocks; at $t = 1$, the off-diagonal coupling Q is fully present. Since $C(t) = (1-t)C(0) + tC(1)$ and the positive semidefinite cone is convex, $C(t) \succeq 0$ for all $t \in [0,1]$.

Note $\{\max_{1\leq i\leq n} |X_i| \leq 1\} = A \cap B$. At $t = 0$, the block-diagonal covariance makes $(X_1, \dots, X_k)$ and $(X_{k+1}, \dots, X_n)$ independent (Gaussian), so $\mathbb{P}_0(A \cap B) = \mathbb{P}_0(A)\mathbb{P}_0(B)$. The goal is to prove that $f(t)$ is non-decreasing.

Transformation to Gamma Structure

The squared Gaussian variables $Z_i = X_i^2/2$ follow a scaled chi-squared law. For $\lambda_i \geq 0$, define the Laplace transform of $Z = (Z_1, \dots, Z_n)$ under $\mathbb{P}_t$:

$$\begin{aligned}\mathcal{L}_t(\lambda) &= \mathbb{E}_t\left[\exp\left(-\sum_{i=1}^n \lambda_i Z_i\right)\right] \\ &= \mathbb{E}_t\left[\exp\left(-X^\top \Lambda X/2\right)\right] \\ &= \det(I + C(t)\Lambda)^{-1/2},\end{aligned}$$

where $\Lambda = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$. Following Royen, differentiate $\log \mathcal{L}_t(\lambda)$ and rewrite the derivative as an expectation under a multivariate gamma law (see Royen Theorem 1 and its gamma mixture representation; Latała–Matlak (2000) §2–§3); the integrand is nonnegative, hence $\mathcal{L}_t(\lambda)$ is non-increasing in t, and hence $f(t)$ is non-decreasing.

Smoothing and Differentiation

To handle the indicator function rigorously, define a smooth approximation: $\phi_\epsilon(x) = 1$ if $|x| \leq 1 - \epsilon$, $\phi_\epsilon(x) = 0$ if $|x| \geq 1 + \epsilon$, and ϕ_ϵ is smooth monotone otherwise. Let $F_\epsilon(Z) = \prod_{i=1}^n \phi_\epsilon\left(\sqrt{2Z_i}\right)$, so that $F_\epsilon \to 1_{\{\max|X_i|\leq 1\}}$ as $\epsilon \to 0$. Using Royen's multivariate gamma representation, $\frac{d}{dt}\mathbb{E}_t[F_\epsilon(Z)]$ is an integral of a nonnegative kernel; dominated convergence then gives ≥ 0. Passing $\epsilon \to 0$ yields the monotonicity of $f(t)$, establishing the Gaussian Correlation Inequality for axis-aligned boxes. The full inequality for all symmetric convex sets follows from Royen (2014).

References:

Royen, T. (2014). *A simple proof of the Gaussian correlation conjecture extended to multivariate gamma distributions.* Far East J. Theor. Stat.

Latała, R., Matlak, D. (2017). *Royen's Proof of the Gaussian Correlation Inequality.* In: Israel Seminar (GAFA) 2014–2016. Springer.