Article

Not peer-reviewed version

The Economy of the Lie Squad

Julio Rives *

Posted Date: 6 March 2026

doi: 10.20944/preprints202603.0536.v1

Keywords: positional notation; Euclidean space; representational cost; Newcomb-Benford law; logarithmic scale; simple lie group; irreducible representation; divisibility

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2026 doi:10.20944/preprints202603.0536.v1

Article

The Economy of the Lie Squad

Julio Rives

Universidad Nacional de Educación a Distancia (UNED), Spain; jrives@dia.uned.es

Abstract

We investigate the mechanisms by which natural systems encode data across multi-dimensional spaces. Integrating principles from information theory, probability, and geometry, we propose that certain Lie Groups govern these encoding processes. We first demonstrate that evenly distributed information becomes computationally unsolvable in higher dimensions. If we do not notice the " curse of dimensionality," it is because nature likely uses geometric positional notation at a rudimentary level. By extending the definition of representational cost to m dimensions using Benford' s Law, we identify a cost minimum at powers of Euler' s number (e^m). We introduce the " Lie Squad" (B3, F4, G2, A2, A1, and E6), a set of six compact simple Lie groups whose irreducible representations coincide with this ideal cost when m matches the group' s algebraic rank. These irreps facilitate a fundamental, rank-invariant number system based on balanced ternary, uniquely encoding integers as the difference of two natural numbers in bijective notation. Finally, we examine the Weyl orders of the Lie Squad members to show that Weyl divisors yield a logarithmic scale consistent with Benford' s Law and the universal number system proposed.

Keywords: positional notation; Euclidean space; representational cost; Newcomb-Benford law; logarithmic scale; simple lie group; irreducible representation; divisibility

MSC: 11A63; 11A67; 11B50; 14L35; 17B10; 17B20; 22E15; 51M05; 51M10; 51M20; 51M25; 60B99; 68P30; 94A15; 94A17

---

1. Introduction

This article analyzes the idea that the laws of physics come from structured mathematical objects. These laws embody economy (the efficient use of resources), information (how data is kept and communicated), and symmetry (what remains unchanged under particular transformations). The goal is to show how nature optimally stores data across spaces of different dimensions, including 2-, 3-, and higher-dimensional spaces. Here, we propose that Lie groups govern the encoding of information.

Quantum mechanics, thermodynamics, and relativity support the view that information is physical. Information is tied to energy and space. Wheeler's 'It from Bit' idea [1] suggests that physical things arise from binary choices. Landauer's Principle [2] proves that erasing one bit of information releases energy. If information equals energy, it can change the geometry of spacetime, according to General Relativity [3]. The Holographic Principle [4] states that information in a region is distributed along its boundary. The Black Hole Information puzzle [5] holds that information absorbed by a black hole remains at its event horizon, thereby limiting the information density and potentially pointing to a fifth state of matter [6]. The Space Memory Theory [7] proposes that information is encoded in the geometry of spacetime, composed of tiny Planck-scale units that mediate quantum interactions.

Information and physics are deeply connected via uncertainty, commonly explored through probability. Entropy, for instance, quantifies how surprising a random event is on average; unexpected events are less probable and more informative. How information is represented is also tied to the likelihood, the space allocated to encode a number between a range' s ends on a harmonic or logarithmic scale [8]; Newcomb-Benford Law (NBL) probabilities can be viewed as normalized likelihoods.

 doi:10.20944/preprints202603.0536.v1

If information, probability, and space are tightly linked, it is natural to expect shared structural features. A probability space models a physical scenario, just as the geometry of spacetime can give rise to quantum probabilities, as demonstrated in the Ryu-Takayanagi formula [9]. Adjusting the entanglement between regions can shift their relative distances. Extending this analogy, Susskind and Maldacena [10] propose that 'probability links' (Einstein-Podolsky-Rosen correlations) are analytically equivalent to 'spatial links' (Einstein-Rosen bridges or wormholes). Similarly, in Optimal Transport Theory [11], the efficient redistribution of probability distributions is tied to curvature and geometric stability.

Following this conceptual thread, this article examines space through the lens of topology (the mathematical study of properties maintained under continuous deformations) and algebra (the study of symbols and rules for manipulating them). In any dimension, a sphere is used to model regular space, measured by area and volume. If a ball is thought of as being constructed from digits, summing the probabilities for each digit is like measuring its volume. This observation elicits questions about how these probabilities change as we add more dimensions. To approach this, it is helpful to view space as both a smooth manifold (a shape that locally resembles flat space and supports calculus) and as a Lie group (a set of continuous symmetry mappings with differentiable operations).

Yet, representing information becomes more challenging as we move to multi-dimensional spaces, where the number of independent information directions increases. In this scenario, familiar notions of distance and size break down. As dimensions rise, most of a shape's volume concentrates near its surface, producing data points so far apart that nearest-neighbor methods turn ineffective. For instance, the volume of a ball shrinks dramatically compared to that of its containing hypercube. Additionally, retaining consistent data density as dimensions rise requires exponentially more points per dimension. This makes data representation increasingly resource-intensive, motivating the search for more efficient information structures.

Given the peculiarities of the high-dimensional Euclidean space, hyperbolic information spaces become relevant. A member of a Lie group can recast data under an efficient positional number system, which is intricately connected to the Newcomb–Benford Law (NBL). This matters because many real-world datasets follow this law, predicting that small digits occur more frequently than large ones [12]. The uniqueness of NBL as a scale-invariant probability distribution on the foremost digits is established in [13], and scale-invariance implies radix-invariance [14]. Using NBL, we reformulate the definition of representational cost in Geometric Positional Notation (GPN), where a number is written as a geometric sequence of place-value digits, to measure how compactly multi-dimensional data can be stored.

Information is closely linked to Lie groups. Lie groups are sets of smooth transformations that keep the basic properties of objects. These changes make it easier to describe a system. For something symmetrical, you need fewer details if you know the Lie group rules. Noether's theorem states that these symmetry rules correspond to quantities that remain constant as the system changes, such as energy under time translation. Moreover, Fisher information quantifies the amount of information that an observable random variable carries about an unknown parameter of its distribution. Specifically, it is the variance of the gradient of the log-likelihood function with respect to the parameter vector [15]. When we can associate the parameters of a probability setup with the elements of the Lie group, Fisher information induces an invariant on the Lie group; the Fisher matrix defines the Riemannian metric tensor on the Lie group manifold and lets us measure the magnitude of the symmetry transformation by the shortest path in the group's Lie algebra.

From this perspective, an information space should appear uniform when acted on by a Lie group [16,17], meaning that all transformations, i.e., translations (sliding), rotations (turning), and reflections (mirroring), act consistently everywhere within the space. Using a smaller group, known as a discrete subgroup (which allows only particular discrete transformations), we can construct repeating space-filling patterns called tessellations (tilings without gaps). Certain Lie groups are fundamental to constructing more complex symmetries. A Simple Lie Group (SLG) is an indivisible

 doi:10.20944/preprints202603.0536.v1

continuous symmetry group; it cannot be broken into smaller, simpler normal subgroups. Any complex, finite, connected Lie group can be built from these simple components. Thus, we focus on the principal examples of SLGs, those that in principle play a decisive role in the configuration of spacetime and matter. The Standard Model, a framework in physics based on Lie groups and gauge theory (a mathematical scheme for describing forces and particles), applies precise rules within a local coordinate system subject to mathematical constraints. A 'gauge' in our context is the ruler that sets the SLG logarithmic information scale.

A chief observation is that six SLGs have basic forms, called irreducible representations (irreps) [18], with dimensions that match Euler's number (e $\approx$ 2.718) raised to the power of the group's rank (the number of independent directions in its inherent algebraic structure). An irrep represents a basic symmetry and can be shown as a matrix that cannot be split into smaller, independent parts. The $E_6 \supset F_4 \supset B_3 \supset G_2 \supset A_2 \supset A_1$ chain, where "$\supset$" means " embeds" , is the sequence listing the thriftiest compact SLGs, each with an irrep size close to the optimal radix choice for its rank. For example, the 52-dimensional irrep of F4 is about e^4. Is this just by chance?

Then we define the specific positional notation system used by the Lie Squad. It must be fundamental, i.e., universal and unambiguous. We do not have many choices. We propose SBN, a zero-free number system that does not depend on whether the irrep order is even or odd, nor on the sign of the integer being written. SBN is based on balanced ternary and encodes every integer as a difference of two naturals. SBN satisfies the NBL for bijective numeration, which is position and length compliant, in addition to being scale-invariant. Moreover, SBN is rank-invariant, meaning that nature does not need to account for the ambient space when adding or removing a spatial dimension. Each component is encoded in SBN radix r^m, where m is the rank, and a vector coding is the composition of the m components. For example, in G2, we write the decimal 987654321 as cgefggffecb$-$dbdfacdgef.

The Weyl group of a Lie group is the collection of symmetries that act simply transitively on the chambers (regions divided by hyperplanes) of the group's root system [19]. Its divisors (numbers that divide the group's Weyl order without remainder) link to the adjoint representation (how the group acts on its own algebra). It is known that the degrees of the independent polynomials on the Cartan subalgebra (the maximal set of commuting elements) that are invariant under the Weyl group are specific divisors of the Weyl order and intrinsic structural constants of the Lie group. Well, we have discovered that these divisors also allow us to find a probability distribution that closely matches NBL with an optimal radix. To this end, we compute the tensor product of the Weyl divisors and search for sequences whose length equals the most informative SLG's irrep size. Then, for all these sequences, we calculate the average sum-of-divisors function for each element. We always find a sequence that approaches the ideal NBL.

We have employed the Anderson-Darling, Kolmogorov-Smirnov, Kuiper, Pearson's chi-squared, Watson U^2, and Cramér-von Mises goodness-of-fit tests for small samples, in particular to verify that the selected ireeps of A1, A2, G2, and B3 are NBL-compliant. However, these methods regularly reject the null hypothesis for medium- or large-sized samples even when there is clear visual agreement. To reduce excess power in samples such as those corresponding to F4 and E6, we use cross-validation. Specifically, we use the aforementioned goodness-of-fit tests across validation set partitions of length 9. However, to strengthen empirical testing and provide sufficient statistical support for the conclusions, we additionally calculate the Relative Root Mean Squared Error (RRMSE= $\sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{(n\sum_{i=1}^{n}\hat{y}_i^2)}}$) regardless of the sample, as described by [20] and interpreted as [21], to provide a normalized measure of the mean absolute deviation between the empirical and expected distributions.

The structure of this article is as follows. First, the curse of dimensionality is examined, demonstrating that evenly distributed information becomes unmanageable in higher dimensions. Subsequently, we discuss the need to consider non-Euclidean information spaces, as natural efficiency favors GPN systems. The concept of representational cost is then extended from one dimension to higher dimensions. It is argued that higher-dimensional algebraic forms should possess an order approximating the optimal radix for their rank. The article introduces the most significant SLGs and

 doi:10.20944/preprints202603.0536.v1

their most informative representations. The 'Lie Squad' refers to a set of six compact SLGs whose algebraic representations closely align with the ideal radix e^m, where m is the algebraic rank of the SLG. These representations give rise to a universal, zero-free, GPN system, SBN, with illustrative examples provided. Finally, we examine the Weyl order of the Lie Squad members and derive the associated logarithmic NBL from the Weyl divisors. The conclusion reviews and summarizes the principal findings.

The results in this paper reveal deep connections among information, probability, and geometry. If information is physical, then probability is also physical, and hence, considering symmetry and divisibility, geometry is physical, too. The way space is structured inside hyperbolic geometry may explain why the NBL, inherently linked to GPN, matches what we observe in reality. This could be why raw samples of certain natural properties appear as data encoded on a logarithmic scale.

2. Curse of Dimensionality

Assuming that information is correlated with space, how higher-dimensional worlds represent data?

First, we need to consider whether it is physically necessary and possible to create and maintain extra dimensions, given the energy and resources required. Physicists often add new dimensions to make an " effective theory" work, even if nature may not have them. Often, adding dimensions is just a quick fix for problems, especially those caused by the proliferation of infinities. For instance, in string theory, keeping extra dimensions stable is difficult because of the steep energy barriers of the resulting " moduli" fields [22]. Therefore, creating a new dimension should only be done if it brings real benefits.

A requirement of productivity can motivate the creation of extra dimensions based on the concept of symmetry. Increasing symmetries guarantee effectiveness by simplification; a part can serve to process the whole.

Extra dimensions could also be motivated by a requirement of randomness based on the principle of maximum entropy. Random walk and Brownian motion are the discrete and continuous versions of a point particle moving in space, randomly changing directions, the archetypal example of elemental translation movement. Per Donsker's invariance principle [23], we can obtain Brownian motion by taking scaled copies of a random walk and a limiting distribution. What is the probability of a random walk going back to the starting position? In 1DE (1-Dimensional Euclidean) space, the probability is 1 (point recurrent); in 2DE, the path returns to the vicinity of the origin with probability 1 (only neighborhood recurrent), but in 3ED and higher spaces, the probability is less than 1 (all points are transient). While the route reaches every point infinitely many times in 2DE, it diverges in 3DE and higher dimensions. Likewise, a Brownian motion path will eventually leave an n-sphere for $n > 2$, regardless of the radius, and will return to it only finitely many times [24]. How many times does a Brownian motion path cross itself? We can find infinitely random points where the path crosses itself exactly $p > 0$ times in 2DE, whereas there are infinitely many points in 3DE where it crosses itself exactly once, but no double or higher self-crossing places exist, and the path never crosses itself in 4DE or higher dimensions. Besides, two uncorrelated Brownian motions will cross infinitely many times in dimensions lower than 4DE, whereas in 4DE and higher dimensions, they will " almost surely" not intersect if we exclude the origin. Furthermore, three mutually independent Brownian motions will jointly coincide at a point infinitely many times in 1DE and 2DE, whereas in 3DE and higher, they will never cross if we exclude the origin. To summarize, 1DE and 2DE spaces behave pretty differently from 3DE and 4DE spaces concerning randomness, and 4DE (and higher) is the most inherently random and unfettered space.

Although greater symmetry and randomness can enrich physics, they come at a computational price. As dimensions multiply, wrangling data becomes a major challenge. Richard Bellman [25] introduced the term " curse of dimensionality" to describe how high-dimensional spaces differ from three-dimensional space; they become vast, sparse, and with most of the volume in the corners.

A key geometric issue in high dimensions is that the volume of a unit m-ball decreases rapidly as m increases, starting at dimension 5. For example, with radius $R = 0.5$, the volume is more than halved with each additional dimension; for $R > 1$, the contraction is even steeper. As m grows, the region containing most data shrinks, making distinctions harder to observe. High-dimensional space becomes compressed.

The problem is consubstantial with geometry and unrelated to the working method or arithmetic. Any fixed, independent, and identically distributed dataset on the one-dimensional Euclidean space induces a product distribution of points in mDE space. Assuming dimensional orthogonality, the Euclidean distance from the center to the corners of the m-cube with edges of length $2R$ is $\sqrt{R^2(1_1^2 + \ldots + 1_m^2)} = R\sqrt{m}$. Likewise, any pair of points (data, quanta, numeral, etc.) separated by a distance 1 in each direction will actually be separated by a distance $\sqrt{1_1^2 + \ldots + 1_m^2} = \sqrt{m}$.

We can examine the issue from a different angle. Since the rate of the m-dimensional volume of a Euclidean ball to the m-cube of diameter and edge $2R$ is

$$\frac{V_m(R)}{(2R)^m} = \frac{\frac{\pi^{\frac{m}{2}}}{\Gamma(1+\frac{m}{2})}R^m}{(2R)^m} = \frac{\frac{\pi^{\frac{m}{2}}}{\Gamma(1+\frac{m}{2})}}{2^m}$$

most of the volume of the hypercube quickly accumulates between the inscribed and the circumscribing hyperspheres as m increases. For example, this rate is 1 for $m = 0$ and $m = 1$. In the Euclidean plane, the ratio of the size of an inscribed 2-ball (a circle) to that of a 2-cube (a square) is $\frac{\pi}{4} \approx 78.54\%$, which intuitively seems sensible. It is $\frac{\pi}{6} \approx 52.36\%$ if $m = 3$, meaning that almost half the volume of a normal cube is outside its inscribed ball, already conveying some distortion. In four dimensions, a great part of the hypercube is practically empty, since $1 - \frac{\pi^2}{32} \approx 69.16\%$ of it is outside the inscribed 4-ball. In E6, the hypercube mainly consists of corners far away twice as much as the boundary of its inscribed 5-sphere, within which only $\frac{\pi^3}{384} \approx 8\%$ of the 6-ball volume resides.

The circle of $2R$ is $\frac{V_2(2R)}{V_2(R)} = 4$ times as much as the size of the circle R, so that 75% of randomly selected points are going to be within a distance in the range $(R..2R]$ from the center. Two floors higher, $1 - \left(\frac{3}{2}\right)^{-4} \approx 80\%$ of the volume occupied by a 4-ball with radius $\frac{3R}{2}$ is outside the concentric 4-ball with radius R. Regardless of the number of dimensions, the bulk of the space moves away from the center. Could we at least avoid this hindrance?

We insist that this is not all about hypercubes or particular shapes. From a statistical or machine learning perspective, consider the formulae of the volume of an m-ball. The radius needed to obtain a fixed volume has an order of growth $\frac{\sqrt[x]{\Gamma(1+\frac{m}{2})}}{\sqrt{\pi}} \propto \sqrt{m}$. For instance, $V_6(0.76) \approx 1$, $V_{12}(0.98) \approx 1$, $V_{18}(1.15) \approx 1$, $V_{24}(1.3) \approx 1$, and $V_{1000}(7.68) \approx 1$. We need a double-exponentially growing amount of examples (rows), proportional to $m^{\frac{m}{2}}$, with every additional dimension, predictor, or feature (column m) to qualitatively cope with the whole range of possibilities while keeping a fixed even-spacing of sample points and averting dispersion of datapoints. In short, equivalent volumes of distinct dimensions differ dramatically with respect to the number of combinations that assure smooth compactness.

Moreover, provided that we have N samples (numbers, data, space atoms, or whatever) uniformly distributed in a unit m-ball centered at the origin, the median distance from it to the closest data point considered as the nearest-neighbor estimate is given by $\left(1 - 1/2^{\frac{1}{N}}\right)^{\frac{1}{m}}$ [26]. As m goes to infinity, such a distance tends to the unit; any pair of points is separated by a similar distance and concentrated around the average distance far from the origin, in the proximity of the surface. To fix ideas, take fewer than 56 points in the unit 4-ball of the 4DE space randomly, and none of them will be within the 4-ball with radius $1/3$. Thus, a m-ball turns into a scrunched-up form where the notion of distance is moot. In general, a random point selected from the volume under consideration will probably be closer to the boundary of the sample space than to any other data point, so that extrapolation from neighboring sample points prevails over interpolation between them.

 doi:10.20944/preprints202603.0536.v1

This scenario demonstrates that the curse of dimensionality results from the basic connection between physical space and probability theory. The idea is intuitively explained in [27], step by step, especially why the concentration of density around the average value causes Euclidean distances to increase as the number of dimensions increases. At the end of the day, the number of random variables we are summing converges, so that each additional data point increases the likelihood that the outcome is a true measure of the mean, which is the law of large numbers in statistics.

As dimensions pile up, space stretches thin and grows patchy. High-dimensional realms are mostly empty, with the center shrinking to a speck and most of the territory drifting far away. Beyond the third dimension, this effect intensifies, and by the eighth, balls and spheres are so minuscule that the space becomes nearly intractable. Even with more randomness and symmetry, nature seems to shun these high-dimensional Euclidean landscapes, making rare exceptions for marvels like the Leech lattice in the 24DE space.

3. Natural Efficiency and Information

To ensure physical persistence and stability [28], we must consider adding dimensions while lessening the effects of the curse of dimensionality via GPN and radix economy.

Why would the universe utilize a positional number system in the first place? GPN denotes a number system in which the contribution of a symbol, usually a digit, to the value of a number is the product of the symbol's value by a factor determined by the place of the symbol [29]. The radix is the number of unique symbols that a positional number system uses to represent numbers. The representation of a number is a sequence of digits; we adopt the convention that the leftmost (leading) nonzero digit is the most significant digit (although the opposite convention is equally valid), where significance is a synonym of weight, worth, or mass.

GPN favors arithmetical calculations and approximations because of the exponentiation of the radix. Although a minus sign and a radix point (i.e., decimal point in radix ten) allow us to include negative numbers and fractions to represent every real number up to arbitrary accuracy, our description will be circumscribed to the integer numbers for the sake of simplicity. For instance, in SPN (Standard Positional Notation), given N, d, p, b whole numbers where $b > 1 \wedge b > d$, the term $d \times b^p$ is to be added up to the value of a decoded number N when we find within the sequence of the number N encoded with radix b, written N_b, the digit d in the pth place starting from 0 and leftwards. For instance, $2011_{\,3} = 2 \times 3^3 + 1 \times 3^1 + 1 \times 3^0 = 58_{\,10}$.

We can choose not to utilize the digit zero; bijective numeration is plausibly " more natural" [30]. Āryabhaṭa numeration distinguishes between odd and even powers to favor the expression of large numbers [31]. Signed-digit representation (e.g., balanced ternary) and negative integer radices [32] handle integers homogeneously (averting the use of a minus sign) and normally improve addition and truncation. Gray (or reflected n-ary) code [33,34] minimizes the number of digit switches between two consecutive numbers and facilitates error correction. Rational as well as algebraic number radices [35,36] constitute alternative positional systems with advantages for specific scopes of application (e.g., geometry) at the expense of a greater complexity of arithmetic operations.

We will focus our work on positional number systems that form geometric sequences to balance simplicity of representation with operational efficiency; some systems are positional as well as non-geometric, such as the mixed-radix system, where the number of permitted digit values depends on the position (factorial number system, Fibonacci coding, binary-coded decimal). The order of magnitude of a number written in GPN approximately corresponds to its logarithm. The logarithm is actually counting the number of occurrences of the same factor, the radix, in repeated multiplication (e.g., the googol number in radix 10 is 100 digits long, and $\log_{24} 1327488$ is between 4 and 5 because $24^4 < 1327488 < 24^5$).

One crucial requirement for a universal number system is a reasonable cost. The representational cost of a number N in radix r, denoted $C(N, r)$, estimates the likelihood of finding an integer between 1 and N encoded on a r-ary logarithmic scale, that is, how much such a datum is compactified. Such a

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2026 doi:10.20944/preprints202603.0536.v1

cost in SPN can be defined as $C(N,r) = r\lfloor \log_r N + 1 \rfloor \in \mathbb{N}$, while bijective numeration uses a similar expression. It can also be viewed as Hartley information [37], namely $C(N,r) = \log_r N^r$. Thus, we will take the representational cost of a number written in GPN to be approximately

$$C(N,r) = \frac{r}{2}\log_r(N+1) \tag{1}$$

so that $C(1,2) = 1$. Note that the cost is a product of two functions of size, the one related to the alphabet ($\frac{r}{2}$) and the other to the number of positions ($\log_r(N+1)$).

For all these systems and coding schemes, we obtain the minimum cost when r is approximately equal to Euler's number, e. We must take the radix r to be a natural number greater than 1. Radix 1 (unary) implies that only one symbol or digit must be repeated in sequence as many times as the number being represented. In such a case, radix economy is ill-defined, and we end up with no coding at all. At the other end, we can choose to shorten the number of positions to be occupied, but then a larger set of symbols is necessary to represent a generic number. In the limit, choosing $r = N+1$ amounts to abandoning the positional number system, requiring a symbol per number, at a cost $C(N,N) \propto N$.

Equation 1 indicates that a certain amount of resources will be consumed. In general, a reduced radix requires wider spatial sequences, and vice versa; larger radices produce narrower data because each position provides more bits of information, meaning that we gain more information when the value of a position becomes known. Binary and negabinary transmit 1 bit (unit of information) per position, ternary and negaternary transmit $\log_2 3 \approx 1.585$ bits per position, quaternary and negaquaternary transmit $\log_2 4 = 2$ bits per position, etc. The greater the radix, the more information we can gain when we unveil the value of a position, but every position involves more symbols preencoded. In other words, we cannot transmit a great amount of information by forming small sequences with a small vocabulary.

4. Probability and Volume

Let us assume that objects of any dimension are liable to support information. If there is no preferred dimension and all the components are handled on an equal footing, the information must be shared and distributed on all the dimensions. How does the curse of dimensionality affect the representational cost of a number?

An m-ball of radius N has volume $V_m N^m$, where V_m is the volume of the unit m-ball. The volume of the unit 1-ball (the ball with $N=1$ of the 1DE space) is $V_1 = 2$, the diameter of the interval. Then, we can express the representational cost of the 1DE in terms of volume, namely $C(N) = \left(\frac{V_1}{2}\right)N = N$, so that $C(1) = 1$.

If we use GPN, we can define the " volume factor" of equation 1 as $\log_r(N+1) = \frac{V_1}{2}$, so that

$$C(N,r) = \frac{r}{2}\log_r(N+1) = \frac{r}{4}V_1 \tag{2}$$

and $C(1,2) = 1$.

Now, NBL and a good radix choice can help contain disproportionate cost growth. Because we are using GPN, the probability of a digit k with coding radix r is

$$\Pr(r,k) = \log_r\left(1+\frac{1}{k}\right) \in \mathbb{R} \tag{3}$$

and

$$\Pr(r,[1,k+1)) = \sum_{k=1}^{k}\log_r\left(1+\frac{1}{k}\right) = \log_r(k+1) \tag{4}$$

is the cumulative distribution function.

Let us assume that the stuff of the one-dimensional space is NBL-compliant. As a generalization of the concepts of length, area, and volume, the volume of a $(k+1)$-cut of the m-ball with center the origin would not be $V_m(k) = \frac{\pi^{\frac{m}{2}}}{\Gamma(1+\frac{m}{2})}\left(\frac{k+1}{r}\right)^m$, which corresponds to the uniform distribution, but (using 4)

$$\Upsilon_m(r,k) = \frac{\pi^{\frac{m}{2}}}{\Gamma(1+\frac{m}{2})}(\Pr(r,[1,k+1)))^m = \frac{\pi^{\frac{m}{2}}}{\Gamma(1+\frac{m}{2})}\log_r^m(k+1)$$

Therefore, the representational cost of an m-dimensional datum is, extrapolating equation 2,

$$C(m,r,k) = \frac{r}{4}\Upsilon_m(r,k) = \frac{\pi^{\frac{m}{2}}}{4\Gamma(1+\frac{m}{2})}\frac{r}{\ln^m r}\ln^m(k+1) \tag{5}$$

We obtain the volume of the unit m-ball by setting $k = r-1$. Note that we have associated the origin with the central nucleus represented by the first digit ($k=1$). For example, in dimension $m=1$, $\Upsilon_1(2,1) = 2$, $\Upsilon_1(3,1) = 2\log_3 2$, $\Upsilon_1(3,2) = 2$, $\Upsilon_1(10,1) = 2\log_{10} 2$, $\Upsilon_1(10,5) = 2\log_{10} 6$, and $\Upsilon_1(10,9) = 2$. Then, the volume of the central nucleus with radix 3 occupies $\frac{\Upsilon_1(3,1)}{\Upsilon_1(3,2)} = 0.63\%$ of the ball (in contrast with $\frac{1}{2} = 0.5\%$), the volume of the central nucleus with radix 10 occupies $\frac{\Upsilon_1(10,1)}{\Upsilon_1(10,9)} = 0.301\%$ of the ball (in contrast with $\frac{1}{9} = 0.111\%$), and $\frac{\Upsilon_1(10,5)}{\Upsilon_1(10,9)} = 0.78\%$ (in contrast with $\frac{5}{9} = 0.55\%$). The new percentages exactly match NBL.

We can generalize the NBL 3 by considering that any numeral induces a likelihood gap on the logarithmic scale. A dataset is said to satisfy the local NBL for SPN if the leading r-ary numeral $R \in \mathbb{N}^+$, of any length, occurs with probability

$$\Pr(r,R) = \log_r(R+1) - \log_r R = \log_r\left(1+\frac{1}{R}\right) \in \mathbb{R}$$

This means that k in equation 5 can be generalized to any number, i.e., the representational cost of a m-tuple of data written in GPN is

$$V_m = \frac{\pi^{\frac{m}{2}}}{\Gamma(1+\frac{m}{2})}$$

$$\Upsilon_m(r,R) = V_m \log_r^m(R+1)$$

$$C(m,r,N) = \frac{r}{4}\Upsilon_m(r,R) \tag{6}$$

Regardless of the number being encoded, we reach the minimal representational cost at $r = \mathrm{e}^m$. If our approach is valid, we should find m-dimensional objects that favor this optimal radix choice for any dimension. The next section focuses on SLGs whose representation is close to such a " natural" radix, where the natural frequency distribution of leading digit $d \in \{1,\ldots,k-1\}$ within an ambient space of rank m will occur with probability

$$\Pr_m(d) = \log_k\left(1+\frac{1}{d}\right) \approx \log_{\mathrm{e}^m}\left(1+\frac{1}{d}\right) = \frac{1}{m}\ln\left(1+\frac{1}{d}\right)$$

The standard cost grows exponentially with the radius (R^m), while the NBL cost grows as $\ln^m R$. Therefore, the standard ball model generally runs into the curse of dimensionality less quickly than the NBL model, at the cost of managing incredibly large volumes (see Figure 1). However, the logarithm moves this lognormal curve leftwards by increasing the radix (see Figure 2). Setting the radix and radius to approximately the same value moves the curve peak vertically, but at the same dimension (see 3).

doi:10.20944/preprints202603.0536.v1

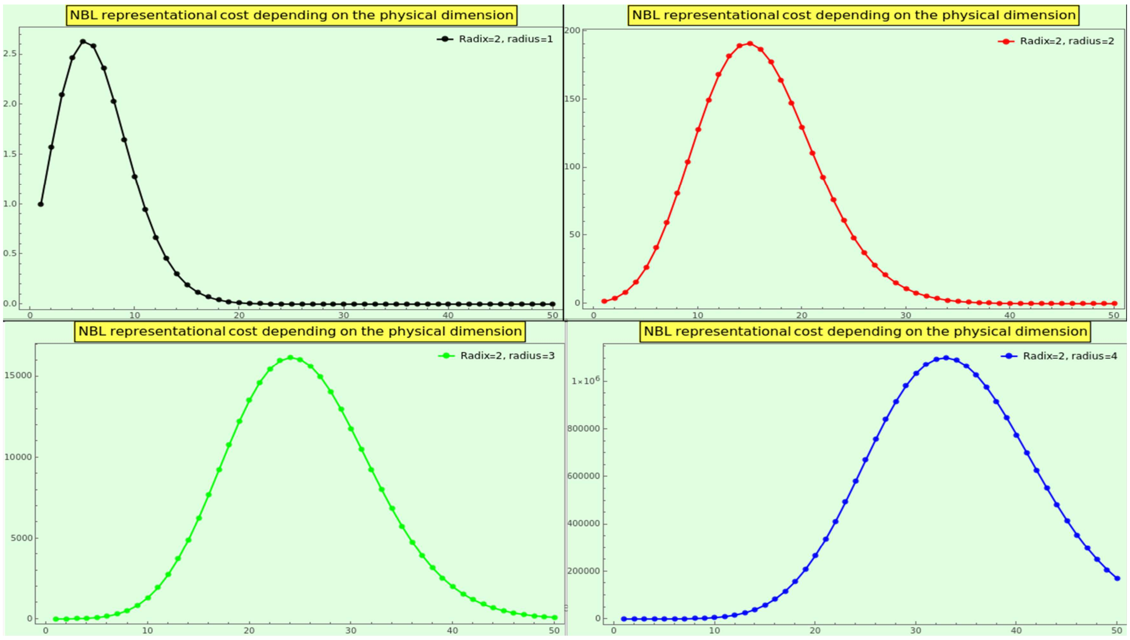

Figure 1. The cost of the standard m-ball grows with respect to the binary logarithmic m-ball as $\sim \left(\frac{R}{\ln R}\right)^m \approx \pi^m(R)$, where $\pi(R)$ is the prime counting function. The logarithm model notices the curse of dimensionality more quickly than the standard ball model. For example, with radix and radius 2, the peak volume is reached at dimensions $m = 15$ and $m = 24$, respectively. However, the respective maximum values are 190.8 and 16186.4. Unlike standard volumes, logarithmic volumes are generally workable.

 doi:10.20944/preprints202603.0536.v1

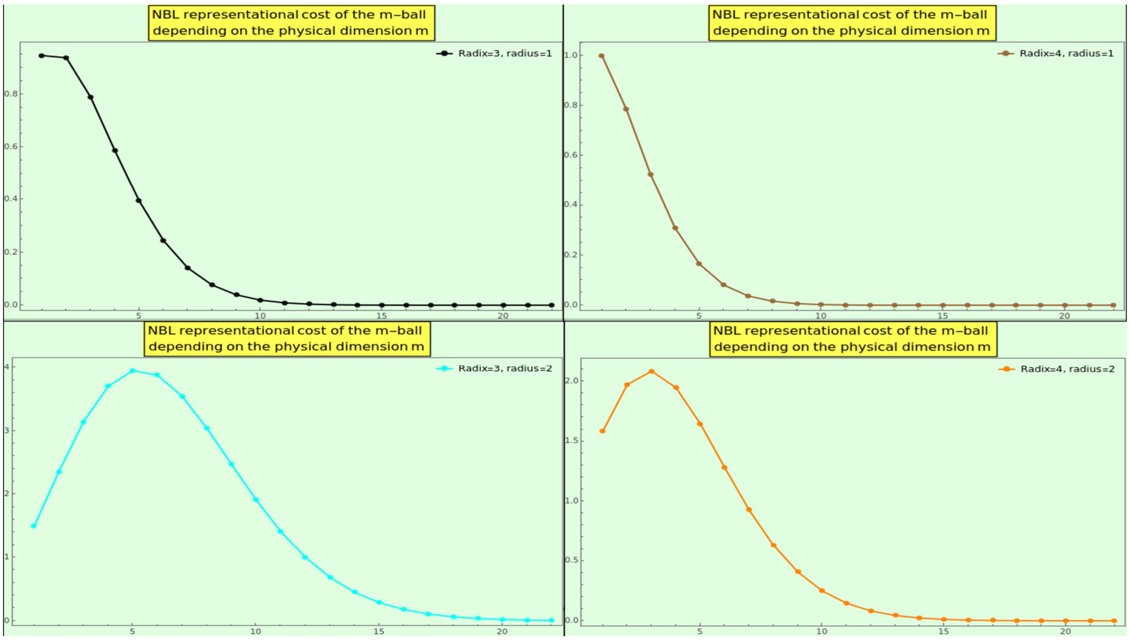

Figure 2. The logarithmic volume shifts the peak of the lognormal curve to the right if we increase the radius or decrease the radix.

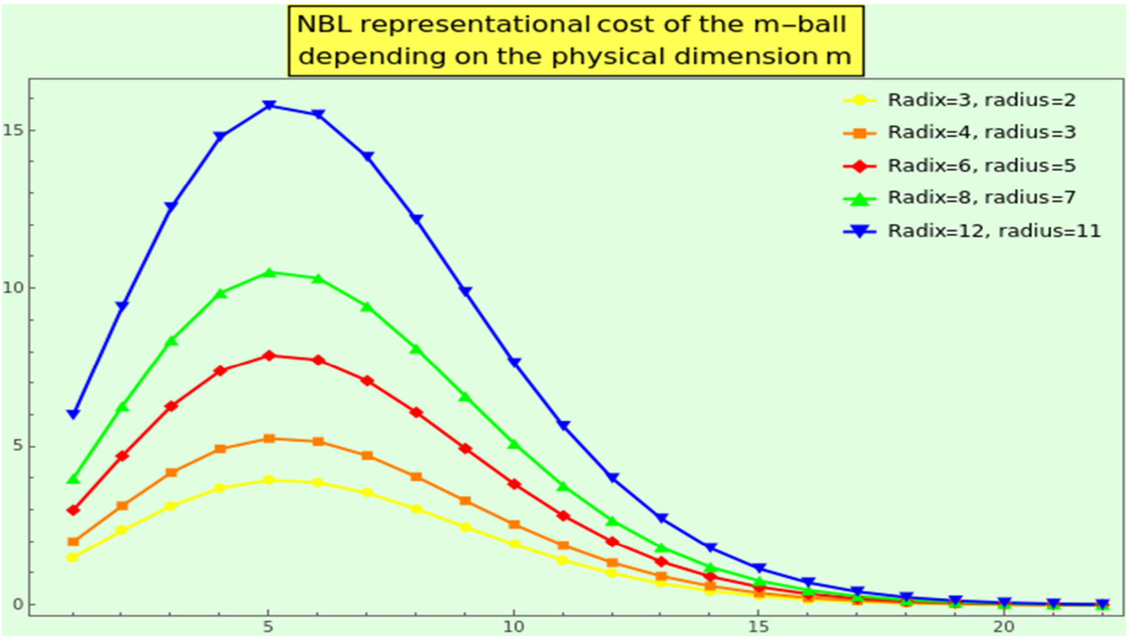

Figure 3. Setting $r = R + 1$ in equation 6 moves the curve peak vertically, but at the same dimension.

The tiniest logarithmic volume, $Y_m(2, 1)$, coincides with the standard volume of the unit ($R = 1$) m-ball.

 doi:10.20944/preprints202603.0536.v1

If the circle of $2R$ was $\frac{V_2(2R)}{V_2(R)} = \left(\frac{2R}{R}\right)^2 = 4$ times as much as the size of the circle R, now

$$\frac{\Upsilon_2(2,2)}{\Upsilon_2(2,1)} = \left(\frac{\ln 3}{\ln 2}\right)^2 \approx 2.51 \tag{7}$$

i.e., 60% of randomly selected points will be at a distance in the range $(1..2]$ from the center. Likewise, 53.4% of randomly selected points will be within a distance $(2..4]$ of the center, 49.2% within $(3..6]$, and so on. The percentage depends on R. Let $\alpha \in \mathbb{R}^{>1}$ be a growth parameter; in the limit,

$$\lim_{R\to\infty} \frac{\Upsilon_m(2,\alpha R)}{\Upsilon_m(2,R)} = \left(\frac{\ln(\alpha R+1)}{\ln(R+1)}\right)^m = 1 \tag{8}$$

That is, this limit tends to 1 regardless of the fixed scale factor α and $m \in \mathbb{N}$, in sharp contrast with the standard $\frac{V_m(\alpha R)}{V_m(R)} = \alpha^m$, which diverges. We have managed to concentrate the mass in the proximity of an m-ball's center.

A volume of stuff is now homogeneous with a bias towards the center of an m-ball, which is the reference framework of a positional coding system. The Euclidean distance from such a center to the corners of the m-cube with edges of length $2R$ will be $R\sqrt{m}$ anyway, but at least the inscribed m-ball is tractable.

5. Radix Economy of Simple Lie Groups

As dimensions rise, data scatter swiftly. Distance loses its meaning, and the landscape is ruled by thinness and sameness. In the language of communication, data points blur amid the static, nearly indistinguishable from noise. This predicament demands a creative remedy.

Yet, in practice, the curse of dimensionality often goes unnoticed. Instead, the Newcomb-Benford Law emerges everywhere. As Benford observed, humans count step by step, but nature leaps geometrically, following the rhythm of $\mathrm{e}^0, \mathrm{e}^1, \mathrm{e}^2, \mathrm{e}^3$, and beyond. This work uncovers an economic logic behind GPN and envisions universal structures where dimension e^m serves as a holistic radix, shaping evolution's grand design.

We can factually export the concept of radix economy and optimal radix choice to any algebraic group, in particular to SLGs, as atoms of finite groups and candidates to characterize the elementary building blocks of the universe. The elements of an SLG can be represented in a matrix. We are not interested in the matrix entries [38], but rather in the dimensions of the irreducible matrices in agreement with the cardinality or amplitude of their potential alphabets. For arbitrary finite-dimensional representations, a fundamental representation is an irreducible faithful collection of states determined by the weight diagram constructed from a choice of simple roots and the mark of the highest weight [39].

In Table 1, we classify a set of 24 compact SLGs regarding their relative group economy, i.e., their algebraical dimension s relative to that of e^m, where m is the algebraic rank. From the irreps of an SLG, we have chosen the defining (adjoint) representation over $\mathbb{R}$ for the classical root systems, and for the exceptional Lie groups, we have selected the fundamental representation with the most informative order, i.e., with the optimal group economy for data representation. The last column indicates the ranking $[1..24]$ in the set of SLGs; the closer $\frac{s}{\mathrm{e}^m}$ to 1, the higher in the hierarchy. For example, the fundamental representations of E7 have dimensions 133, 8645, 365750, 27664, 1539, 56, and 912, the latter being the most informative, and that we have taken into consideration.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2026 doi:10.20944/preprints202603.0536.v1

Table 1. Economy of the 24 most significant SLGs.

SLG	Algebraic Rank (m)	algebraical dimension (s)	Relative Group Economy	Hierarchy Position
$\mathrm{SU}(2), \mathrm{Sp}(1)$	2	3	0.406	14
$\mathrm{SU}(3)$	3	8	0.398	15
$\mathrm{SU}(5)$	5	24	0.162	17
$\mathrm{SU}(9)$	9	80	0.010	23
$\mathrm{Spin}(3)$	3	3	0.149	18
$\mathrm{SO}(4,1)$	5	10	0.067	20
$\mathrm{SO}(8)$	8	28	0.009	24
A_1	1	3, (2), (adjoint)	1.104	5
A_2	2	8, $(1,1)$, (adjoint)	1.083	4
A_3	3	15	0.747	8
A_4	4	24	0.440	13
A_8	8	80	0.027	22
B_2	2	10	1.353	11
B_3	3	21, $(0,1,0)$, (fundamental, adjoint, faithful for SO(7), but not faithful for Spin(7))	1.046	1
C_3	3	21, $(2,0,0)$, (adjoint)	1.046	1
B_4, C_4	4	36	0.659	10
B_6, C_6	6	78	0.193	16
B_8, C_8	8	136	0.046	21
D_4	4	28	0.513	12
D_7	7	91	0.083	19
G_2	2	7, $(1,0)$, (fundamental, faithful)	0.947	3
F_4	4	52, $(0,0,1,0)$, (fundamental, adjoint, faithful)	0.952	2
E_6	6	$\frac{702}{2}$, $(0,0,1,0,0,0) \oplus (0,0,0,1,0,0)$, (fundamental)	0.870	6
E_7	7	912	0.832	7
E_8	8	3875	1.300	9

In general, the exceptional Lie groups are the groups that conform the best to their algebraic rank, that is, the dimension of the underlying Euclidean space. In absolute terms, the SLG with the best economy is the B3 ($\simeq$C3) root system. Funnily enough, F4, second in the rank, is the only one in which the most economic fundamental representation tallies with the adjoint representation, obtained by linearizing, i.e., taking the differential of, the action on itself by conjugation. The 7-dimensional irreducible representation of G2 follows in the ranking. Note that this podium of economy consists of faithful irreps. Remember that an irrep is faithful if the kernel of the representation is trivial, i.e., maps only the identity to the identity, and a faithful irrep captures the full structure of the group without " wasting" any distinct elements. A2, A1, and E6 are the next thriftiest SLGs. Note also that E7, A3, and E8 complete the series but provide significantly worse efficiency. In particular, from the perspective of the economy, E8 is the least " beautiful" among the exceptional Lie groups [40,41].

 doi:10.20944/preprints202603.0536.v1

The 6 most economically active groups induce an inclusion chain, namely $\mathsf{E}_6 \supset \mathsf{F}_4 \supset \mathsf{B}_3 \supset \mathsf{G}_2 \supset \mathsf{A}_2 \supset \mathsf{A}_1$. Thus, F4 embeds in E6, B3 embeds in F4, and so on. All are compact. Some notes about the members of the Lie Squad follow.

- B3. Its root diagram (and the root diagram of the dual C3) outlines a star whose Voronoi cell is the space-filling (3 dimensions) rhombic dodecahedron (18 dimensions) [39]. The simply-connected covering group associated with the algebra is $\mathsf{Spin}(7)$, a compact group that is the double cover of the seven-dimensional special orthogonal (rotation) centerless compact group, $SO(7)$, which provides the symplectic structure of Hamiltonian systems. $\mathsf{Spin}(7)$ is key to fully understanding general relativity and quantum gravity, given that Spin(7)-manifolds are Ricci-flat. Specifically, gluons transform in the 21 representation of the color symmetry group.
- F4. It is simply-connected and compact. The atomic constituent of this four-dimensional body-centered cubic lattice is the peerless self-dual 24-cell. It is the union of two hypercubic lattices, each lying in the center of the other, spanned by the ring of Hurwitz quaternions. The F4 lattice supports the 3 only regular space-filling tessellation (or honeycomb) of the four-dimensional Euclidean space, giving rise to the densest known regular sphere packing in rank four, with the kissing number 24 [42].
- G2. It is simply-connected and compact. It is the smallest of the five exceptional SLGs. The root diagram outlines a hexagram (i.e., a Star of David) with 6 dimensions plus the origin. Essentially related to the octonions, it appears in string and Yang-Mills theories of supersymmetric atomic physics and gravity [43] and as a vehicle of dual superconductivity processes [44]. In M-theory, G2-manifold compactifications yield four-dimensional theories with minimal ($N = 1$) supersymmetry. Notoriously, G2 embeds the subgroups of the standard model of physics, to wit, $U(1)$, $SU(2)$, and $SU(3)$, constituting the simplest gauge group that allows confinement without a center [45].
- A2. Its associated simply-connected and compact group is the special unitary group, $SU(3)$, i.e., the Lie group of 3×3 unitary matrices with determinant 1, which drives the strong force [46], i.e., the quark model. The root diagram depicts a hexagon (6 dimensions) in the plane (2 dimensions), yielding the hexagonal close-packed lattice. The group is also isomorphic to the union of two quaternions. Computationally, we can assign A2 to the 8-binary representations of a byte; for instance, the hypothesis of a logarithmic distribution of mantissa errors points to a binary-power radix of 2^3 if optimal storage use is additionally required [47].
- A1. It is the simply-connected and compact special unitary group, $SU(2)$, that is, the Lie group of 2×2 unitary matrices with determinant 1, the one driving the weak force. The group is also isomorphic to the unitary quaternions $\mathsf{U}(\mathbb{H})$, and hence to the 3-sphere [48]. The root diagram outlines the sequence $\{-1, 0, 1\}$, the main root plus its opposite (2 dimensions) on a line (1 dimension), leading to the balanced ternary numeral system [49]. The lattice version is $\mathbb{Z}$, trivially defined in the 1DE space R ($\cong \mathbb{R}$). In physics, it is associated with the angular momentum and spin, or with a three-dimensional rotation.
- E6. Both the complex and the compact real form of this exceptional Lie group are simply-connected. Geometrically, it can result in the 2_{22} honeycomb, a uniform tessellation of the six-dimensional Euclidean space, a lattice constructed from 2_{21} polytopes [50]. Closely related to quaternions and octonions, E6 conveys the notion of triality [51], which establishes a symmetric super-geometry via a volumetric structure between three vector spaces, just as duality associates complementary vector spaces. This root system appears in particle physics as a candidate gauge group in grand unification theories and in supergravity.

In general, for the 6 most efficient Lie groups, the leading digit 1 will approximately occur with probability $\Pr_m(1) = \frac{\ln 2}{m}$, meaning that the probability of the first digit is in any dimension given by the fraction of $\ln 2$ corresponding to the rank. A1, A2, G2, B3, F4, and E6 encode the unit with probabilities $\Pr_{\mathsf{A}_1}(1) = \log_3 2 = 63.09\%$, $\Pr_{\mathsf{G}_2}(1) = \log_7 2 = 35.62\%$, $\Pr_{\mathsf{A}_2}(1) = \log_8 2 = 33.33\%$, $\Pr_{\mathsf{B}_3}(1) = \log_{21} 2 = 22.77\%$, $\Pr_{\mathsf{F}_4}(1) = \log_{52} 2 = 17.54\%$, and $\Pr_{\mathsf{E}_6}(1) = \log_{351} 2 = 11.83\%$, respectively.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2026 doi:10.20944/preprints202603.0536.v1

The nearness of this Lie Squad to an optimal coding can be the cause of the universal compliance with NBL, by which, to put it simply, the universe counts geometrically. Moreover, this outlook can serve as a predictive tool.

In Table 2, we include the Weyl order and corresponding divisors. The Weyl group, generated by simple reflections on the root system, is finite and acts faithfully (each element has a distinct effect on the set) to permute the set of roots, preserving inner products and the root (and weight) lattices, and divides the corresponding rank-dimensional space into specific regions termed Weyl chambers. The order of the Weyl group indicates the size of the solvable symmetry group of the corresponding polytope. For instance, the Weyl group for B3 (and C3) consists of signed permutations of 3 coordinates, and the Weyl order of F4 is 1152, double the size of the symmetry group of the 24-cell. The table also includes the corresponding polytopes that embody the abstract properties of the SLG. A " polytope" in the context of SLG's irreps refers to the convex hull of the weights of that particular representation, where the " weights" are vectors in the weight lattice, a geometrical space associated with the Lie group's structure. For a certain group of representations, these weights form a pattern, and the convex hull of these weights yields a geometric shape, the polytope.

Table 2. Weyl characterization of the members of the Lie Squad and their polytopes.

SLG and irrep label	Dimension	Polytope description	Weyl order	Weyl divisors
B_3 $(0,1,0)$	21	Rhombic dodecahedron's dual	$2^3 3! = 48$	$d(48) = 10$
F_4 $(0,0,1,0)$	52	Rectified 24-cell	1152	$d(1152) = 24$
G_2 $(1,0)$	7	Hexagram	12	$d(12) = 6$
A_2 $(1,1)$	8	Regular hexagon	$3! = 6$	$d(6) = 4$
A_1 (2)	3	Line segment	$2! = 2$	$d(2) = 2$
E_6 $(0,0,1,0,0,0) \oplus (0,0,0,1,0,0)$	351	Specific polytope	51840	$d(51840) = 80$

6. Fundamental Positional Notation

If the geometry of an SLG encodes information, its number system must be fundamental. Let us assume that 'fundamental' means 'lack of ambiguity'. For instance, SPN is not fundamental because we need an ad hoc symbol to indicate whether a number is negative or positive.

In principle, we have two candidates, namely balanced positional notation and bijective numeration.

The former is a non-standard, unambiguous number system; unlike standard positional notation, every integer has a unique representation, assuming leading zeros are ignored. It requires an odd radix; it uses a set of digits centered around zero, including both positive and negative values. This eliminates the need for a separate minus sign to represent negative numbers. For a given radix r (e.g., 3, 7, 21), the digits range from $\frac{(1-r)}{2}$ to $\frac{(r-1)}{2}$. We represent negative digits with an underbar, such as $\underline{3}$ for -3. The sign of a number is determined by the sign of its leftmost non-zero digit; the sign is inherent in the number's representation. The negative of a number is obtained by flipping the signs of all its digits. It has the " truncation property" , meaning that rounding a fractional number to the nearest integer is exactly the same operation as simple truncation (discarding the fractional part), which minimizes cumulative errors in calculations. Moreover, balanced positional notation simplifies some arithmetic operations because carries are shorter or eliminated.

To convert from balanced positional notation to SPN, we multiply each digit by its corresponding power of the radix and sum the results. To convert a SPN number to balanced positional notation, we repeatedly divide the number by the radix, ensuring the remainder falls within the balanced digit range. If it is out of the range, then add 1 to the quotient, and the remainder minus the radix is the new remainder. Reading the remainders from bottom to top gives the balanced positional notation.

The archetype of balanced positional notation is balanced ternary notation. We can use the digits $\{\underline{1},0,\overline{1}\}$ to represent the trits of a balanced ternary system with values $\{-1,0,1\}$. For example,

 doi:10.20944/preprints202603.0536.v1

$17_{10} \equiv \overline{1}\underline{1}0\underline{1}$. For one thing, this system is optimal from an informational perspective, since 3 is the integer closest to Euler's number. Additionally, the merits of this uniform self-contained number system are multiple. Balanced ternary is " extremely easy," said the pioneer Thomas Fowler [49], " The Jewel in the Triple Crown" [52], and " Perhaps the prettiest number system of all" [36].

Furthermore, for the following reasons, it is the most efficient representation system for elements of $\mathbb{Z}$, just as binary is the simplest GPN for elements of $\mathbb{N}$. The leading nonzero trit indicates the sign of a number (e.g., $-8_{10} \equiv \underline{1}01$), and hence arithmetic is realized in direct code. Balanced ternary can operate on signed numbers of variable length, avoiding the typical overflows that radix complement entails. Comparing any two numbers is trivially accomplished by parsing their sequences from left to right and applying the rule $\underline{1} \prec 0 \prec \overline{1}$. Error detection is a self-checking process, since the parity of a balanced ternary integer matches the parity of the sum of all its trits (for example, the parity of 17 is the parity of $1 - 1 - 1$, that is, odd). Addition is realized as usual, but with symmetric ± 1 carry. Subtraction is the addition of the negated subtrahend. A one-digit multiplication table has no carries. General multiplication only implies negation and (shifted) addition. Divisibility by 2 (parity) requires checking the parity of the sum of all trits, and divisibility by 3 is equivalent to a 0 in the unit position. Transactions (of mass, money, goods, etc.) are minimized with respect to other numeral systems.

It turns out that the balanced ternary is naturally associated with the A1 SLG. Interestingly, A1 is fundamental because every rational number is uniquely the difference between two sums of distinct and non-overlapping powers of 3 (just as every natural number is uniquely a sum of distinct powers of 2), as an ancient theorem that Euler stated in terms of generating functions [53]

$$\overset{[}{n}= 0]\infty\prod\left(q^{-3^n}+1+q^{3^n}\right) = \overset{[}{n}= -\infty]\infty\sum q^n.$$

For example, $17_{10} \equiv \left\{\underline{1}0\underline{1}, \overline{1}000\right\} = \{-10, 27\}$ and $-8_{10} \equiv \left\{\underline{1}00, \overline{1}\right\} = \{-9, 1\}$.

A serious inconvenience of taking balanced positional notation as a fundamental representation of the SLGs is that it is only defined over odd radices, since it uses a symmetric set of digits centered around zero.

An alternative to balanced positional notation is bijective numeration, a GPN that uses a string of r digits from the set $\{1, 2, \cdots, k, \cdots, r\}$ (where $1 \le k \le r$) to encode natural numbers in a unique way. The merits of bijective numeration are multiple, too. Because it is a zero-free representation, it is very compact, especially for smaller radices. An ordered list of bijective numbers is automatically in " shortlex" order (the shortest number appears first, and by digit value as a tie-breaker). Unlike SPN, unary bijective numeration is not a special case; the no-code case is a sound continuation of a radix-reduction process. The greatest inconvenience of this notation is that we cannot extend it to negative integers without involving ambiguity [54] or adding a sign symbol.

We can combine balanced ternary with bijective numeration in what we have called Symmetric Bijective Notation. Since balanced ternary allows us to write every integer as a unique difference between two positive integers, we can bijectively represent an integer as the difference between a pair of numbers encoded as bijective numeration. Table 3 below contains several examples of SBN adopted by the Lie Squad.

Table 3. Examples of how to represent integers employing the Lie Squad SBN systems.

Number	A1	A2	G2	B3	F4	E6
−9	−23	−11	$-ab$	$-l$	$-i$	−|9
45	2223 − 323	121 − 44	$add - ea$	$dw - bs$	$aC - J$	|81 − |36
−945		33 − 1714	$cf - beef$	$bh - cfh$	$A - rJ$	|27 − |2|270
4321			$beacb - fdag$	$rxt - gcz$	$bvK - QF$	|18|271 − |6|162
−23456				$hwm - cwbl$	$afr - iOv$	|8|226 − |75|165
345678				$cskqh - ybtl$	$cNFJ - apOb$	|4|110|270 − |1|178|327
−123456789					$SzDB - qMAIC$	|51|330|171 − |3|1|6|81

The (2) irrep (adjoint) of A1 is a real representation that can be realized using 3×3 matrices acting on a three-dimensional real vector space, hence employing ternary notation with digits $\{1, 2, 3\}$ corresponding to their numeric value.

The $(1,1)$ irrep (adjoint) of A2 is a real representation that can be realized using matrices of order 8 acting on an 8-dimensional real vector space, hence employing octonary notation with digits $\{1,2,\cdots,8\}$ corresponding to their numeric value.

G2 is a 14-dimensional Lie group, meaning a generic element requires 14 independent real parameters to specify it. The matrices are of order 7 and act on a 7-dimensional complex vector space whose elements are 7-component complex vectors. However, the $(1,0)$ irrep of G2 is a real representation that can be realized using matrices whose entries are real numbers. These real matrices act on a 7-dimensional real vector space. Moreover, when we diagonalize an element of $(1,0)$, we reveal 7 values that are genuinely crucial " axes" of information about that particular transformation. We use the first seven letters of the English alphabet as digits for this irrep.

The $(0,1,0)$ (fundamental, adjoint, faithful for SO(7), but not faithful for Spin(7)) irrep of B3 describes a 21-dimensional real vector space employing unvigesimal notation, with digits corresponding to consonant letters of the English alphabet.

The $(0,0,1,0)$ (fundamental, adjoint, faithful) irrep of F4 describes a 52-dimensional real vector space consisting of the 48 root vectors plus the rank 4. We use the 26 letters of the English alphabet as digits for this irrep, both lower- and upper-case.

We have considered a couple of conjugate fundamental 351-dimensional irreps of E6 as the most informative, namely those with labels (highest weights) $(0,0,1,0,0,0)$ and $(0,0,0,1,0,0)$. (Alternatively, we can say the representation of $(0,0,1,0,0,0) \oplus (0,0,0,1,0,0)$ has dimension $\frac{702}{2}$.) Any 351-dimensional vector space must have a basis consisting of 351 linearly independent vectors. Choosing these vectors to be " unit vectors" simply means we have selected an orthonormal basis, the common choice in linear algebra. These basis vectors, on the one hand, span the entire 351-dimensional space, and on the other hand, provide us with an alphabet of 351 elements to express any integer. For every element in the E6 group, the representation assigns a specific 351×351 matrix that acts on the vectors in that 351-dimensional space. Finally, we can express every component of a 351-dimensional vector, in turn, as a 351-radix number. We use the naturals between 1 and 351 as digits of this irrep, preceded by the character " |" .

7. The Role of the Weyl Group

The Weyl group describes the symmetries of a root system, which is an intrinsic property of the Lie algebra and, by extension, of the Lie group itself. Specifically, the divisors of the Weyl order are associated with the adjoint representation and play a role in the structure of the Lie group.

The Weyl group acts simply transitively on the set of Weyl chambers in the root space. Since the total number of chambers divides the space into equal regions, any Weyl subgroup corresponds to a larger, composite region formed by merging specific chambers. If, for example, the Weyl order is divisible by a high power of 2, the Weyl group contains large 2-groups, corresponding to reflection subgroups generated by subsets of simple roots. In particular, for A1 and A2, the Weyl groups are the symmetric groups S_{n+1}, with order $(n+1)!$, so the divisors of this factorial number, 2 and 6, respectively, dictate the possible subgroup structures within the Weyl symmetry.

More interestingly, for a simple Lie algebra of rank m, the ring of polynomial functions on the Cartan subalgebra that are invariant under the Weyl group is generated by m algebraically independent polynomials. The degrees of these fundamental invariant polynomials, $(d_1, d_2, \cdots, d_n)$, are intrinsic structural constants of the group. The product of these degrees equals the order of the Weyl group, and hence, these degrees are specific divisors of the Weyl order. For A1, A2, G2, B3, F4, and E6, for instance, the degrees of invariants are respectively 2, 2 and 3 (with product 6), 2 and 6 (with product 12), 2, 4, and 6 (with product 48), 2, 6, 8, and 12 (with product 1152), and 2, 5, 6, 8, 9, and 12 (with product 51840).

Closely related to these degrees are the group exponents, denoted $(m_1, m_2, \cdots, m_n)$, where $m_i = d_i - 1$. These determine the topology of the Lie group via its cohomology ring, and the sum of the exponents equals the number of positive roots, which relates to the dimension of the group.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2026 doi:10.20944/preprints202603.0536.v1

Furthermore, we have found that the divisors of a SLG allow us to derive the logarithmic scale provided by NBL. For that purpose, we use the average sum-of-divisors function on a sequence obtained by calculating the tensor multiplication of the Weyl divisors with themselves.

The divisors of the Weyl order of the A1' s adjoint irrep (that with label (2)) are 1 and 2, i.e., $d(|W(\mathsf{A}_1)|) = d(2) = 2$ is the cardinality of the vector space of divisors đ of $|W(\mathsf{A}_1)|$, namely $\text{đ}(\mathsf{A}_1) = \{1,2\}$. By calculating the tensor product $\text{đ}(\mathsf{A}_1) \otimes \text{đ}(\mathsf{A}_1)$, flattening, and sorting, we obtain a sequence of 3 distinct divisors, $\text{đ}_{\mathsf{A}_1^2} = (1,2,4)$, which is precisely the algebraical dimension of the adjoint irrep of A1. Now, for each element $n \in \text{đ}_{\mathsf{A}_1^2}$, calculate its average partial sum, i.e.,

$$\frac{[k = 1]n \sum \text{đ}_{\mathsf{A}_1^2}(k)}{n}$$

Thus, we obtain the sequence $\left(\frac{[k=1]1\sum \text{đ}_{\mathsf{A}_1^2}(1)}{1}, \frac{[k=1]2\sum \text{đ}_{\mathsf{A}_1^2}(2)}{2}, \frac{[k=1]3\sum \text{đ}_{\mathsf{A}_1^2}(3)}{3}\right) = \left(1, \frac{1+2}{2}, \frac{1+2+4}{3}\right) = \left(1, \frac{3}{2}, \frac{7}{3}\right)$. When normalized (and reversed), this distribution approaches NBL in radix 3 with RRMSE 2.33% (see Figure 4 on the left).

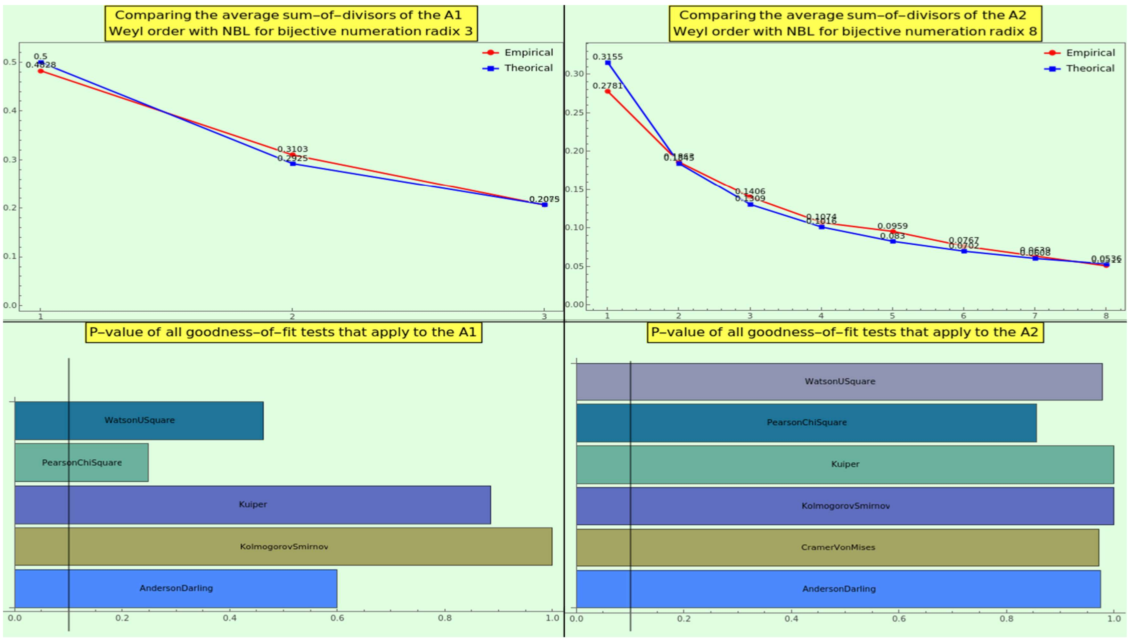

Figure 4. The sequences generated from the divisors of A1 and A2 Weyl orders are NBL-compliant. At the top-left, we display the probability mass functions of the theoretical (NBL for bijective numeration with radix 3) and experimental (average sum-of-divisors of the A1 Weyl order) data. At the top-right, we display the probability mass functions of the theoretical (NBL for bijective numeration with radix 8) and experimental (average sum-of-divisors of the A2 Weyl order) data. At the bottom, we show the corresponding goodness-of-fit p-values; we cannot reject the hypothesis that the distributions generated by the divisors are NBL-compliant.

The divisors of the Weyl order of the A2' s adjoint irrep (that with label $(1,1)$) are 1, 2, 3, and 6, i.e., $d(|W(\mathsf{A}_2)|) = d(6) = 4$ is the cardinality of the vector space of divisors đ of $|W(\mathsf{A}_2)|$, namely $\text{đ}(\mathsf{A}_2) = \{1,2,3,6\}$. By calculating the $\text{đ}(\mathsf{A}_2) \otimes \text{đ}(\mathsf{A}_2)$ tensor product, flattening, and sorting, we obtain the sequence $(1,2,2,3,3,4,6,6,6,6,9,12,12,18,18,36)$ of 16 divisors. By taking alternate divisors, this sequence produces two sequences of 8 divisors, which is precisely the algebraical dimension of the

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2026 doi:10.20944/preprints202603.0536.v1

adjoint irrep of A2. Calculate the average partial sums, $\frac{\sum_{k=1}^{n} \text{đ}_{A_2^2}(k)}{n}$, for both sequences. In particular, the average partial sums of

$$\text{đ}_{A_2^2} = (2, 3, 4, 6, 6, 12, 18, 36)$$

produce the sequence $\left(2, \frac{5}{2}, 3, \frac{15}{4}, \frac{21}{5}, \frac{11}{2}, \frac{51}{7}, \frac{87}{8}\right)$. When normalized (and reversed), this distribution approaches NBL in radix 8 with RRMSE 3.5% (see Figure 4 on the right).

The divisors of the Weyl order of the G2' s adjoint irrep are $\text{đ}(G_2) = \{1, 2, 3, 4, 6, 12\}$, and $d(|W(G_2)|) = d(12) = 6$. By calculating the $\text{đ}(G_2) \otimes \text{đ}(G_2)$ tensor product, flattening, and sorting, we obtain the sequence of 6^2 divisors

(1,2,2,3,3,4,4,4,6,6,6,6,8,8,9,12,12,12,12,12,12,16,18,18,24,24,24,24,36,36,36,48,48,72,72,144)

By taking one out of every $\left\lfloor \frac{6^2}{7} \right\rfloor = 5$ divisors (where $\lfloor\rfloor$ is the floor function), this sequence produces five sequences (those starting from the second to the sixth element) of seven divisors, which is the algebraical dimension of the most informative irrep of G2, namely the fundamental and faithful irrep with label $(1, 0)$ of G2. Calculate the average partial sums, $\frac{\sum_{k=1}^{n} \text{đ}_{G_2^2}(k)}{n}$, for the five sequences of seven divisors. In particular, the average partial sums of the sequence starting at the fourth element

$$\text{đ}_{G_2^2} = (3, 6, 8, 12, 18, 36, 72)$$

produce the sequence $\left(3, \frac{9}{2}, \frac{17}{3}, \frac{29}{4}, \frac{47}{5}, \frac{83}{6}, \frac{155}{7}\right)$. When normalized (and reversed), this distribution approaches NBL in radix 7 with RRMSE 2.18% (see Figure 5 on the left).

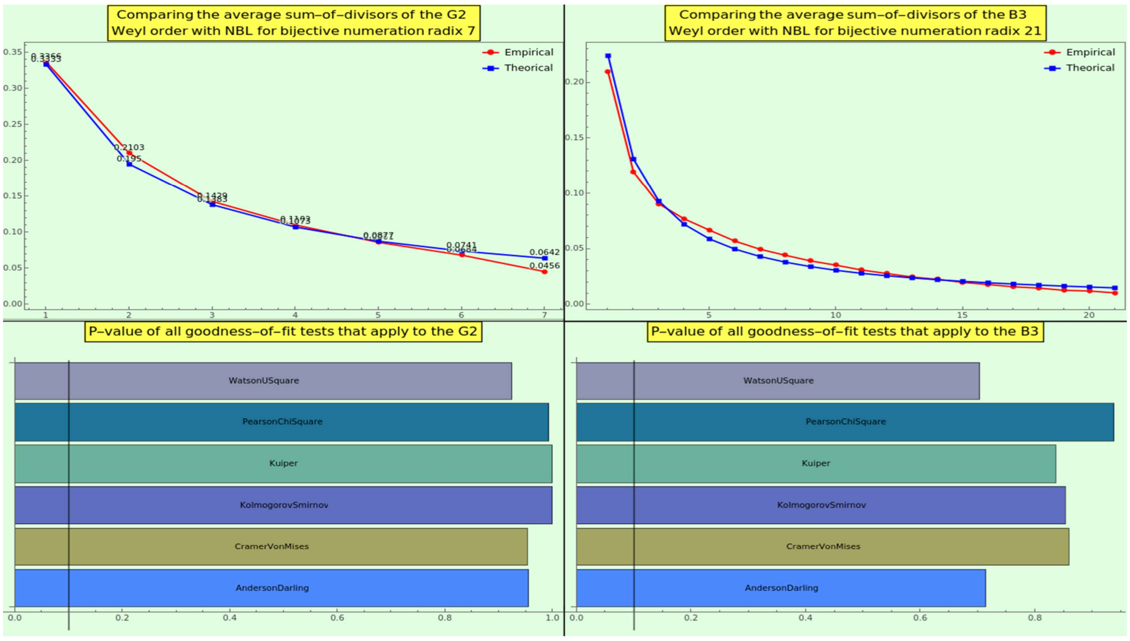

Figure 5. The sequences generated from the divisors of G2 and B3 Weyl orders are NBL-compliant. At the top-left, we display the probability mass functions of the theoretical (NBL for bijective numeration with radix 7) and experimental (average sum-of-divisors of the G2 Weyl order) data. At the top-right, we display the probability mass functions of the theoretical (NBL for bijective numeration with radix 21) and experimental (average sum-of-divisors of the B3 Weyl order) data. At the bottom, we show the corresponding goodness-of-fit p-values; we cannot reject the hypothesis that the distributions generated by the divisors are NBL-compliant.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 6 March 2026 doi:10.20944/preprints202603.0536.v1

The divisors of the Weyl order of the B3' s fundamental adjoint irrep (with label $(0, 1, 0)$) are $đ(\mathsf{B}_3) = \{1, 2, 3, 4, 6, 8, 12, 16, 24, 48\}$ This set has cardinality 10. By calculating the $đ(\mathsf{B}_3) \otimes đ(\mathsf{B}_3)$ tensor product, flattening, and sorting, we obtain the sequence of 10^2 divisors

(1 2 2 3 3 4 4 4 6 6 6 6 8 8 8 8 9 12 12 12 12 12 12 16 16 16 16 16 18 18 24 24 24 24 24 24 24 24 32 32 32 32 36 36 36 48 48 48 48 48 48 48 48 48 64 64 64 72 72 72 72 96 96 96 96 96 96 96 96 128 128 144 144 144 144 144 192 192 192 192 192 192 256 288 288 288 288 384 384 384 384 576 576 576 768 768 1152 1152 2304)

By taking one out of every $\left\lfloor \frac{10^2}{21} \right\rfloor = 4$ divisors, this sequence produces four sequences (those starting from the 17th to the 20th element) of 21 divisors, which is the algebraical dimension of the most informative irrep of B3, precisely the irrep with label $(0, 1, 0)$. Calculate the average partial sums, $\frac{\sum_{k=1}^{n} đ_{\mathsf{B}_3^2}(k)}{n}$, for the four sequences of 21 divisors. In particular, the average partial sums of the sequence starting at the 20th element,

$$đ_{\mathsf{B}_3^2} = (12, 16, 16, 24, 24, 32, 36, 48, 48, 64, 72, 96, 96, 128, 144, 192, 256, 288, 384, 768, 2304)$$

produce a distribution that, when normalized (and reversed), approaches NBL in radix 21 with RRMSE 1.85% (see Figure 5 on the right).

The divisors of the Weyl order of the F4' s fundamental, adjoint, and faithful irrep (with label $(0, 0, 1, 0)$) are

$$đ(\mathsf{F}_4) = \{1, 2, 3, 4, 6, 8, 9, 12, 16, 18, 24, 32, 36, 48, 64, 72, 96, 128, 144, 192, 288, 384, 576, 1152\}$$

with cardinality $d(|W(\mathsf{F}_4)|) = d(1152) = 24$. By calculating the $đ(\mathsf{F}_4) \otimes đ(\mathsf{F}_4)$ tensor product, flattening, and sorting, we obtain a sequence of 24^2 divisors. By taking one out of every $\left\lfloor \frac{24^2}{52} \right\rfloor = 11$ divisors, this sequence produces eleven sequences (those starting from the 5th to the 15th element) of 52 divisors, which is the algebraical dimension of the most informative irrep of F4, precisely the irrep with label $(0, 0, 1, 0)$. Calculate the average partial sums, $\frac{\sum_{k=1}^{n} đ_{\mathsf{F}_4^2}(k)}{n}$, for the eleven sequences of 52 divisors. In particular, the average partial sums of the sequence starting at the 7th element

$đ_{\mathsf{F}_4^2} =$ (8 16 24 32 36 48 64 72 96 128 144 144 192 216 288 288 384 384 432 576 576 768 768 864 1152 1152 1296 1536 1728 2048 2304 2304 3072 3456 4096 4608 5184 6144 6912 9216 9216 12288 13824 18432 20736 27648 36864 49152 73728 110592 221184 1327104)

produce a distribution that, when normalized (and reversed), approaches NBL in radix 52 with RRMSE 3.44% (see Figure 6 at the top-left), indicating very good agreement. However, the null hypothesis that the datasets have the same distribution is rejected at the 10% level by all goodness-of-fit tests, except for Pearson's chi-squared test (p-value = 0.6245).

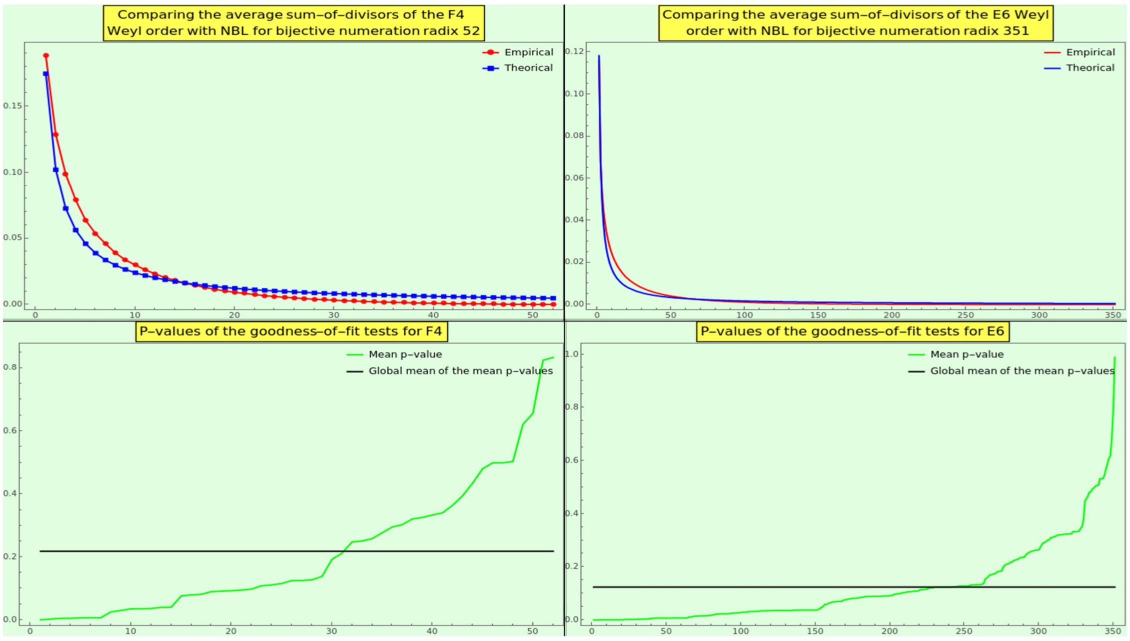

Figure 6. The sequences generated from the divisors of F4 and E6 Weyl orders are NBL-compliant. At the top-left, we display the probability mass functions of the theoretical (NBL for bijective numeration with radix 52) and experimental (average sum-of-divisors of the F4 Weyl order) data. At the top-right, we display the probability mass functions of the theoretical (NBL for bijective numeration with radix 351) and experimental (average sum-of-divisors of the E6 Weyl order) data. At the bottom, we list, in order, the corresponding goodness-of-fit p-values. These are obtained as the mean of the Anderson-Darling, Kolmogorov-Smirnov, Kuiper, Cramér-von Mises, and Watson U^2 goodness-of-fit tests for a series of random elements of length 9, calculated 52 times for F4 and 351 times for E6. Based on the results, we cannot reject the hypothesis that the distributions generated by the divisors are NBL-compliant.

To validate the hypothesis that this sequence generated from the divisors is NBL-compliant, we randomly select 9 integers between 1 and 52, compare the values on both distributions for these 9 positions employing the Anderson-Darling, Kolmogorov-Smirnov, Kuiper, Cramér-von Mises, and Watson U^2 goodness-of-fit tests, and calculate the mean p-value. We repeat the process 52 times to ensure that the arithmetic mean of these 52 goodness-of-fit p-values exceeds the 10% significance level. Because the result of this process yields a value of 21.77%, we cannot reject the prior hypothesis (see Figure 6 at the bottom-left).

The divisors of the Weyl order of the E6' s adjoint irrep (with label $(0,1,0,0,0,0)$ or its conjugate $(0,0,0,0,1,0)$) are

$$\text{đ}_{\mathsf{E}_6} = \{1\ 2\ 3\ 4\ 5\ 6\ 8\ 9\ 10\ 12\ 15\ 16\ 18\ 20\ 24\ 27\ 30\ 32\ 36\ 40\ 45\ 48\ 54\ 60\ 64\ 72\ 80\ 81\ 90\ 96\ 108\ 120\ 128\ 135\ 144\ 160\ 162\ 180\ 192\ 216\ 240\ 270\ 288\ 320\ 324\ 360\ 384\ 405\ 432\ 480\ 540\ 576\ 640\ 648\ 720\ 810\ 864\ 960\ 1080\ 1152\ 1296\ 1440\ 1620\ 1728\ 1920\ 2160\ 2592\ 2880\ 3240\ 3456\ 4320\ 5184\ 5760\ 6480\ 8640\ 10368\ 12960\ 17280\ 25920\ 51840\}$$

with cardinality $d(|W(\mathsf{E}_6)|) = d(51840) = 80$. By calculating the$\text{đ}(\mathsf{E}_6) \otimes \text{đ}(\mathsf{E}_6)$ tensor product, flattening, and sorting, we obtain a sequence of 80^2 divisors. By taking one out of every $\left\lfloor \frac{80^2}{351} \right\rfloor = 18$ divisors, this sequence produces eighteen sequences (those starting from the 83rd to the 100th element) of 351 divisors, which is the algebraical dimension of the most informative irrep of E6, precisely the irrep with label $(0,0,1,0,0,0)$) of E6 (or its conjugate $(0,0,0,1,0,0)$). Calculate the average partial sums,

 doi:10.20944/preprints202603.0536.v1

$\frac{\sum_{k=1}^{n} đ_{\mathsf{E}_6^2}(k)}{n}$, for the eighteen sequences of 351 divisors. In particular, the average partial sums of the sequence starting at the 97th element

$đ_{\mathsf{E}_6^2}$ = (45 54 64 75 90 96 108 128 144 160 162 180 200 216 240 270 288 300 324 360 360 400 432 450 480 486 540 576 600 640 648 720 720 768 810 864 900 960 972 1080 1080 1152 1152 1215 1296 1296 1440 1440 1536 1600 1620 1728 1728 1920 1920 1944 2048 2160 2160 2304 2400 2430 2592 2592 2700 2880 2880 3072 3240 3240 3240 3456 3600 3645 3840 3888 4050 4320 4320 4320 4608 4800 4860 5120 5184 5184 5760 5760 5760 5832 6400 6480 6480 6480 6912 7200 7290 7680 7776 7776 8640 8640 8640 8640 9216 9600 9720 10240 10368 10368 10800 11520 11520 11520 11664 12800 12960 12960 12960 13824 14400 14400 14580 15360 15552 16200 17280 17280 17280 17280 18225 19200 19440 19440 20480 20736 20736 21600 23040 23040 23328 24300 25600 25920 25920 25920 26244 27648 28800 29160 30720 31104 31104 32400 32805 34560 34560 34560 36450 38400 38880 38880 40960 41472 43200 43200 46080 46080 46656 48600 51200 51840 51840 51840 51840 55296 57600 58320 58320 61440 62208 64800 64800 69120 69120 69120 72900 76800 77760 77760 77760 82944 86400 86400 87480 92160 93312 93312 97200 103680 103680 103680 103680 110592 115200 116640 116640 122880 124416 129600 129600 138240 138240 139968 145800 155520 155520 155520 155520 165888 172800 172800 174960 184320 186624 194400 207360 207360 207360 207360 221184 230400 233280 233280 248832 259200 259200 259200 276480 276480 279936 291600 311040 311040 311040 331776 345600 349920 349920 373248 388800 388800 414720 414720 414720 437400 466560 466560 466560 497664 518400 518400 552960 552960 559872 583200 622080 622080 656100 691200 699840 699840 746496 777600 777600 829440 829440 874800 921600 933120 933120 995328 1036800 1049760 1105920 1166400 1166400 1244160 1244160 1382400 1399680 1399680 1492992 1555200 1658880 1658880 1749600 1866240 1866240 1990656 2073600 2099520 2239488 2332800 2488320 2488320 2764800 2799360 2799360 3110400 3317760 3317760 3499200 3732480 3981312 4147200 4199040 4665600 4976640 4976640 5598720 5598720 6220800 6635520 6998400 7464960 8294400 8398080 9331200 9953280 11059200 11197440 12441600 13996800 14929920 16796160 18662400 19906560 22394880 24883200 29859840 33592320 37324800 44789760 55987200 67184640 89579520 111974400 167961600 298598400 2687385600)

produce a distribution that, when normalized (and reversed), approaches NBL in radix 351 with RRMSE 1.0% (see Figure 6 at the top-right). However, the null hypothesis that the datasets have the same distribution is rejected at the 10% significance level by all goodness-of-fit tests owing to the distribution's length.

To validate the hypothesis that the sequence generated from the divisors is NBL-compliant, we randomly select 9 integers from the 351 elements, compare the values on both distributions for these 9 positions employing the Anderson-Darling, Kolmogorov-Smirnov, Kuiper, Cramér-von Mises, and Watson U^2 goodness-of-fit tests, and calculate the mean p-value. We repeat the process 351 times to ensure that the arithmetic mean of these 351 goodness-of-fit p-values exceeds the 10% significance level. Because the result of this process yields a value of 12.37%, we cannot reject the prior hypothesis (see Figure 6 at the bottom-right).

What all this means is that a SLG of the Lie Squad can quite approximately realize the logarithmic encoding of a number on its own, which can explain the extensive presence of NBL on raw datasets and natural science in general.

8. Conclusions

Galileo proposed that nature is fundamentally mathematical. Less emphatically, Einstein, in a letter to Reichenbach (Doc. AEA 20-117, 8 April 1926), noted that connecting geometry to theories is a subjective choice. String M-Theory requires such a high degree of mathematical soundness that many advocates propose that mathematics dictates physics, i.e., our universe is a specific solution to the gamut of potential universes generated by mathematical possibilities. Wigner' s " unreasonable effectiveness" of mathematics and Tegmark' s view of the universe as a mathematical model both

 doi:10.20944/preprints202603.0536.v1

emphasize the role of calculus in reality. Theories such as Quantum Graphity and Pregeometry propose that spacetime springs from discrete configurations and combinatorics at the Planck scale, whereas Geometric Unity presents physics as a projection from multidimensional systems. Our study accordingly points to topology and algebra having a potential material status, as evidenced by facts that lie midway between mathematics and physics.

We presuppose that information is physical. When used as a measure of surprise or likelihood, information can be expressed as entropy or Hartley information. Normalized likelihood gives rise to probability, particularly NBL. The representational cost of a number subsumes the NBL and is a measure of extent. This bridge between information and space is supported by the overarching consensus in quantum mechanics that space reflects entanglement probabilities, and in quantum gravity research, that information is fundamental.

Surely, the association of information with space should be independent of the dimension of the latter. Despite randomness and symmetry driving a dimensional explosion, generating new dimensions is not free; information within high-dimensional Euclidean spaces entails considerable handicaps. While points can be close or far apart in low dimensions, in high dimensions, the center of a hyper-object is effectively empty. The distance between the opposite corners of a unit hypercube becomes large, spreading the points out widely, and every point is roughly equidistant from all other points. Although a sphere inside a cube touches the sides in two and three dimensions, as dimensions increase, the distance from the center of the cube to its vertices grows much faster than the radius of the sphere. The high-dimensional Euclidean space is an exaggeration.

To address the curse of dimensionality, we investigate how to extrapolate the NBL in ambient spaces of 2, 3, and m dimensions. We define the NBL cumulative probability distribution of a digit k with radix r in the m-dimensional space as the one-dimensional NBL cumulative probability distribution to the mth power, i.e., $\Pr_m(r, [1, k+1)) = (\Pr(r, [1, k+1)))^m = \ln_r^m(k+1)$. This rule is also valid for any numeral $k = R$, not just for digits $k < r$.

We recast the representational cost of a m-tuple of data written in GPN as $C(m, r, R) = \frac{r}{4}\Upsilon_m(r, R)$, where $\Upsilon_m(r, R) = V_m \log_r^m(R+1)$ measures the total logarithmic volume and $V_m = \frac{\pi^{\frac{m}{2}}}{\Gamma\left(1+\frac{m}{2}\right)}$ is the standard volume of the unit m-ball. Compared with the volume of the mDE space, we have replaced the uniform distribution with the NBL. Specifically, we have replaced the radius R with the r-ary logarithmic span of the interval between the origin and R, i.e., $R \mapsto \Pr(r, [1, R+1))$. We factually obtain the standard (Euclidean) volume of the m-ball by setting $R = r-1$. We evaluate the rate of a k-cut to the entire unit m-ball, and notice that the stuff is now concentrated within the origin's neighborhood instead of being displaced to the $(m-1)$-sphere. This approach diminishes the curse of dimensionality; a logarithmic volume of an m-ball is generally tractable, and we can move the peak of the lognormal curve of logarithmic volumes to the right if we decrease the radix. The proof that the information we perceive stems from natural sources that encode data logarithmically, not linearly, is the overwhelming presence of NBL we perceive in our (in principle) four-dimensional world.

We can describe the space using the topological plus algebraic observables of the SLGs, namely, connectivity, compactness, structure, homogeneity, and extension. We have learned that the leading numeral R of the most informative irrep of an SLG occurs with probability $\Pr_m(R) \approx \frac{1}{m}\ln\left(1+\frac{1}{R}\right)$. When we focus on the most important SLGs of the lower algebraic ranks, we find that the most thrifty groups are B3, F4, G2, A2, A1, and E6, with irreps whose algebraic dimension approximates the ideal ideal e^m quite closely, where e is Euler's number and m is the algebraic rank. The abundance of B3, as well as A2 and A1 in our world, agrees with this result. Besides, because the Lie Squad forms a chain of embeddings, we conjecture that G2, F4, and E6 play a crucial role in the cosmos. Further, if a being belonging to our world is the embodiment of an algebraic structure via geometry, F4 might be the cradle we live in.

SBN is the universal, unambiguous GPN that the Lie Squad employs. It is primarily based on balanced ternary and encodes every integer as a difference of two naturals written in bijective notation. SBN is scale, position, length, and rank invariant. In general, invariance must be interpreted

 doi:10.20944/preprints202603.0536.v1

as noise-tolerance of any kind and a guarantee of computability. In particular, because SBN preserves the algebraic rank, it is a number system valid for any SLG, and hence, different Lie groups might coexist in nature.

Finally, we examine the Weyl groups of the Lie Squad. The Weyl group describes the symmetries of a root system (e.g., 2 is always a divisor), which is an intrinsic property of a Lie algebra and, by extension, the Lie group itself. Therefore, it defines the group's basic structure and is not associated with a specific irrep. Well, we knew the Weyl divisors govern the geometry of the fundamental chamber, the degrees of invariant polynomials, and its topology (cohomology). In this article, we argue that the Weyl order is also critical to create exact proportions that allow the Lie group represent the information on a logarithmic scale, so that an SLG can reproduce its NBL on its own simply from the calculus of Weyl divisors.

This manner in which space and information are related is reminiscent of the way entropy changes, ΔS, during the expansion (or compression) of an ideal gas with an increment of volume ΔV and pressure ΔP at constant temperature, or during heating (or cooling) at constant pressure or volume when the temperature varies. Specifically, compare the expression for $C(m, r, R)$ with $\Delta S = nk_X \log\left(1 + \frac{\Delta X}{X_0}\right)$, where n is the number of moles of gas, and k_X, ΔX, and X_0 represent the physical constants, increments, and initial values, respectively, of volume, pressure, or temperature. Now, think of a potential difference, or a temperature gap. When two systems are brought into contact, their information is shared; mightn' t we be observing how the transitory states recode the data on a common SBN logarithmic scale, eventually smoothing out the gaps?

Furthermore, why are electron orbitals for hydrogen-like atoms filled in by principal quantum number n=1,2,3 in order? Why are those eigenvalues not treated on an equal footing by nature? Why are the features producing the least information more stable? Because the universe has finite resources. SBN explains the tendency towards the lowest-energy states and why the lightest atoms are more abundant than the heavier ones. It seems that proximity (distance from the nucleus) and smallness are foundational, as well as progressive incrementality (what Brouwer would call constructability).

9. Postscript

This apopemptic chapter includes additional information about the context of our activity during the article's development.

Data Availability Statement: Only genuine contributions are included in the article. Subsequent inquiries can be directed to the author. Specifically, the programming code used to generate the data presented in this study is available upon request. The programming framework used is WolframAlpha, whose language is based on Mathematica and its associated online services. Although these are closed source, Wolfram Research has released a parser for the language, originally developed in C++ and rewritten in Rust, under the open-source MIT License.

Acknowledgments: We hereby express our tribute to the personalities and entities that have influenced this work. The UNED instilled a scientific instinct and a general method to problem-solving. Working for Indra enabled a practical approach to addressing crises and an awareness of the productive character of our universe. We greatly thank José Mira for his motivational capacity. We apologize in advance for omitting laudable references. We greatly appreciate those who offer comments and critical insights on this version.

Conflicts of Interest: This research received no external funding. Therefore, the author (ORCID 0000-0003-3980-5829) declares no conflicts of interest. The author asserts that exclusively scientific significance drives the work's high-level goals. It contains no known minor or significant incongruities, errors, or inaccuracies. The author did not receive support from any organization for the submitted work and declares that he has no competing financial or nonfinancial interests directly or indirectly related to the work submitted for publication. No personal relationships have influenced the content of this work. It has not been published elsewhere in any form or language, in whole or in part. The author claims to have committed no intentional ethical wrongdoing related to this paper, including self-plagiarism, plagiarism, far-fetched self-citations, conflict of interest, inaccurate authorship declarations, and unacceptable biases concerning the references. This work respects the rights of third parties, including copyright and moral rights.

 doi:10.20944/preprints202603.0536.v1

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviations (in alphabetical order)	
GPN	Geometric Positional Notation
irrep	irreducible representation
nDE	n-dimensional Euclidean
NBL	Newcomb-Benford Law
RRMSE	Relative Root Mean Squared Error
SBN	Symmetric Bijective Notation
SLG	Simple Lie Group
SPN	Standard Positional Notation
List of MSC (Mathematics Subject Classification) epigraphs	
11A63	Radix representation; digital problems
11A67	Other representations
11B50	Sequences
14L35	Classical groups (geometric aspects)
17B10	Representations, algebraic theory (weights)
17B20	Simple, semisimple, reductive (super)algebras (roots)
22E15	General properties and structure of real Lie groups
51M05	Euclidean geometries (general) and generalizations
51M10	Hyperbolic and elliptic geometries (general) and generalizations
51M20	Polyhedra and polytopes; regular figures, division of spaces
51M25	Length, area and volume
60B99	None of the above, but in this section
68P30	Coding and information theory (compaction, compression, models of communication, encoding schemes, etc.)
94A15	Information theory, general
94A17	Measures of information, entropy

References

1. Wheeler, J.A. Information, Physics, Quantum: The Search for Links. In *Proceedings III International Symposium on Foundations of Quantum Mechanics*; Physics Department, Princeton University, 1989; pp. 354–358.
2. Landauer, R. Information is Physical. In Proceedings of the Workshop on Physics and Computation, Dallas, TX, USA, USA, 1992; pp. 1–4. https://doi.org/10.1109/PHYCMP.1992.615478.
3. Einstein, A. Cosmological Considerations in the General Theory of Relativity. *Sitzungsber. Preuss. Akad. Wiss. Berlin (Math. Phys.)* **1917**, *1917*, 142–152.
4. Barbón, J.L.F. Black Holes, Information and Holography. *Journal of Physics: Conference Series* **2009**, *171*, 012009. https://doi.org/10.1088/1742-6596/171/1/012009.
5. Okon, E.; Sudarsky, D. The Black Hole Information Paradox and the Collapse of the Wave Function. *Foundations of Physics* **2015**, *45*, 461–470. https://doi.org/10.1007/s10701-015-9877-6.
6. Vopson, M.M. The Mass-Energy-Information Equivalence Principle. *AIP Advances* **2019**, *9*, 095206, [https://pubs.aip.org/aip/adv/article-pdf/doi/10.1063/1.5123794/19733686/095206_1_online.pdf]. https://doi.org/10.1063/1.5123794.
7. Neukart, F.; Brasher, R.; Marx, E. The Quantum Memory Matrix: a Unified Framework for the Black Hole Information Paradox. *Entropy* **2024**, *26*, 1039. https://doi.org/10.3390/e26121039.
8. Rives, J. The Code Underneath. *Axioms (CC BY license, creativecommons.org/licenses/by/4.0/)* **2025**, *14*, 106. https://doi.org/10.3390/axioms14020106.
9. Ryu, S.; Takayanagi, T. Aspects of Holographic Entanglement Entropy. *Journal of High Energy Physics* **2006**, *2006*, 045. https://doi.org/10.1088/1126-6708/2006/08/045.
10. Maldacena, J.; Susskind, L. Cool Horizons for Entangled Black Holes. *Fortschritte der Physik* **2013**, *61*, 781–811. https://doi.org/10.1002/prop.201300020.
11. Barbachoux, C.; Kouneiher, J. Analytical and Geometric Foundations and Modern Applications of Kinetic Equations and Optimal Transport. *Axioms* **2025**, *14*. https://doi.org/10.3390/axioms14050350.

doi:10.20944/preprints202603.0536.v1

12. Hill, T.P. The Significant-Digit Phenomenon. *The American Mathematical Monthly* **1995**, *102*, 322–327. https://doi.org/10.1080/00029890.1995.11990578.
13. Pinkham, R.S. On the Distribution of First Significant Digits. *The Annals of Mathematical Statistics* **1961**, *32*, 1223–1230. https://doi.org/10.1214/aoms/1177704862.
14. Hill, T.P. Base-Invariance Implies Benford's Law. *Proceedings of the American Mathematical Society* **1995**, *123*, 887–895.
15. Lehmann, E.L.; Casella, G. *Theory of Point Estimation*; Springer Texts in Statistics, Springer New York, 2003.
16. Chirikjian, G.S. *Stochastic Models, Information Theory, and Lie Groups, Volume 1: Classical Results and Geometric Methods*; Applied and Numerical Harmonic Analysis, Birkhäuser Boston, 2009.
17. Chirikjian, G.S. *Stochastic Models, Information Theory, and Lie Groups, Volume 2: Analytic Methods and Modern Applications*; Applied and Numerical Harmonic Analysis, Birkhäuser Boston, 2011.
18. Hall, B. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*; Graduate Texts in Mathematics, Springer International Publishing, 2015.
19. Helgason, S. *Differential Geometry, Lie Groups, and Symmetric Spaces*; Crm Proceedings & Lecture Notes, American Mathematical Society, 2001.
20. Nigrini, M.J. *Digital analysis using Benford's Law*; 2000.
21. Despotovic, M.; Nedic, V.; Despotovic, D.; Cvetanovic, S. Evaluation of empirical models for predicting monthly mean horizontal diffuse solar radiation. *Renewable and Sustainable Energy Reviews* **2016**, *56*, 246–260. https://doi.org/https://doi.org/10.1016/j.rser.2015.11.058.
22. nLab authors. Moduli Stabilization, 2025. Revision 26.
23. Schondorf, E. The Wiener Measure and Donsker's Invariance Principle. In Proceedings of the The University of Chicago Mathematics REU 2019; of Chicago, U., Ed., 2019, Full program.
24. Mörters, P.; Peres, Y. *Brownian Motion*; Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2010.
25. Bellman, R.E. *Dynamic Programming*; Dover Books on Computer Science Series, Dover Publications, 2003.
26. Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*; Springer Series in Statistics, Springer New York, 2009.
27. Rowe, C. The Higher Dimensions Series IV: The Probability Theory Connection. *Medium* **2020**.
28. Barrow, J.D. *Pi in the Sky - Counting, Thinking, and Being*; Clarendon Press, 1992.
29. Wikipedia contributors, the Free Encyclopedia. Positional Notation, 2022. [Online; accessed 11-November-2022].
30. Forslund, R.R. A Logical Alternative to the Existing Positional Number System. *Southwest Journal of Pure and Applied Mathematics* **1995**, *1*, 27–29.
31. Fleet, J.F. Aryabhata's System of Expressing Numbers. *Journal of the Royal Asiatic Society of Great Britain and Ireland* **1911**, *43*, 109–126. https://doi.org/10.1017/S0035869X00040995.
32. Cthaeh, T. Numeral Systems: Everything You Need to Know. *Probabilistic World* **2021**.
33. Gilbert, E.N. Gray Codes and Paths on the n-Cube. *Bell System Technical Journal* **1958**, *37*, 815–826, [https://onlinelibrary.wiley.com/doi/pdf/10.1002/j.1538-7305.1958.tb03887.x]. https://doi.org/https://doi.org/10.1002/j.1538-7305.1958.tb03887.x.
34. Herter, F.; Rote, G. Loopless Gray Code Enumeration and the Tower of Bucharest. In Proceedings of the 8th International Conference on Fun with Algorithms (FUN 2016); Demaine, E.D.; Grandoni, F., Eds., Dagstuhl, Germany, 2016; Vol. 49, *Leibniz International Proceedings in Informatics (LIPIcs)*, pp. 19:1–19:19. https://doi.org/10.4230/LIPIcs.FUN.2016.19.
35. Khmelnik, S.I., The Papers of Independent Authors, Volume 7; DNA - Lulu.com, 2008; chapter Computer Engineering: Positional Codes of Complex Numbers and Vectors, pp. 19–44.
36. Knuth, D.E. *Art of Computer Programming, Volume 2: Seminumerical Algorithms*; Pearson Education, 2014.
37. Klir, G.J.; Folger, T.A. *Fuzzy Sets, Uncertainty, and Information*; Prentice-Hall International editions, Prentice-Hall International, 1988.
38. Manack, C.; Miller, S.J. Leading Digit Laws on Linear Lie Groups. *Research in Number Theory* **2015**, *1*, 1–19. Article number: 22.
39. Wangberg, A.; Dray, T. Visualizing Lie Subalgebras Using Root and Weight Diagrams. *Loci* **2009**, *1*. https://doi.org/10.4169/loci003287.
40. Lisi, A.G. An Exceptionally Simple Theory of Everything, 2007, [arXiv:hep-th/0711.0770].
41. Lisi, A.G.; Weatherall, J.O. A Geometric Theory of Everything. *Scientific American* **2010**, *303N6*, 30–37.
42. Coxeter, H.S.M. *Regular Polytopes*; Dover Books on Mathematics, Dover: New York, NY, 1973.

43. Acharya, B.S. M Theory, G2-Manifolds and Four-dimensional Physics. *Classical and Quantum Gravity* **2002**, *19*, 5619–5653. https://doi.org/10.1088/0264-9381/19/22/301.
44. Cossu, G. Dual Superconductivity in G2 Group. In Proceedings of the XXIVth International Symposium on Lattice Field Theory, 2006, Vol. LAT2006, p. 63, [arXiv:hep-lat/hep-lat/0609061]. https://doi.org/10.22323/1.032.0063.
45. Holland, K.; Minkowski, P.; Pepe, M.; Wiese, U.J. Exceptional confinement in G(2) gauge theory. *Nuclear Physics B* **2003**, *668*, 207–236, [hep-lat/0302023]. https://doi.org/10.1016/S0550-3213(03)00571-6.
46. Wikipedia contributors, the Free Encyclopedia. Quantum Chromodynamics, 2020.
47. Schatte, P. On Mantissa Distributions in Computing and Benford's law. *Journal of Information Processing and Cybernetics* **1988**, *24*, 443–455.
48. Frankel, R. The HyperSphere, from an Artistic point of View. MIT Computer Science and Artificial Intelligence Lab, 2005, [https://groups.csail.mit.edu/mac/users/rfrankel/fourd/FourDArt.html]. Project MAC ("Switzerland") - MIT Project on Mathematics and Computation.
49. Glusker, M.; Hogan, D.M.; Vass, P. The Ternary Calculating Machine of Thomas Fowler. *IEEE Annals of the History of Computing* **2005**, *27*, 4–22. https://doi.org/10.1109/MAHC.2005.49.
50. Wikipedia contributors, the Free Encyclopedia. 2-21 Polytope, 2020.
51. Wangberg, A. The Structure of E6. PhD thesis, Oregon State University, 2007.
52. Hayes, B. Third Base. *American Scientist (reprint)* **2001**, *89*, 490–494.
53. Andrews, G. Euler's De Partitio Numerorum. *Bulletin (New Series) of the American Mathematical Society* **2007**, *44*.
54. Wikipedia contributors, the Free Encyclopedia. Bijective Numeration, 2020.