Article

Not peer-reviewed version

A Curious Effect of Benford's Law for Bijective Numeration

Julio Rives *

Posted Date: 21 March 2025

doi: 10.20944/preprints202503.1625.v1

Keywords: Probability Mass Function; Newcomb-Benford Law; Positional Notation; Harmonic scale; Kempner's curious series; Bijective Numeration; Logarithmic scale

Article

A Curious Effect of Benford's Law for Bijective Numeration

Julio Rives

Dept. of Artificial Intelligence, Univ. Nacional de Educacián a Distancia (UNED), Paseo Senda del Rey, 7, Madrid 28040, Spain; jrives@dia.uned.es)

Abstract: We assume that the probability mass function $\Pr(Z) = (2Z)^{-2}$ $(Z \in \mathbb{Z}^+)$ is at Newcomb-Benford Law's root and the origin of positional notation. Under its tail, we find that the harmonic (global) $\mathbb{Q}$-NBL for bijective numeration is $\Pr(\underline{b}, q) = \left(qH_{\underline{b}}\right)^{-1}$, where q is a quantum $(1 \leq q \leq \underline{b})$, H_n is the nth harmonic number, and $\underline{b}$ is the bijective base. Under its tail, the logarithmic (local) $\mathbb{R}$-NBL for bijective numeration is $\Pr(\underline{r}, d) = \log_{\underline{r}+1}(1 + 1/d)$, where $d \leq \underline{r} \ll \underline{b}$, being d a digit of a local complex system's bijective radix $\underline{r}$. We generalize both lows to calculate the probability mass of the leading quantum/digit of a chain/numeral of a given length and the probability mass of a quantum/digit at a given position, verifying that the global and local NBL are length- and position-invariant in addition to scale-invariant. In the framework of bijective numeration, we also prove that the sums of Kempner's series conform to the global Newcomb-Benford Law and suggest a natural resolution for the precision of a universal positional notation system.

Keywords: Probability Mass Function (PMF); Newcomb-Benford Law; Positional Notation (PN); Harmonic scale; Kempner's curious series; Bijective Numeration (BN); Logarithmic scale

MSC: 11A63 - Radix representation; digital problems; 11A67 - Other number representations; 40-11 - Research data for problems pertaining to sequences, series, summability; 40A05 - Convergence and divergence of series and sequences; 60A99 - None of the above, but in "Foundations of probability theory"; 60E05 - Probability distributions: general theory; 68P30 - Coding and information theory; 94A17 - Measures of information, entropy

---

1. Introduction

The fiducial Newcomb-Benford law (NBL) [1,2] is well-established for standard positional notation (PN) [3]. It states that the first digits of randomly chosen original data typically outline a logarithmic curve in diverse fields irrespective of their physical units; series of raw natural values usually are nearly scale-invariant, i.e., geometric [4]. The law must be valid for any place-value numeral system if it is fundamental. This paper provides the formulas for bijective numeration (BN) [5], suggesting that NBL might account for an elementary principle.

Suspicion about the authenticity of the number cero [6] suggests that BN [7] is likely more natural than standard PN, the number system we use daily. BN is a zero-free and unambiguous number system, contrasting with standard PN; every natural number has a unique representation using the symbols $\{1, 2, \cdots, \underline{r}\}$. (Watch the notation; we display the radix underlined to denote BN instead of standard PN.)

Every NBL rule for standard PN has a bijective counterpart. BN is a notably more efficient notation than standard PN due to its compactness, especially for the smaller radices $\underline{r} \in \mathbb{N}^+$. An enumerated list of $\underline{r}$-ary bijective numerals is automatically in "shortlex" order, with $\underline{r}^l$ bijective $\underline{r}$-ary numbers of length $l \geq 1$ [8]. For instance, the sequence of 11 ternary bijective numbers $\{3, 11, 12, 13, 21, 22, 23, 31, 32, 33, 111\}_{\underline{3}}$, from decimal $3 = 3 \times 3^0$ to $13 = 1 \times 3^2 + 1 \times 3^1 + 1 \times 3^0$, has $3^2 = 9$ numbers of length 2.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 March 2025 doi:10.20944/preprints202503.1625.v1

BN is a natural consequence of NBL, which can originate from a primordial inverse-square law [9]. This "canonical" probability mass function (PMF) might be the sought-after cause of NBL [10]. Under the tail of the canonical PMF, we find information based on the concept of likelihood; probability is a relative likelihood. The theory has fundamentally a twofold manifestation, namely the NBL for the global ("rational", discrete, and harmonic) and local ("real", continuous, and logarithmic) domains.

Departing from $\Pr(Z) = (2Z)^{-2}$ (where $Z \in \{\pm1, \pm2, \pm3, \cdots\}$), the probability of a natural variable falling into $[s, t)$ is a harmonic likelihood, namely the bucket's width $\psi(t) - \psi(s) = H_{t-1} - H_{s-1}$ relative to the base's support width $\psi(\underline{b}) - \psi(1) = H_{\underline{b}}$, where $s < t \leq \underline{b}$, $\{s, t, \underline{b}\} \in \mathbb{N}^+$, and H_n is the nth harmonic number. The base $\underline{b}$ is a global referent that changes the status of a number to a globally computable elemental entity denominated a "quantum". When the bucket is $[q, q+1)$, we obtain the global NBL of a generic quantum q, $\Pr(\underline{b}, q) = \frac{1}{qH_{\underline{b}}}$, where $q \in \mathbb{N}^+$ and $1 \leq q \leq \underline{b}$. Thus, the mass of a quantum is a rational number representing the harmonic improbability of an elemental gap. When $\underline{b}$ has a colossal value, a "coding source" must establish a local referent $\underline{r} \ll \underline{b}$ to normalize its information separated from the surrounding environment, changing the status of a quantum to a locally computable elemental entity denominated a "digit". The probability of a quantum falling into $[i, j)$ is the bin's width $\ln j/i$ relative to the radix support's width $\ln \underline{r}$. When the bin is $[d, d+1)$, we arrive at $\log_{\underline{r}+1}(1 + 1/d)$, where $d \in \mathbb{N}^+$ is a digit such that $1 \leq d \leq \underline{r}$.

The only symbol in the bijective unary is 1; contrary to standard PN, this numeral system is not exceptional and can result from the natural continuation of a radix reduction process. In the bijective binary, 1 occupies $2/3 \approx 66.7\,\%$ (global) and $\ln 2/\ln 3 \approx 63.1\,\%$ (local), while 2 occupies $1/3 \approx 33.3\,\%$ and $\ln 1.5/\ln 3 \approx 36.9\,\%$, respectively. Likewise, 1 in the bijective nonary occupies 35.3 % (global) and 30.1 % (local), while 9 occupies 3.9 % and 4.6 %, respectively. The harmonic plot is always steeper than the logarithmic one.

This double-scale theory unites efficiency and entropy. "The lowest digits maintain distinctness from the surroundings thanks to their solid entropic support. The more significant digits are vulnerable and give rise to more transitions..." From the information perspective, whereas the first quanta or digits make the difference irrespective of the string length, the last ones might provide negligible, even arbitrary, information [11].

We back this claim by delving into the sums of Kempner's curious harmonic series [12], which echo the bijective harmonic scale traced by the global NBL. This outcome is absolute because every Kempner series is infinite, and the calculations consider every possible numerical chain. Short numerical chains or low quantum densities are accessible and cheap, producing heavy harmonic terms that condense the space. In contrast, long chains or high densities are "rare" and deliver slender harmonic terms.

For instance, while the specific digits involved in a constraining numeral do not matter, the length of such a numeral does; the law favors 0.234 against 123.4 and this against 12.345 because 12345 is less probable than 1234 and this is less probable than 234. Likewise, while removing the harmonic decimal terms that include less than 10 % of 5's in the denominator makes a harmonic series converge, missing the terms including $\geq 10\,\%$ of 5's does not impede the divergence of the depleted harmonic series.

Our study of the depleted and constrained harmonic series allows us to conclude that a conservative policy of significands and positions might exist inherent to universal computation, restricting the resolution of PN to preclude unnecessary precision. Consequently, we conjecture that the natural span of a positional system in base $\underline{b}$ is $\underline{b}^{\underline{b}}$, a measure of the physical quantity of tractable numerals. Beyond this computational resolution, quanta could be haphazard for practical purposes.

This article's field of study is information, probability, and number theory. First, we describe the global and local NBL theory for BN, unveiling that both laws are length- and position-invariant. Then, we substantiate that the set of Kempner's "curious" series conforms to the global NBL for BN. Further, we surmise a universal resolution resolution, i.e., the prospect of a position threshold ascribed to a natural PN system.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 March 2025 doi:10.20944/preprints202503.1625.v1

2. The Global NBL for BN

This subsection follows the same plot thread developed in [9]'s section "The rational (global) version of NBL".

A sample of numerical chains encoded using bijective $\underline{b}$-ary satisfies the global NBL if the leading quantum falls in bucket $[s,t)$ relative to the area swept by base $\underline{b}$ with probability

$$\Pr(\underline{b},[s,t)) = \frac{H_{t-1} - H_{s-1}}{H_{\underline{b}}} \in \mathbb{Q}$$

where $1 \leq s < t \leq \underline{b}$ and $s,t,\underline{b} \in \mathbb{N}^+$. When $s = q$ and $t = q+1$ we obtain the probability with base $\underline{b}$ of leading quantum q.

Definition 1. *A leading quantum q is said to satisfy the global NBL for bijective $\underline{b}$-ary numeration if it occurs with probability*

$$\Pr(\underline{b},q) = \frac{1}{qH_{\underline{b}}} \in \mathbb{Q}\ (1 \leq q \leq \underline{b}, \{q,\underline{b}\} \in \mathbb{N}^+)$$

Thus, NBL for the standard PN in base $b+1$ corresponds to NBL for bijective $\underline{b}$-ary numeration.

Example 1. *We obtain* $\Pr(\underline{1},1) = 100\,\%$, $\Pr(\underline{3},1) = 6/11 \approx 54.5\,\%$, $\Pr(\underline{3},2) = 3/11 \approx 27.3\,\%$, $\Pr(\underline{3},3) = 2/11 \approx 18.2\,\%$, $\Pr(\underline{\mathrm{A}},1) = 0.34142$, *and* $\Pr(\underline{\mathrm{A}},\mathrm{A}) = 0.03414$, *where "$\underline{\mathrm{A}}$" symbolizes the bijective decimal base. Owing to* $\Pr(\underline{2},1) = 2/3$ *and* $\Pr(\underline{2},2) = 1/3$, *the odds* $o(2:1|\underline{2}) = \Pr(\underline{2},2)/\Pr(\underline{2},1) = ½$ *constitute an essential sharing out.*

The entropy of PMF (1), $\widetilde{E}(\underline{b})$, is the expected value (weighted arithmetic mean) of the harmonic likelihood function, namely $\psi(x) - \psi(1) = H_{x-1}$ (where ψ is the digamma), evaluated at x's probability mass reciprocal.

Definition 2. *The global NBL entropy is given by the formula*

$$\widetilde{E}(\underline{b}) = \widehat{E}(\Pr(\underline{b},q)) = \sum_{q=1}^{q=\underline{b}} \frac{H_{qH_{\underline{b}}-1}}{qH_{\underline{b}}}\ \text{harmt}$$

Example 2. $\widetilde{E}(\underline{1}) = 0$, $\widetilde{E}(\underline{2}) = 0.90914\,\text{harmt}$, $\widetilde{E}(\underline{3}) = 1.35432\,\text{harmt}$, $\widetilde{E}(\underline{10}) = 2.47676\,\text{harmt}$, *and* $\widetilde{E}(\underline{100}) = 4.2269\,\text{harmt}$.

When $\underline{b}$ acquires a gargantuan value, we can take the summation as an integral and the harmonic number function as the natural logarithm, so that the "differential entropy" [13] of the global NBL approximately tends to

$$\int_1^{\underline{b}} \frac{\ln(q \ln \underline{b})}{q \ln \underline{b}}\,\mathrm{d}q = ½ \ln \underline{b} + \ln(\ln \underline{b})$$

Thus, the global entropy is finite, which agrees with the Bekenstein bound in physics [14].

The probability of picking a chain of any length starting with c is the likelihood gap it induces on the $\underline{b}$-ary harmonic scale.

Definition 3. *A leading numerical chain c is said to satisfy the global NBL for bijective $\underline{b}$-ary numeration if it occurs with probability*

$$\Pr(\underline{b},c) = \frac{H_c}{H_{\underline{b}}} - \frac{H_{c-1}}{H_{\underline{b}}} = \frac{1}{cH_{\underline{b}}} \in \mathbb{Q}$$

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 March 2025 doi:10.20944/preprints202503.1625.v1

This occurrence probability becomes (1) when c is a base's quantum.

Example 3. *The probability that a bijective decimal chain starts with* 11 *(e.g.,* .111*) and* AA *(e.g.,* AAAA*) is* $\Pr(\underline{A}, 11) = 1/(11H_{10}) \approx 0.03104$ *and* $\Pr(\underline{A}, \mathrm{AA}) = 1/(110H_{10}) \approx 0.003104$*, respectively.*

Definition (3) allows us to derive the following PMF.

Definition 4. *A numerical dataset is said to satisfy the global NBL for BN if the probability of picking a length-l bijective $\underline{b}$-ary chain starting with the quantum q, where $1 \leq q \leq \underline{b}$ and $\{\underline{b}, l, q\} \in \mathbb{N}^+$, is*

$$\Pr(\underline{b}, l, q) = \frac{\sum_{k=q\underline{b}^{l-1}+\frac{\underline{b}^{l-1}-1}{\underline{b}-1}}^{k=(q+1)\underline{b}^{l-1}+\frac{\underline{b}^{l-1}-1}{\underline{b}-1}-1} \frac{1}{k}}{H_{\frac{\underline{b}^{l+1}-1}{\underline{b}-1}-1} - H_{\frac{\underline{b}^{l}-1}{\underline{b}-1}-1}} \in \mathbb{Q}$$

Example 4. *The probability of running into* 1 *to* 3 *as the first quantum of a bijective ternary chain with length* 5 *is* $\{0.46565, 0.30602, 0.22833\}$*, and the chances of choosing* 1 *to* A *as the first quantum of a bijective decimal chain with length* 2 *is* $\{0.2842, 0.1688, 0.1205, 0.09377, 0.07677, 0.065, 0.05637, 0.04976, 0.04454, 0.04031\}$.

Watch that (4) boils down to (1) if $l = 1$, meaning that the global NBL is length-invariant beside base-invariant.

Definition (3) also allows us to derive the following PMF.

Definition 5. *A numerical dataset is said to satisfy the global NBL for BN if the probability of runing into q as the p-th quantum of a bijective $\underline{b}$-ary chain, where $1 \leq q \leq \underline{b}$ and $\{\underline{b}, q, p\} \in \mathbb{N}^+$, is*

$$\Pr(\underline{b}, q, p) = \frac{\sum_{k=\frac{\underline{b}^{p-1}-1}{\underline{b}-1}}^{k=\frac{\underline{b}^{p}-1}{\underline{b}-1}-1} \frac{1}{\underline{b}k+q}}{H_{\frac{\underline{b}^{p+1}-1}{\underline{b}-1}-1} - H_{\frac{\underline{b}^{p}-1}{\underline{b}-1}-1}} \in \mathbb{Q}$$

Example 5. *The probability of getting* 1 *to* 3 *as the fifth quantum of a bijective ternary chain is* $\{0.335011, 0.333327, 0.331662\}$*, and the chances of encountering* 1 *to* A *as the second quantum of a bijective decimal chain is* $\{0.1183, 0.113, 0.1083, 0.1041, 0.1004, 0.09694, 0.09381, 0.09094, 0.08829, 0.08583\}$.

Watch that (5) boils down to (1) if $p = 1$, meaning that the global NBL is position-invariant beside base-invariant. Figure 1 shows the PMF of various bijective bases for consecutive positions and the hyperbolic progression of the bijective ternary digits as the position increases.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 March 2025 doi:10.20944/preprints202503.1625.v1

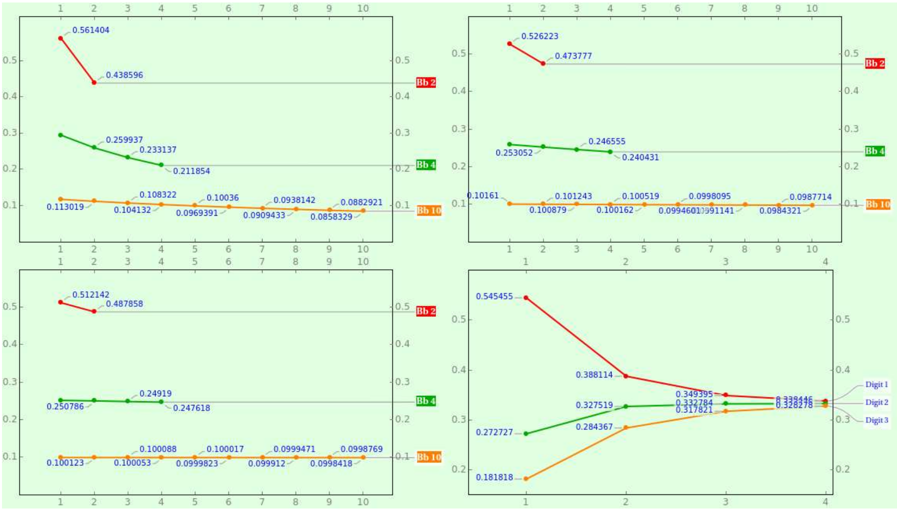

Figure 1. Leading quantum's PMF for bijective bases $\underline{2}$, $\underline{4}$, and $\underline{A}$ (decimal) at positions 2 (top-left), 3 (top-right), and 4 (bottom-left) using (5). The plots quickly tend to the uniform distribution. On the bottom right, we show the probability plot of the bijective ternary digits as a function of their position; only the first few quanta make a difference concerning information.

3. The Local NBL for BN

This subsection follows the plot thread developed in [9]'s section "The fiducial (local) NBL".

The ratio between the area under the hyperbola delimited by the bin $[d_1, d_2)$ and the radix support $[1, \underline{r} \ll \underline{b})$ is

$$\Pr(\underline{r}, [d_1, d_2)) = \log_{\underline{r}+1} \frac{d_2}{d_1} \in \mathbb{R} \; \left(1 \le d_1 < d_2 \le \underline{r}, \{d_1, d_2, \underline{r}\} \in \mathbb{N}^+\right)$$

We arrive at the NBL for BN by putting $d_1 = d$ and $d_2 = d + 1$.

Definition 6. *A sample of numerals expressed in bijective $\underline{r}$-ary PN satisfies the local NBL if the leading digit d occurs with probability*

$$\Pr(\underline{r}, d) = \log_{\underline{r}+1}\left(1 + \frac{1}{d}\right) \in \mathbb{R} \; (1 \le d \le \underline{r})$$

Thus, the NBL with radix $r + 1$ corresponds to the bijective $\underline{r}$-ary numeration's NBL.

Example 6. *The standard ternary system assigns to* 1 *and* 2 *the probabilities* 63 % *and* 37 %*, which is the PMF of bijective binary numeration. In the usual case where the radix is* $r = 10$*, the standard decimal system assigns to digits* 1 *and* 9 *probabilities of* 30.1 % *and* 4.6 %*. In contrast, the bijective decimal numeration assigns to digits* 1 *and* $\mathrm{A} \equiv 10$ *probabilities of* 28.9 % *and* 4.0 %*. On the other hand, the local bijective ternary numeration assigns to* 1, 2, *and* 3 *the probabilities* 50 %, 29 %, *and* 21 %*, contrasting with the percentages* 54.5 %, 27.3 %, *and* 18.2 % *the global bijective ternary numeration assigns.*

The entropy of PMF (6) for radix $\underline{r}$, $\widetilde{e}(\underline{r})$, is the expected value (weighted arithmetic mean) of the likelihood function ($\ln(x)$) evaluated at x's probability mass reciprocal.

Definition 7. *The local NBL entropy is given by the formula*

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 March 2025 doi:10.20944/preprints202503.1625.v1

$$\widetilde{e}(\underline{r}) = \widehat{E}(\Pr(\underline{r}, d)) = \sum_{d=1}^{d=\underline{r}} \log_{\underline{r}+1}\left(1+\frac{1}{d}\right) \ln\left(\frac{1}{\log_{\underline{r}+1}\left(1+\frac{1}{d}\right)}\right) \text{ nat}$$

Example 7. $\widetilde{e}(\underline{1}) = 0$, $\widetilde{e}(\underline{2}) = 0.65846$, $\widetilde{e}(\underline{3}) = 1.03247$, $\widetilde{e}(\underline{10}) = 2.08134$, *and* $\widetilde{e}(\underline{100}) = 3.84099$.

Because $\underline{r} < \underline{b}$ and we assume that $\underline{b}$ is a positive natural number, the local entropy is finite, in agreement with the Bekenstein bound.

Note that (6) is also valid for the unary numeral system, unlike in standard PN; $\underline{r} = 1$ assigns the probability of 100 % to 1. The scale of a system "encoding" data in the bijective unary is linear, i.e., has no curvature. In BN, (re)coding from unary into $\underline{r}$-ary means summing the number of ones and executing an iterative procedure based on Euclidean division; Figure 2 describes the encoding algorithm that converts the representation of 1567 into $1233231_{\underline{3}}$.

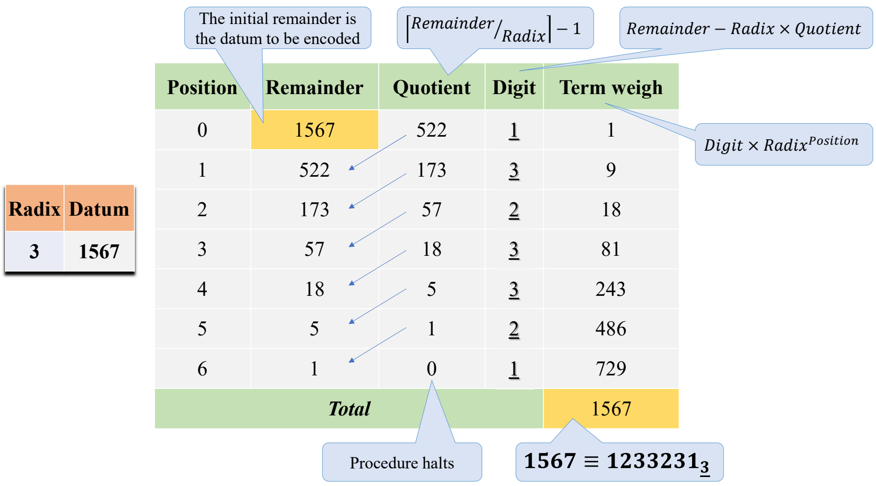

Figure 2. Data encoding in BN, which eludes zero as a symbol. Note that we use the ceiling function; specifically, we use $\lceil\rceil - 1$ instead of the floor function $\lfloor\rfloor$, which standard PN uses.

We can generalize the PMF given by (6) to the probability of getting a leading $\underline{r}$-ary numeral $n \in \mathbb{N}^+$ of any length. It is the likelihood gap it induces on the logarithmic scale.

Definition 8. *A leading numeral n is said to satisfy the local NBL for bijective $\underline{r}$-ary numeration if it occurs with probability*

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

Example 8. *The probability that a bijective decimal numeral starts with "2A1", say $2.A1_{\underline{10}}$ or $2A17_{\underline{10}}$, is* $\log_{11}\left(1+\frac{1}{301}\right) = 0.13832\,\%$.

Definition (8) allows us to derive the following PMF.

Definition 9. *A numerical dataset is said to satisfy the local NBL for BN if the probability of picking a length-l bijective $\underline{r}$-ary numeral starting with the digit d, where $1 \le d \le \underline{r}$ and $\{\underline{r}, l, d\} \in \mathbb{N}^+$, is*

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 March 2025 doi:10.20944/preprints202503.1625.v1

$$\Pr(\underline{r}, l, d) = \sum_{k=d\underline{r}^{l-1}+\frac{\underline{r}^{l-1}-1}{\underline{r}-1}}^{k=(d+1)\underline{r}^{l-1}+\frac{\underline{r}^{l-1}-1}{\underline{r}-1}-1} \log_{\frac{\underline{r}^{l+1}-1}{\underline{r}^{l}-1}}\left(1+\frac{1}{k}\right) \in \mathbb{R}$$

Example 9. *The probability of picking* 1 *to* 3 *as the first digit of a bijective ternary numeral with length* 5 *is* $\{0.465312, 0.306147, 0.228541\}$*, and the probability of choosing* 1 *to* A *as the first digit of a bijective decimal numeral with length* 2 *is* $\{0.2797, 0.1685, 0.1209, 0.09442, 0.07746, 0.06567, 0.057, 0.05036, 0.0451, 0.04084\}$.

Owing to (9) boils down to (6) if $l = 1$, the local NBL is length-invariant beside radix-invariant. Figure 3 shows the PMF of various bijective radices for consecutive lengths and the hyperbolic progression of the bijective ternary digits as the numeral's length expands.

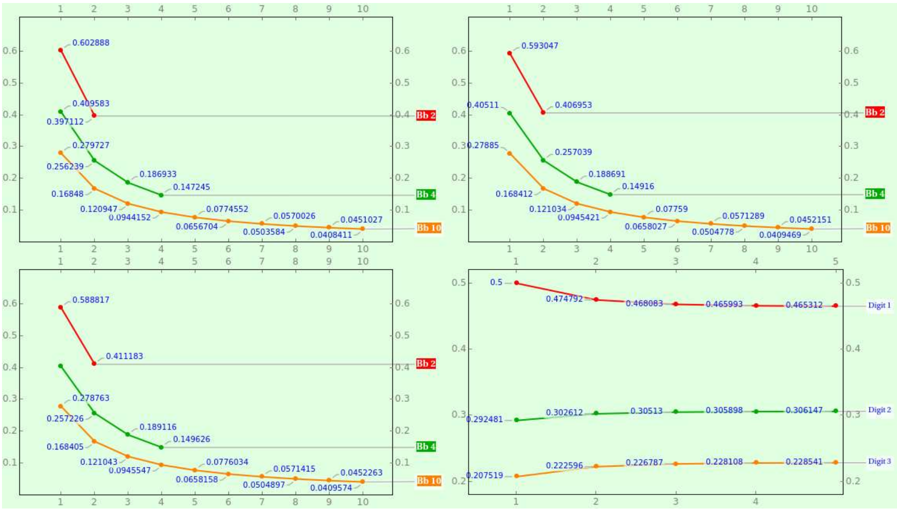

Figure 3. Leading digit's PMF for numerals in bijective radices $\underline{2}$, $\underline{4}$, and $\underline{A}$ (decimal) with lengths 2 (top-left), 3 (top-right), and 4 (bottom-left) using (9). On the bottom right, we show the probability plot of the bijective ternary digits as a function of the numeral's length; the probability mass gap between consecutive digits tends to stabilize.

Definition (8) also allows us to derive the following PMF.

Definition 10. *A numerical dataset is said to satisfy the local NBL for BN if the probability of runing into* d *as the* p*-th digit of a bijective* $\underline{r}$*-ary numeral, where* $1 \leq d \leq \underline{r}$ *and* $\{\underline{r}, d, p\} \in \mathbb{N}^{+}$*, is*

$$\Pr(\underline{r}, d, p) = \sum_{k=\frac{\underline{r}^{p-1}-1}{\underline{r}-1}}^{k=\frac{\underline{r}^{p}-1}{\underline{r}-1}-1} \log_{\frac{\underline{r}^{p+1}-1}{\underline{r}^{p}-1}}\left(1+\frac{1}{\underline{r}k+d}\right) \in \mathbb{R}$$

Example 10. *The chance of picking* 1 *to* 3 *as the fifth digit of a bijective ternary numeral is* $\{0.335006, 0.333327, 0.331667\}$*, and the probability of choosing* 1 *to* A *as the second digit of a bijective decimal numeral is* $\{0.1177, 0.1126, 0.1081, 0.1041, 0.1004, 0.09707, 0.09402, 0.09121, 0.08862, 0.0862\}$.

Owing to (10) boils down to (6) if $p = 1$, the local NBL is position-invariant beside radix-invariant.

4. Depleted and Constrained Harmonic Series

The global NBL for BN suddenly appears in the set of Kempner's curious series. We say a series is curious when the infinite summation of a harmonic series, divergent, is depleted by constraining its terms to satisfy specific convergence conditions. For example, consider the harmonic series missing the terms where "66" appears in their denominator. Most researchers in this fieldwork use decimal representation, but we can generalize the results to any base. Although their terminology refers to the items of a unit fraction's denominator as digits, for us, these are quanta of a chain because we are handling terms of a harmonic series.

The point is that most depletions result in an absolute mass because a harmonic series is on the verge of divergence. In particular, a harmonic series becomes convergent by omitting a single quantum. For example, the shrunk harmonic series without the terms in which "4" appears anywhere in the decimal representation of the denominator is K_4 of the Kempner series. Offhand, convergence comes up because we withdraw most of the terms; $1/10$ of the terms contain a "4" if the random variable ranges from 0 to 9, 20% have at least one "4" if the random variable ranges from 0 to 99, and eventually, most of the terms of any random chain with 100 quanta will contain at least one "4" and will not sum.

Notwithstanding, this explanation is misleading. The weight of the long chains containing a given quantum is lower than that of the short chains. A K_N series converges slowly [15] due to the relative and geometrically low contribution of large numerical chains containing N to the total.

Table 1 summarizes the outcomes of approximated calculations from 1 (K_1) to A $\equiv$ 10 (K_A). The most stunning feature of the Kempner summations (second column) is that they outline a curve that decreases harmonically.

Table 1. These are the absolute and relative masses of the Kempner series compared to NBL averaged over the first nine positions.

Decimal quantum (q)	Kempner K_q summations	Kempner M_q mass	NBL average weight W_q (9 positions)
1	16.1770	13.00	12.91
2	19.2573	10.92	10.95
3	20.5699	10.22	10.26
4	21.3275	9.86	9.89
5	21.8346	9.63	9.65
6	22.2056	9.47	9.48
7	22.4935	9.35	9.36
8	22.7264	9.25	9.25
9	22.9207	9.18	9.16
A	23.1034	9.10	9.09
Total	212.6158	100	100

Every quantum eliminates the same number of terms. $K_1 < K_2 < \cdots < K_A$ means not that "1" is in more terms than "2" or "3" but a heavier mass attributed to the terms with the minor quanta; if we take out $1/1$, the resulting summation is smaller than when we take out $1/2$ or $1/3$, and "A" is the quantum that contributes less to the total. Note also that although "A" is taken as "0" for calculation purposes, the value of K_A proves that BN is underneath.

Considering that a Kempner series is infinite and the set of Kempner series embraces all quanta q represented in bijective decimal, how could we find a better proof that a default probability potential outlines a hyperbolically decreasing function of q?

Since a curious series converges by default of unit fraction terms, the mass share of a quantum globally depends on the reciprocals of the Kempner summations; the third column of the table includes K_q's reciprocals normalized to 100 % (e.g., K_1's relative mass is $M_1 = \left(K_1 \sum_{q=1}^{A} 1/K_q\right)^{-1} \approx 13\%$). We must underline the relevance of these summations and percentages, reflecting the mass of every

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 March 2025 doi:10.20944/preprints202503.1625.v1

quantum irrespective of where it is, in contrast with the global NBL, which indicates the probability mass of a quantum at a given position in a given base.

We introduce two caveats to analyze the NBL weights (fourth column). First, the Kempner distribution conforms with NBL via the average of NBL distributions for different positions, which is NBL, too [16]. For instance, W_1 is, in principle, the average of quantum 1's probabilities at first (34.14%), second (11.89%), third (10.18%), fourth (10.01%), et cetera position according to (5). Second, because the distribution of the nth quantum quickly tends to be uniform (10% for each of the ten quanta from the fifth position), we must suspect that there exists a threshold position above which the contributions to the quantum's weight do not count; otherwise, the resulting mean distribution will end up reaching uniformity despite the differences that the Benford distribution makes at the first positions. Consequently, the last column calculates W_q as the NBL frequency averaged only over the first nine positions. Averaging ten positions also gives an excellent approximation (with a mean error of .091 %) to the distribution of Kempner masses, but nine positions deliver the minimal total mean error of .024 %.

Can we extrapolate this result in $\underline{b} = \underline{A}$ to any value of $\underline{b}$? If affirmative, PN would ignore a natural significand's quanta from the $\underline{b}$th place, agreeing with claims often made by mathematicians [17], physicists [18,19], and engineers [20] about the illogicality of a PN system carrying excessive digits in calculations of any type, regardless of the discipline.

We surmise that a bijective b-ary chain c that fulfills $\log_{\underline{b}} c > \underline{b}$ is physically elusive. The universe in base $\underline{b}$ would cope with at most $\underline{b}$ nesting levels, each distinguishing between b possible quanta. The "physical resolution"

$$\mathring{R}(\underline{b}) = \underline{b}^{\underline{b}}$$

would estimate the scope of quanta a computational system like the cosmos can naturally operate, much as a native resolution describes the number of pixels a screen can display.

In [21], the author contrives an efficient algorithm for summing a series of harmonic numbers whose denominator contains no occurrences of a particular numerical chain. As a result of the calculations, a harmonic series in base b omitting a chain of length n (regardless of its specific quanta) might converge approximately to

$$b^n \ln b.$$

This conjecture means that the contribution of linearly more extended chains to an endless series is geometrically lesser. For instance, the harmonic series where we impede the occurrence of the decimal numeral "314159" is about 2302582.334, whereas the same sum omitting "only" "3" is 22.921, 10^5 times as low. Thus, large numerical chains would be exponentially inconsequential.

More general constraints allow several occurrences of a given quantum to calculate summations positively. Let $S(n, q, b)$ be the sums of the b-ary reciprocals of naturals that have precisely n instances of the quantum q. For example, omitting the terms whose denominator in decimal representation contains one or more 6 is the particular case $S(0, 6, 10)$. The sequence of values S decreases and tends to

$$\lim_{n\to\infty} S(n, q, b) = b \ln b \approx \ln \mathring{R}(b)$$

regardless of q [22].

Except for the gap from $n = 0$ to $n = 1$, where the total increases, the summation falls as we raise the constraining quantity of quanta. What is the reason? It is not that we get more terms with n qs than terms containing $n + 1$ qs, but that the longer the chain, the lighter the contribution. Furthermore, when $n \gg 1$, $S(n, q, b) \gtrsim S(n + 1, q, b)$, whereas if $n \gtrsim 1$, $S(n, q, b) \gg S(n + 1, q, b)$, i.e., increments of n near the origin produce significant drops and vice versa, increments of n far from the origin produce negligible drops. Although we have not statistically tested "the number of quanta" for compliance with

NBL, we can again conclude that while small is a synonym for solid and discernible, huge numerical chains are fragile and hardly convey differences.

Instead of imposing absolute constraints, we can allow in a term arbitrarily many quanta q irrespective of the position and number so long as the proportion of qs remains below a fixed parameter $\lambda \in [0..1]$. In [23], the authors prove that the series converges if and only if $\lambda < 1/b$. In decimal, while Kempner's original series implies $\lambda = 0$, where no term containing a given quantum contributes to the summation, the complete harmonic series means $\lambda = 1$, where any density is allowed, i.e., we keep all the reciprocals.

For instance, if we consider the constraint "allow a rate of $\lambda = 5\,\%$ of 7s at most", the term $1/98765432109876543210$ disappears (10 % of 7s), but neither $1/98654321098$ (no 7s) nor $1/98865432109876543210$ (5 % of 7s) does. While the series converges in $\lambda \in [0..1/10)$, it no longer converges above the threshold $\lambda = 1/10$. Note that the archetype of the Pareto law appears naturally; on average, 90% of the unit fractions, those with the highest quantum density, offset the remaining 10%. Moreover, this result engages with our surmise concerning the physical resolution $\mathring{R}(b)$ of a universal computational system. Again, densities of b quanta or more are intractable. A PN system must restrict itself to chains with less than b quanta to guarantee the operability of coded data and avoid overflow conditions.

5. Conclusion

The canonical PMF is a probability inverse-square law that, taken as a brute fact, allows us to derive the global and local NBL. The canonical PMF, NBL subsidiaries, and PN run in parallel.

All the NBL formulas of standard PN are translatable to BN. For example, the standard and bijective decimal system global and local laws are similar but different, meaning that the precision of NBL is nonessential, while the supporting positional scale is what matters. In particular, the global NBL with standard base $\underline{b}+1$ corresponds to the global NBL with bijective base $\underline{b}$ (def. 1 and def. 3), and the local NBL with standard radix $\underline{r}+1$ corresponds to the local NBL with bijective radix $\underline{r}$ (def. 6 and def. 8).

In contrast with standard PN, the BN expressions of NBL have a very high level of generality, to the point that both are length- (def. 4 and def. 9) and position-invariant (def. 5 and def. 10), in addition to other well-known invariances. We also provide the formula for the harmonic entropy of the global NBL (def. 2), which is not evident, and the logarithmic entropy of the local NBL (def. 7).

We have proved that the Kempner distribution reflects the global version of NBL for BN and confirms a fundamental tendency toward minor numbers. The study of constrained harmonic series mainly teaches us that the specific digits involved in the restraining chain do not matter, whereas its length does. More generally, long chains or high densities of quanta are "rare" and deliver slender harmonic terms. In contrast, short chains or low quantum densities are regular and cheap, producing heavy harmonic terms that lead to convergence of the series if eliminated. In other words, only usual and economic constraints can impede the divergence of a harmonic series. More generally, positions with decreasingly lower exponents on a harmonic or logarithmic scale have exponentially less and less weight; a natural resolution in PN plausibly exists.

Regardless of the number system, we must conceive of positional scales as hyperbolic spaces in a broad sense, harmonic in the first place, and logarithmic in the second place. NBL correlates accessibility with smallness, proximity, scarcity, or brevity. The generality of the NBL expressions for BN reinforces the thesis that this law is comprehensively universal.

Funding: This research received no external funding.

Data Availability Statement: This article includes original contributions. Readers can contact the author for further inquiries. The programming code producing the data presented in this study is available on request.

Acknowledgments: We herein express our tribute to some personalities and entities influencing this work. The UNED incited a scientific instinct and a generalistic approach to problem-solving. Working for Indra enabled a pragmatic approach to figuring out any crisis and awareness of the productive character of our universe. We greatly thank José Mira for his capacity for motivation. We apologize in advance for omitting laudable references. We sincerely appreciate those offering comments and critical insight on this version.

Conflicts of Interest: The author (ORCID 0000-0003-3980-5829) declares no conflicts of interest. The author asserts that only scientific rigor, significance, and clarity drive the high-level goals of this work. It contains no known minor or significant incongruencies, 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, partially or in whole. 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 third-party rights such as copyright and moral rights.

References

1. Benford, F. The Law of Anomalous Numbers. *Proceedings of the American Philosophical Society* **1938**, *78*, 551–572.
2. Hill, T.P. A Statistical Derivation of the Significant-Digit Law. *Statistical Science* **1995**, *10*, 354–363.
3. Wikipedia contributors, the Free Encyclopedia. Positional Notation, 2022. [Online; accessed 11-November-2022].
4. Berger, A.; Hill, T.P. What is Benford's Law? *Notices of the AMS* **2017**, *64*, 132–134. https://doi.org/10.1090/noti1477.
5. Wikipedia contributors, the Free Encyclopedia. Bijective Numeration, 2020.
6. Rives, J. The Zero Delusion. *ResearchGate (Prepint)* **2023**, p. 60. Universe Intelligence, https://doi.org/10.13140/RG.2.2.20063.74400.
7. Foster, J.E. A Number System without a Zero-Symbol. *Mathematics Magazine* **1947**, *21*, 39–41.
8. Forslund, R.R. A Logical Alternative to the Existing Positional Number System. *Southwest Journal of Pure and Applied Mathematics* **1995**, *1*, 27–29.
9. Rives, J. The Code Underneath. *Axioms* **2025**, *14*, 106. https://doi.org/10.3390/axioms14020106.
10. Berger, A.; Hill, T. Benford's Law Strikes Back - No Simple Explanation in Sight for Mathematical Gem. *The Mathematical Intelligencer* **2011**, *33*, 85–91. https://doi.org/10.1007/s00283-010-9182-3.
11. Gisin, N. Indeterminism in Physics, Classical Chaos and Bohmian Mechanics: Are Real Numbers Really Real? *Erkenntnis* **2021**, *86*, 1469–1481. https://doi.org/10.1007/s10670-019-00165-8.
12. Kempner, A.J. A Curious Convergent Series. *The American Mathematical Monthly* **1914**, *21*, 48–50.
13. Wikipedia contributors, the Free Encyclopedia. Differential entropy, 2024. [Online; accessed 10-September-2024].
14. Bekenstein, J.D. A Universal Upper Bound on the Entropy to Energy Ratio for Bounded Systems. *Physical Review D* **1981**, *23*, 287–298. https://doi.org/10.1103/PhysRevD.23.287.
15. Baillie, R. Sums of Reciprocals of Integers Missing a Given Digit. *The American Mathematical Monthly* **1979**, *86*, 372–374.
16. Berger, A.; Hill, T.P. The Mathematics of Benford's Law: a Primer. *Statistical Methods and Applications* **2021**, *30*, 779–795. https://doi.org/10.1007/s10260-020-00532-8.
17. Wildberger, N.J. Finite versus Infinite and Number Systems. In *Sociology and Pure Mathematics*; Insights into Mathematics, YouTube Channel, 2022.
18. Dowek, G. Real Numbers, Chaos, and the Principle of a Bounded Density of Information. In Proceedings of the Theory and Applications. - Computer Science Symposium in Russia; A. A., B.; Shur, A.M., Eds., Springer, Berlin, Heidelberg, 2013; Vol. 7913, *LNTCS*, [978-3-642-38535-3].
19. Santo, F.D., Undecidability, Uncomputability, and Unpredictability; The Frontiers Collection, Springer International Publishing: Cham, 2021; chapter Indeterminism, Causality and Information: Has Physics Ever Been Deterministic?, pp. 63–79. https://doi.org/10.1007/978-3-030-70354-7_5.
20. Bailey, B. The Cost of Accuracy. *Semiconductor Engineering (Low Power - High Performance)* **2018**.
21. Schmelzer, T.; Baillie, R. Summing a Curious, Slowly Convergent Series. *The American Mathematical Monthly* **2008**, *115*, 525–540.
22. Farhi, B. A Curious Result Related to Kempner's Series. *The American Mathematical Monthly* **2008**, *115*, 933–938. https://doi.org/10.2307/27642640.

 doi:10.20944/preprints202503.1625.v1

23. Lubeck, B.; Ponomarenko, V. Subsums of the Harmonic Series. *The American Mathematical Monthly* **2018**, *125*, 351–355.