Applications of the first digit law to measure correlations

R. Gramm, J. Yost, Q. Su, and R. Grobe

Phys. Rev. E **95**, 042136 — Published 24 April 2017

DOI: 10.1103/PhysRevE.95.042136

Applications of the first digit law to measure correlations

R. Gramm, J. Yost, Q. Su and R. Grobe

Intense Laser Physics Theory Unit and Department of Physics
Illinois State University, Normal, IL 61790-4560 USA

The quasi-empirical Benford law predicts that the distribution of the first significant digit of random numbers obtained from mixed probability distributions is surprisingly meaningful and reveals some universal behavior. We generalize this finding to examine the joint first-digit probability of a pair of two random numbers and show that undetectable correlations by means of the usual covariance-based measure can be identified in the statistics of the corresponding first digits. We illustrate this new measure by analyzing the correlations and anti-correlations of the positions of two interacting particles in their quantum mechanical ground state. This suggests that by using this new measure, the presence or absence of correlations can be determined even if only the first digit of noisy experimental data can be measured accurately.

1. Introduction

There are unfortunately many situations where the experimentally unavoidable noise permits an accurate measurement to only a few leading digits of a quantum mechanical observable. The astounding significance and even universality of the distribution of the nonzero leftmost digit of "natural" data points has triggered a rather impressive number of recent studies [1-6]. These studies range from analyzing measurements in accounting-related [3], geographical [7], biological [8], chemical and physical [9,10] systems. These recent developments date back to earlier findings by Newcomb [11] in 1881 and Benford [12] in 1938, who discovered independently that the first digit (d=1,2,... 9) of almost any set of measured data does not occur with equal likelihood [i.e., $P(d) \neq 1/9$] as one might naively expect. In fact, it was observed that about five times as many numbers have d=1 as their first digit than those that start with d=9. This at first unexpected non-uniform distribution of the first digit is now believed to be so universal for most data sets, that deviations from this first digit distribution is being used routinely by state governments and other institutions to scan data for possible tax or accounting fraud [3] or the possibility of manipulation of voter data [13].

Due to its counter-intuitive nature, the law was first ignored for decades as a mere curious observation. However, the number of works analyzing the first-digit distribution increased more than exponentially, and now there are numerous archives available that catalog these publications [14]. Amusingly, an article [15] was published in 2016 that collected data on the number of citations received by those articles that cited the original papers of Newcomb and Benford. It was reported that the leading digits of these citations were distributed according to the very law they described.

The distribution of the first digit P(d) for any set of random variables x can be computed directly from the underlying probability density $\rho(x)$. For example, if the random variable x is uniformly distributed between $0<x<1$, then we obtain an equal likelihood $P(d) = 1/9$, as one would intuitively expect. However, we know now that this is a rather exceptional case, as most other densities $\rho(x)$ lead to a non-uniform first-digit probability P(d), where $P(d) > P(d+1)$.

During the last decade, it was conjectured that a rather specific first-digit distribution, given by the Benford logarithmic function $P_B(d) \equiv \log_{10}(1+1/d)$, has some remarkable universal features and is generic to the numbers from many real-world sources. In fact, the first digits of the set of numbers obtained from many general sequences, such as α^n (with $\log_{10} \alpha$ an irrational number and n=1,2, …), are precisely distributed according to $P_B(d)$.

Furthermore, if data from different distributions are superimposed its seems that the averages also approach $P_B(d)$. In 2000 Shao and Ma [16] examined three widely used distribution functions in quantum physics. They showed that the first digits of variables x distributed according to the Bose-Einstein density [$\rho(x) \sim (\mathrm{Exp}(\beta x)-1)^{-1}$] follow exactly the function $P_B(d)$, whereas the first digit probability for the Boltzmann-Gibbs [$\sim \mathrm{Exp}(-\beta x)$] as well as the Fermi-Dirac density [$\sim(\mathrm{Exp}(\beta x)+1)^{-1}$] fluctuate slightly around the Benford function in a periodic manner. In fact, when averaged over the logarithm of the inverse temperature $1/\beta$, even those corresponding first digit distributions converge to Benford's law. The authors also concluded that this law seems to present a general pattern for physical statistics and might be even more fundamental and profound in nature.

To the best of our knowledge, the vast literature on Benford's law has been focusing exclusively on the relationship between various probability densities of *single* random numbers on the distribution of the first few digits. The purpose of this note is four-fold. First, we aim at extending our understanding of the first digit law by examining how potential dependencies between *two* random numbers can manifest themselves in the correlations among the pairs of the corresponding first digits. Second, we suggest that using the joint first digit distribution (JFD) for two sets of random numbers can serve as a more robust indicator of possible correlations than other standard measures such as the usual Pearson product-moment coefficient based on the covariance. Third, by solving the Schrödinger equation for the ground state wave function for bound two-particle systems, we provide a concrete physical system where these correlations can occur due to the interactions between the particles. Fourth, we show analytically that while in general the first-digit distributions of powers of uniformly distributed random numbers are not Benford distributed, their distribution becomes logarithmic when averaged over the exponent. This is consistent with other attempts to provide evidence for the universality of the Benford law.

The article is structured as follows. In section 2 we introduce the concept of the joint first-digit distribution (JFD) and examine its predictions for Gaussian correlated data. We provide an example of power-law related random numbers and suggest that two new JFD-based parameters K and Ξ can provide a better measure of correlations than the traditional covariance-based correlation measure. In section 3 we provide a quantum mechanical example where the JFD can be used to detect the correlation of two interacting particles in their ground state wave function. In section 4 we provide a brief outlook on future challenges.

2. Transfer of correlations onto the joint first-digit distribution (JFD)

Mathematically, the distribution of the first digit P(d) for any set of random variables x can be easily computed from the underlying probability density ρ(x) of the random variable x as

$$P(d) = \Sigma_{n=-\infty}^{\infty} \; [\int_a^b dx\, \rho(x) + \int_c^e dx\, \rho(x)] \qquad (2.1)$$

where the integration limits extend from $a \equiv -(d+1)\,10^n$ to $b \equiv -d10^n$ to cover negative x and from $c \equiv d\,10^n$ to $e \equiv (d+1)10^n$ for positive x. P(d) is therefore the sum of the contributions due to positive and negative numbers x, i.e., $P(d) = P_{neg}(d) + P_{pos}(d)$.

To provide three brief examples, if the random variable x is uniformly distributed between 0<x<1, [i.e. ρ(x)= θ(x,0,1), where θ(x,a,b) is the generalized unit-step function that is θ=1 if a<x<b and θ=0 otherwise], then we obtain an equal likelihood $P(d) = \Sigma_{n=-\infty}^{\infty} \int dx\, \theta(x,0,1) = \Sigma_{n=-\infty}^{-1} \; 10^n = 1/9$, as one would intuitively expect. If the random variable is standard normal distributed, $\rho(x) = (2\pi)^{-1/2}\exp(-x^2/2)$, we obtain, for example, P(1)≈0.36, P(2)≈0.13 and P(9)≈0.06. Our third example is discussed in the Appendix A where we examine the relationship between the first digit distribution of random numbers $y=x^{\alpha}$, obtained from uniformly distributed variables x and the Benford law $P_B(d) \equiv \log_{10}(1+1/d)$. We show that this law is confirmed in the limit for the exponent $\alpha \rightarrow \pm\infty$. It is also obtained when we average the probabilities P(d) over all exponents α.

Next we will generalize this to the case of two random numbers x and y. The corresponding joint probabilities of the pair of first digits $P(d_x,d_y)$ ($d_x = 1,2, \ldots 9$ and $d_y = 1,2, \ldots 9$) can be obtained from the original joint density ρ(x,y)

$$P(d_x,d_y) = \Sigma_{nx=-\infty}^{\infty} \Sigma_{ny=-\infty}^{\infty} \; [\int_{ax}^{bx} dx \int_{ay}^{by} dy\, \rho(x,y) + \\ + \int_{ax}^{bx} dx \int_{cy}^{ey} dy\, \rho(x,y) + \int_{cx}^{ex} dx \int_{ay}^{by} dy\, \rho(x,y) + \int_{cx}^{ex} dx \int_{cy}^{ey} dy\, \rho(x,y)] \qquad (2.3)$$

where the integration limits are again from $ax \equiv -(d_x+1)\,10^{nx}$ to $bx \equiv -d_x10^{nx}$ and from $cx \equiv d_x\,10^{nx}$ to $e_x \equiv (d_x+1)10^{nx}$ and similarly for y, $ay \equiv -(d_y+1)\,10^{ny}$ to $by \equiv -d_y10^{ny}$ and from $cy \equiv d_y\,10^{ny}$ to $e_y \equiv (d_y+1)10^{ny}$. We will abbreviate the *j*oint *f*irst-digit *d*istribution $P(d_x,d_y)$ as JFD from now on.

There are two limiting cases for $P(d_x,d_y)$ that are of interest. It is obvious that if the density $\rho(x,y)$ describes two random numbers that are completely uncorrelated, i.e., $\rho(x,y) = \rho_x(x)\, \rho_y(y)$, then the corresponding 9×9 matrix $P(d_x,d_y)$ is simply given by the direct product of the corresponding two individual first digit probabilities associated with each random variable, $P(d_x,d_y) = P_x(d_x)\, P_y(d_y)$, where $P_x(d) = P_{xneg}(d) + P_{xpos}(d)$ and similarly for $P_y(d)$. On the other hand, if the two random variables are absolutely correlated, for example as in $\rho(x,y) = \rho_x(x)\, \delta(x-y)$, then the joint probability becomes a diagonal 9×9 matrix of the form $P(d_x,d_y) = \delta_{dx,dy} \Sigma_{n=-\infty}^{\infty} [\int_a^b dx\, \rho_x(x) + \int_c^e dx\, \rho_x(x)]$.

There are also two illustrative special cases where each variable is uniformly distributed, $\rho_x(x)= \theta(x,0,1)$ and similarly for y. If x and y are uncorrelated, i.e. $\rho(x,y) = \theta(x,0,1)\theta(y,0,1)$ we obtain $P(d_x,d_y) = 1/81$ and for the fully correlated case, i.e. $\rho(x,y) = \theta(x,0,1)\, \delta(x-y)$, we obtain $P(d_x,d_y) = \delta_{dx,dy}\, 1/9$.

In order to study more general cases of random numbers that are neither fully correlated nor completely independent, we have to employ numerical techniques. As a test joint density with variable degree of correlation W^{-1} we have examined the density $\rho(x,y;W) = N\, Exp[-(x-y)^2/(2W^2)]\, \theta(x,0,1)\theta(y,0,1)$, where N is the corresponding normalization factor, $N^{-1} \equiv \int_0^1 \int_0^1 dxdy\, Exp[-(x-y)^2/(2W^2)]$. Obviously, we have the two limits $\rho(x,y;W\to 0) = \delta(x-y)\, \theta(x,0,1)$ and $\rho(x,y;W\to\infty) = \theta(x,0,1)\theta(y,0,1)$.

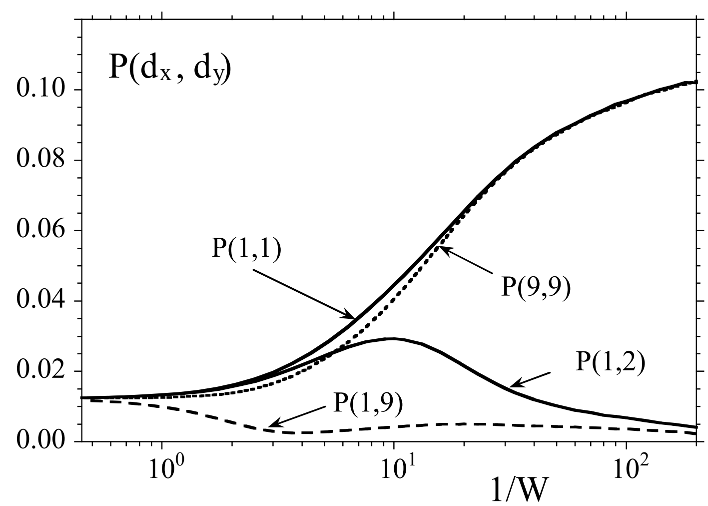

Figure 1 The joint first digit (JFD) probability $P(d_x,d_y;W)$ obtained from pairs of random numbers with a density $\rho(x,y;W) = N\, Exp[-(x-y)^2/(2W^2)]\, \theta(x,0,1)\theta(y,0,1)$ as a function of the correlation W^{-1} for the four pairs (d_x=1, d_y=1), (d_x=9, d_y=9), (d_x=1, d_y=2) and (d_x=1, d_y=9).

In Figure 1 we display $P(d_x,d_y;W)$ for the four pairs ($d_x=1$, $d_y=1$), ($d_x=9$, $d_y=9$), ($d_x=1$, $d_y=9$) and ($d_x=1$, $d_y=2$) as a function of W^{-1}. In order to confirm the accuracy of the four probabilities we compared the results directly obtained from the double integration and the two infinite summations based on Eq. (2.1) with the counting data obtained from 100000 pairs of random numbers (x,y). These can be easily obtained from $\rho(x,y;W)$ by employing a generalized rejection method. In this numerical technique the two uniformly distributed random numbers ξ_1 and ξ_2 become x and y, if the density $\rho(\xi_1,\xi_2;W) > \xi_3$, where ξ_3 is a third uniformly distributed random number. Even though in this method we can map three random numbers (ξ_1,ξ_2,ξ_3) into at most two random numbers (x,y), it is still rather efficient as uniformly distributed random numbers can be rather easily generated on a computer based on linear-congruence techniques [17].

The four curves in Figure 1 show that the original degree of correlation W^{-1} has a direct impact on the correlation among the corresponding first digits. As the correlation W^{-1} increases, $P(1,1;W)$ increases from $P(1,1;\infty) = 1/81$ [corresponding to $P(1)=1/9$] to $P(1,1;0) = 1/9$, whereas the off-diagonal probability for two different first digits decreases from $P(1,9;\infty) = 1/81$ to $P(1,9;0) = 0$. In other words, the 9×9 matrix $P(d_x,d_y;W)$ becomes diagonal as the correlation grows. The transition occurs close to W=0.5, where the corresponding half width of the Gaussian joint density becomes less than the domain from $0<x<1$ and $0<y<1$. Comparing the different behavior of the two off-diagonals P(1,2) and P(1,9) suggests an interesting non-monotonic decrease with increasing W^{-1}.

2.2 The JFD-based measures of correlation K and Ξ

In order to have an unambiguous and direct quantitative measure for the degree of correlations contained in the JFD, we employ the canonical basis state decomposition [18,19] on the joint probability $P(d_x,d_y)$. If the first digits are entirely uncorrelated, then it can be decomposed into a direct product as $P(d_x,d_y) = p_x(d_x)\, p_y(d_y)$, where the subscripts x and y are necessary as in general $p_x(d) \neq p_y(d)$. The definition for correlation employed here is based on the assumption that the more products are required to express the joint density, the larger is the degree of correlation:

$$P(d_x,d_y) \;=\; \Sigma_{n=1}^{9}\, w_n\, p_x(d_x;n)\, p_y(d_y;n) \qquad (2.4)$$

Here we assume that the marginal probabilities are normalized $\Sigma_{d=1}^{9}\, p_x(d;n) = 1$ and $\Sigma_{d=1}^{9}\, p_y(d;n) = 1$ and that they are orthogonal with respect to each other, $\Sigma_{d=1}^{9}\, p_x(d_x;n)\, p_y(d_y;m) = 0$ if $n\neq m$. While we obviously have the total probability $\Sigma_{n=1}^{9}\, w_n = 1$, not all of the coefficients w_n have to be nonzero. In fact, if the digits are uncorrelated, we have $w_1 = 1$ while all other w_n (n=2, 3,... 9) vanish.

We identify here the amount of correlation with the effective number of different products $p_x(d_x;n)\, p_y(d_y;n)$ that is necessary to construct the exact joint probability density $P(d_x,d_y)$. A similar measure was introduced in 1994 [20] to provide a global quantitative and non-operator-specific measure of correlations in multi-particle quantum systems [21]. The average weight would be computed as $\Sigma_{n=1}^{9}\, w_n\, w_n$. The inverse of this weight would therefore correspond to the number of effectively non-zero weights. So we define here the correlation K by the effective number of non-vanishing weight factors, $K(d_x,d_y) \equiv [\Sigma_{n=1}^{9}\, w_n^{2}]^{-1}$. If the probabilities $P(d_x,d_y)$ are perfectly uncorrelated, we have K=1. The highest possible degree of correlation would be K=9, meaning that $P(d_x,d_y)$ can only be expressed as the sum over 9 products $p_x(d_x;n)\, p_y(d_y;n)$ and $w_n=1/9$.

As we show in the Appendix B, it is possible numerically, to construct uniquely the canonical probabilities $p_x(d;n)$ and $p_y(d;n)$ and the corresponding weight factors w_n from an arbitrary 9×9 joint probability matrix $P(d_x,d_y)$. By using the spectral information of two auxiliary positive definite matrices, we can calculate the degree of correlation K.

As an alternative second measure of correlation of the 81 elements of the JFD, we can also introduce the quantity

$$\Xi(x,y) \equiv (81/8)\, \Sigma_{dx,dy}\, [P(d_x,d_y) - P(d_x)P(d_y)]^2 \qquad (2.5)$$

where $P(d_x) \equiv \Sigma_{dy=1}^{9}\, P(d_x,d_y)$ and $P(d_y) \equiv \Sigma_{dx=1}^{9}\, P(d_x,d_y)$. This correlation Ξ obviously vanishes only if all possible first digit-pairs are completely uncorrelated to each other. For example, in the Gaussian example discussed above in Sec. 2.1 for the fully correlated case (W=0) we obtain $\Xi = 1$ and for no-correlations ($W=\infty$) we find $\Xi = 0$

2.3 Comparison of the JFD with the usual Pearson correlation coefficient χ

There are, of course, numerous expectation value based definitions to quantify the degree of correlations between two random numbers x and y. While the general concept of correlation is related to the degree of dependence of two random variables, a linear dependence is just a particular kind of correlation. Possibly the most common measure for this is the Pearson product-moment correlation coefficient χ based on the ratio of the covariance of the two variables and the product of their standard deviations, i.e.

$$\chi(x,y) \equiv \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle \,/\, (\Delta x \; \Delta y) \qquad (2.6)$$

where $\Delta x \equiv \langle (x - \langle x \rangle)^2 \rangle)^{1/2}$ is the standard deviation, and similarly for Δy. This coefficient $[-1 \leq \chi \leq 1]$ is zero if the variables are linearly independent and $\chi = 1$ (-1) if they are perfectly correlated (anti-correlated).

However, as χ quantifies the degree of linear dependence, it can be zero even if y is *completely* determined by x. To provide a concrete example, the pairs x and $y=x^2$ lead to $\chi=0$ if $\rho(x) = \rho(-x)$ as $\chi(x,x^2) = \langle x(x^2 - \langle x^2 \rangle) \rangle / (\Delta x \; \Delta y) = 0$. We suggest that using the JFD for two sets of random numbers can possibly serve as a more robust indicator of correlations than the Pearson product-moment coefficient based on the covariance. We will illustrate this using an example of two random numbers that are direct functions of each other and the first one is uniformly distributed for $-1<x<1$.

For example, for the power-law relationship $y = |x|^{\alpha}$ the Pearson coefficient χ is not able to reflect the α-dependent relationship between x and y as $\chi(x,y)$ vanishes for any exponent α. Here the numerator in $\chi(x,|x|^{\alpha})$ is given by $\langle x \, (|x|^{\alpha} - (\alpha+1)^{-1}) \rangle = 0$. However, in contrast to χ, the JFD and K and Ξ are capable of recovering the correlation as a function of α. In this case, the joint probability density is given by $\rho(x,y) = \theta(x,-1,1)/2 \; \delta(y-|x|^{\alpha})$. For $\alpha=1$ we find that $P(d_x,d_y) = \delta_{dx,dy} \; 1/9$ is diagonal, corresponding to a maximal correlation degree K=9 and also $\Xi=1$. Furthermore, these two parameters are capable of sensing rather small changes in α. For example, for α close to 1.1 Ξ has decreased by 50% to $\Xi=0.5$. In the opposite limit of large α, the JFD becomes completely uncorrelated as it approaches the fully factorized form $P(d_x,d_y) = 1/9 \log_{10}(1+1/d_y)$ and correspondingly K=1 and $\Xi=0$. This logarithmic dependence is fully consistent with the Benford-like limit as we derived in Appendix A for power-laws. This decrease of the correlation in $P(d_x,d_y)$ is expected as the first digit of x and $|x|^{\alpha}$ becomes completely independent of each other in this limit.

In Figure 2 we display the overall decay of the JFD correlation Ξ as a function of both positive as well as negative exponents α. It is interesting to note that the overall decreasing behavior of Ξ with |α| is accompanied with small maxima for integer values of the exponents α. These maxima might reflect some regularities due to symmetries that lead to the vanishing of $P(d_x,d_y)$ for some specific pairs of d_x and d_y. For example, for α =2, only 24 of the 81 matrix elements for $P(d_x,d_y)$ are non-zero. Here $P(d_x=9,d_y)$ is only nonzero for $d_y=8$ or $d_y=9$. This corresponds to the fact that when we square any number that starts with 9, then the mantissa of the product has to begin with either 8 or 9.

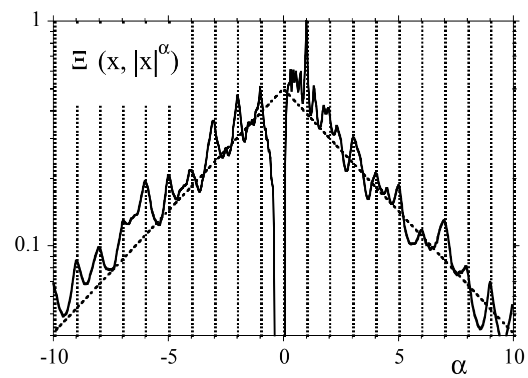

Figure 2 The correlation Ξ of the joint first digit (JFD) probability $P(d_x,d_y)$ obtained from pairs of random numbers x and $y=|x|^{\alpha}$, where $-1<x<1$ is uniformly distributed as a function of the exponent α. Here the corresponding Pearson correlation coefficient χ is inappropriate to detect a nonlinear correlation as it vanishes for any α. For comparison, the dashed line is $\Xi(\alpha) = \mathrm{Exp}(-|\alpha|/4)/2$.

On average, the overall decreasing behavior of the correlation Ξ among the first digits seems to follow for about |α|>2 an exponential decay law. To illustrate this, we have accompanied the data with the function $\Xi(\alpha) = 0.5\ \mathrm{Exp}(-|\alpha|/4)$. We mention that it is not uncommon that a power law-like relationship between two random variables can lead to exponential decays of their correlations.

It is also of interest to note the unusual small window for negative values of α close to zero, where the correlation Ξ seems to vanish entirely. This can be easily understood if we consider that for α=0, all variables y take the same value y=1 and therefore $P(d_x,d_y) = 1/9\ \delta_{dy,1}$ such that $P(d_x) = 1/9$ and $P(d_y) = \delta_{dy,1}$. This observation is also directly related to the case discussed in Appendix A and the discontinuities of the graphs for P(d=1) and P(d=9) in Figure A.

3. Joint first-digit distributions for correlated two-particle ground state wave functions

In this section, we will examine the JFD together with some other measures of correlations for a pair of random variables x and y that follow a bivariate normal distribution given by the general form

$$\rho(x,y) = [2\pi\sigma_x\sigma_y(1-\rho^2)^{1/2}]^{-1} \ \exp\{ -[x^2/\sigma_x^2 + y^2/\sigma_y^2 - 2\rho xy/(\sigma_x\sigma_y)] / [2(1-\rho^2)] \} \qquad (3.1)$$

This kind of distribution is rather generic and emerges in many different physical contexts. For example, this distribution can be realized by the density of the lowest energetic stationary eigenstate of two particles, that are bound by an attractive external harmonic oscillator and which interact with each other through an attractive or repulsive force. This would provide us with a concrete physical realization for the correlation transfer from x and y to the JFD, where the origin of the correlation is directly related to the magnitude of the inter-particle force. The quantum Hamiltonian of the two particles in a harmonic oscillator that are mutually coupled via another quadratic binding force of strength γ is then given by

$$H(\gamma) = p_x^2/2 + p_y^2/2 + x^2/2 + y^2/2 + \gamma\,(x-y)^2/2 \qquad (3.2a)$$

For $\gamma>0$, the two particles attract each other, whereas for $\gamma<0$, we can simulate the effect of mutual repulsion, such as between two particles of equal charge. However, in the latter case a bound system is only possible as long as the repulsion does not exceed the individual binding of each particle by the common quadratic potential, requiring $\gamma > -\frac{1}{2}$. In this limiting case, the total potential $V(x,y,\gamma = -\frac{1}{2}) = \frac{1}{4}(x+y)^2$, making the motion in the relative coordinate x–y entirely force-free. So, the overall permitted range to study two mutually bound particles is $-\frac{1}{2} < \gamma < \infty$. In the absence of any interaction ($\gamma=0$) we obtain as the lowest eigenvalue $E(\gamma=0) = 1$ and the simple uncorrelated product $\Psi(x,y;\gamma) \sim \exp(-x^2/2)\exp(-y^2/2)$ for the ground state wave function defined as $H(\gamma)\,\Psi(x,y;\gamma) = E(\gamma)\,\Psi(x,y;\gamma)$.

To construct analytically the ground state for the interacting system ($\gamma\neq0$), we can introduce the relative and center of mass coordinates $R \equiv (x+y)/2$, $r \equiv x-y$, $P \equiv p_x+p_y$ and $p \equiv (p_x-p_y)/2$ that can decouple the Hamiltonian into

$$H(\gamma) = \tfrac{1}{2}[P^2/2 + 4\,R^2/2] + 2\,[p^2/2 + (2\gamma+1)\,r^2/8] \qquad (3.2b)$$

which takes the energy $E(\gamma) = ½\, \kappa_+$ with $\kappa_+ \equiv ((2\gamma+1)^{1/2}+1)$ and $\kappa_- \equiv ((2\gamma+1)^{1/2}-1)$. The ground state wave function is therefore given by

$$\begin{aligned}\Psi(x,y;\gamma) &= N \exp[-R^2] \exp[-(2\gamma+1)^{1/2} r^2/4] \\ &= N \exp[-\kappa_+ (x^2+y^2)/4 + \kappa_- xy/2] \qquad (3.3)\end{aligned}$$

The normalization factor N can be obtained from the normalization condition $N^{-2} = \iint dx\, dy\, |\Psi(x,y;\gamma)|^2 = 2\pi\, [\kappa_+^2 - \kappa_-^2]^{-1/2}$.

While the wave function is deterministic, we can rely on the usual statistical (Born) interpretation of the wave function, which is directly related to the distribution of the measured (random) positions of the two particles found in consecutive realizations of the same state. In this sense do the different realization come from the random draws of the initial conditions characteristic of the ground state wave function.

Next, the density $\rho(x,y;\gamma) = N^2 \exp[-\kappa_+ (x^2+y^2)/2 + \kappa_- xy]$ can be converted into the corresponding joint first digit probability $P(d_x,d_y;\gamma)$. Similarly as in Section 2, we can either simulate random numbers that are distributed according to $\rho(x,y;\gamma)$ and then count the fraction of pairs according to their first digits, or determine $P(d_x,d_y;\gamma)$ numerically according to Eq. (2.3).

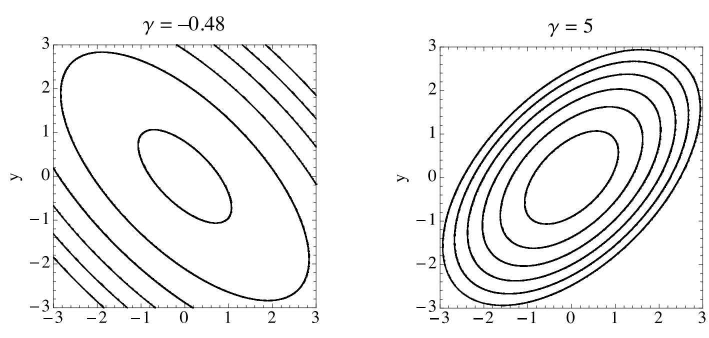

Figure 3 The contour plots spatial probability density for the two particles in the ground state for attractive (left) and repulsive (right) interactions. For both densities, the six contour levels were chosen at $\rho(x,y) = 10^{-n}$ with n=-6,-5, ...-1).

To show the direct impact of the inter-particle force γ on the spatial probability density $\rho(x,y;\gamma)$, we graph its contour plot in Figure 3 for $\gamma = -0.48$ (repulsion) and $\gamma = 5$ (attraction). While in the absence of any forces the distribution is radially symmetric around the origin, in the repulsive regime the Gaussian is deformed and peaked along the x=-y diagonal. In the attractive for regime, the particles positions are more likely to match each other thus favoring the diagonal x=y. Naturally, the forces have also a significant impact on the uncertainty of the position of both particles, which decreases monotonically from $\Delta x = \infty$ to $\Delta x = 1/2$ as γ increases from -1/2 to ∞. As a reference, in the absence of any interaction (γ=0) the spatial width $\Delta x \equiv \langle x^2 \rangle^{1/2}$ is $2^{-1/2} = 0.71$. For γ=-0.48 we measured it as Δx =1.22 whereas for γ =10 it is Δx = 0.55.

In Figure 4, we compare the covariance-based quantity χ and the JFD-based parameters K and Ξ as measures for the two-particle correlation as a function of the force-strength γ. In the small parameter domain ($-1/2<\gamma<0$) where the force is repulsive, we find consistently a negative χ showing the anti-correlation, while the parameters K and Ξ are positive. In the limit of no interaction γ=0, the two particles are independent of each other and we consistently find Ξ=0, χ=0 and K=1. As the force is increased, we observe the growth of all three measures. While χ increases rather rapidly for small forces γ as it approaches its maximum value (=1), (for example χ=0.5 for just γ=4.06, but Ξ is only 5.4×10^{-4}), the other two parameters Ξ and K increase more slowly. Once χ is within a few percent of its maximum value 1, Ξ and K still grow and are therefore more suited to detect changes in γ for the larger force regime.

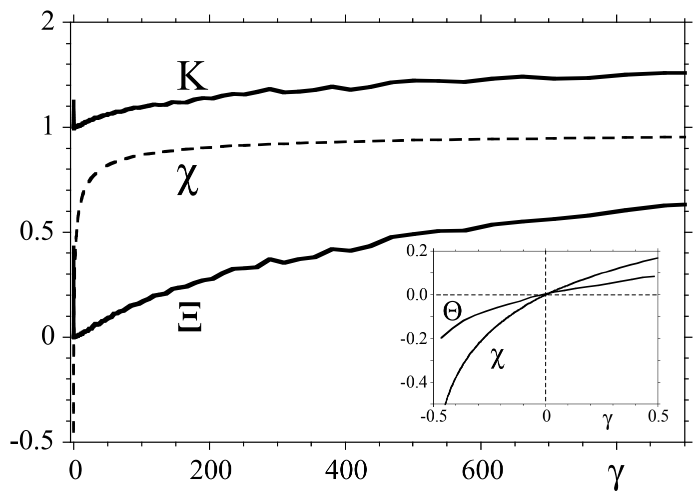

Figure 4 Three measures of correlations χ, Ξ and K between two particles in the ground state wave function of a harmonic oscillator as a function of the inter-particle force strength γ. For better graphical clarity, the parameter Ξ was multiplied by a factor of 10. In the inset we display the parameter Θ (defined in the text) in the short transition range for $-0.49<\gamma<3$.

Finally, let us examine if also the JFD is capable of distinguishing between the repulsive (anti-correlated) and the attractive (correlated) force regime. We could try to follow the similar line of thought that led to the definition of the Pearson coefficient χ of Eq. (2.5). We could use the average deviation of the first digits from their expectation value as a measure. This would lead to a quantity $\Theta \equiv \langle(d_x - \langle d_x\rangle)(d_y - \langle d_y\rangle)\rangle / (\Delta d_x \, \Delta d_y)$ with the standard deviation $\Delta d_x \equiv \langle(d_x - \langle d_x\rangle)^2\rangle)^{1/2}$ and similarly for Δd_y. However, it turns out that the difference between correlated and anti-correlated pairs (x,y) is often manifest in the sign of each random variable. Two positions x and y that have opposite signs (due to the particles' repulsion) could still both have the same (positive) first digit. In other words, as long as we do not take the signs into account, any difference between anti-correlation and correlation might be difficult to be observed with Θ. In fact, for the range from $-0.49<\gamma<0.49$, Θ remains positive.

However, if we permit the first digits d_x and d_y in the definition of Θ to carry the sign of their respective variables x and y, then this modified parameter based on $P(d_x,d_y)$ can discriminate between the attractive and repulsive force region. We have illustrated this in the inset to Figure 4, where we have displayed the "correlation" Θ in the range for $-0.49<\gamma<3$. For large values of the attraction γ, Θ approaches 1, as basically both particles prefer identical positions, i.e., x=y, such that $\Theta \rightarrow \langle(d_x - \langle d_x\rangle)(d_x - \langle d_x\rangle)\rangle / (\Delta d_x \, \Delta d_x) = 1$.

4. Summary

While there are numerous investigations on the statistics of the first digits for single random numbers, this is the first study that suggests that the concept of the first digit can be applied also to a pair of two (possibly correlated) random numbers in a meaningful way. We have examined some examples that suggest that the corresponding joint first-digit distribution (JFD) can serve as a potentially more generally applicable measure of possible correlations between two random numbers than the usual Pearson correlation coefficient that solely quantifies the degree of linear dependence. We have also provided a concrete physical example from quantum mechanical bound states where the correlation was provided by either attractive or repulsive forces between two particles. Here the JFD can be used to measure the degree of correlation.

The mapping of a number onto its first digit is for us a very non-trivial and certainly non-linear map that involves naturally a large amount of information loss. This is similar to binning used often as a data pre-processing technique, where the original data that fall into a certain small interval are replaced by an integer value that is representative of this interval. This aggregation may reduce the impact of noise on the processed data, but this mapping usually respects the magnitude of these numbers. In the case of first digits, even numbers that differ by several orders of magnitude in size can be counted towards the same "bin" out of a maximum of only nine bins. Despite these complexities, it is remarkable to us that nevertheless some properties such as correlations seem to be preserved to a certain degree.

This work obviously opens the door to many further investigations. For example, while the first digit statistics for single random number has revealed some rather universal behavior (as indicated by the quasi-empirical Benford law), it would be interesting to examine whether similar universal structures can also be detected for the joint-first digit distribution of pairs of random numbers.

Our concrete quantum mechanical illustration of the JFD was restricted to the stationary state, where we could show that the degree of the correlation among the two positions depends on the magnitude of the (either attractive or repulsive) force between them as characterized by the coupling parameter. We would expect that this relationship to be even valid if the system was initiated in a completely uncorrelated state and would be allowed to develop nontrivially in time. Quite likely, the correlation as measured by the JFD would then build up in time as the two particles evolve.

In the case of the usual Benford law for single random numbers, it turns out that there are even (weak) correlations between the consecutive digits [2], as the joint probability to find d_1 as the first and d_2 as the second digit is not simply given by the product of the individual probabilities. Following the same idea, it might be interesting to examine if and how correlations between two random numbers can affect these correlations among consecutive digits.

Acknowledgements

We would like to thank Profs. Ansher, Christensen, Harris, Martin and Rutherford for helpful discussions. This work has been supported by the NSF and the NSFC (#11529402).

Appendix A

Here we examine the first digit probability for random numbers y that are given as powers from uniformly distributed (positive) random number x, i.e. $y = x^{\alpha}$, where the exponent covers the entire range $-\infty<\alpha<\infty$. As a first step, we have to obtain the density $\rho(y;\alpha)$ from the transformation law for probability densities, $\rho(y)\,dy = \rho(x)\,dx$. Using $dy/dx = \alpha\, x^{\alpha-1}$ and $\rho(y) = \rho(x)\,|dx/dy|$, we obtain for positive exponents $\rho_{>}(y;\alpha) = \alpha^{-1}\, y^{1/\alpha-1}$ with $0\leq y\leq 1$, and for $\alpha<0$ we have $\rho_{<}(y;\alpha) = |\alpha^{-1}|\, y^{1/\alpha-1}$ with $1\leq y<\infty$. Using $\int dy\, y^{1/\alpha-1} = \alpha\, y^{1/\alpha}$, one can easily check the required normalization $\int_0^1 dy\, \alpha^{-1}\, y^{1/\alpha-1} = 1$ for $0<\alpha$ and $\int_1^{\infty} dy\, |\alpha^{-1}|\, y^{1/\alpha-1} = 1$ for $\alpha<0$. Next we calculate

$$
\begin{aligned}
P(d;\alpha) &= \Sigma_{n=-\infty}^{\infty} \int_c^e dy\, \rho_{<,>}(y;\alpha) \\
&= \Sigma_{n=-\infty}^{-1} [(d+1)^{1/\alpha}\, 10^{n/\alpha} - d^{1/\alpha}\, 10^{n/\alpha}] \\
&= [(d+1)^{1/\alpha} - d^{1/\alpha}]\, \Sigma_{n=1}^{\infty}\, (10^{-1/\alpha})^{n}] \\
&= [(d+1)^{1/\alpha} - d^{1/\alpha}]\, (10^{1/\alpha} - 1)^{-1} \qquad \text{(A1)}
\end{aligned}
$$

where $c \equiv d\, 10^{n}$ to $e \equiv (d+1)10^{n}$. A similar derivation leads to the same functional form for $\alpha<0$. One can also test this expression and easily confirm numerically the required normalization $\Sigma_{d=1}^{9}\, P(d;\alpha) = 1$ for any positive or negative α.

In Figure A we graph $P(d;\alpha)$ for d=1,2 and 9 as a function of the exponent α. For $\alpha=1$ we recover of course $P(d;\alpha=1) = 1/9$ as characteristic of the original uniformly distributed random number, but the likelihood of starting for the mantissa of y to begin with d=1 increases as α increases. At the same time, as α increases from 1 the probability $P(d=9;\alpha=1)$ decreases. In fact, the expression $P(d;\alpha) = [(d+1)^{1/\alpha} - d^{1/\alpha}]/(10^{1/\alpha} - 1)$ simplifies for both limits $\alpha\rightarrow\pm\infty$ exactly to the Benford form $P(d;\pm\infty) = \log_{10}(1+1/d)$, as $A \rightarrow \infty$. Because of this limiting behavior, the average probability, defined as $P(d) = \lim_{(A\rightarrow\infty)} (2A)^{-1} \int_{-A}^{A} d\alpha\, P(d;\alpha)$ naturally approaches the Benford function, $P(d) = \log_{10}(1+1/d)$.

We also see an interesting discontinuity of $P(d;\alpha)$ for d=1 and d=9 at $\alpha=0$. In this special case, all of the random numbers y are very close to 1, $y=x^{0}=1$, and therefore for an infinitesimally small positive ε, we have $x^{-\varepsilon} \approx 1.00...$ and $x^{\varepsilon} \approx 0.99...$ This leads to the discontinuity, $P(d=1;\alpha=-\varepsilon)=1$ and $P(d=1;\alpha=\varepsilon)=0$ as α changes its sign and correspondingly $P(d=9;\alpha=-\varepsilon)=0$ and $P(d=9;\alpha=\varepsilon)=1$. At this transition, all other P(d) vanish for d=2,3, ...8.

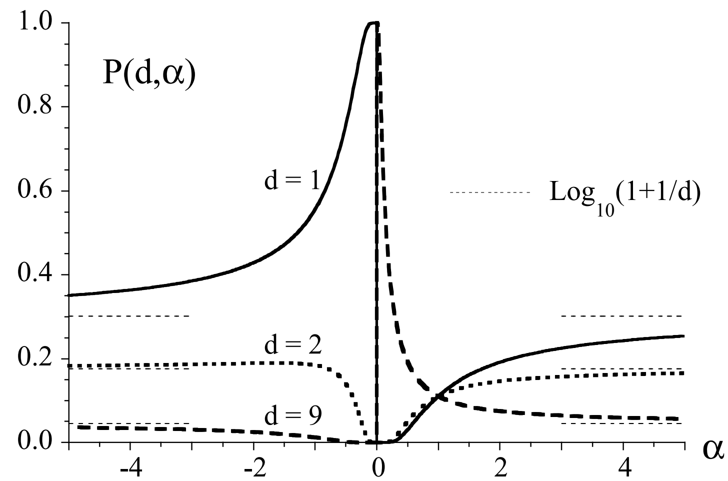

Figure A The probabilities P(d,α) that the first digit (d=1,2 or 9) of the random number y is d=1, 2 or d=9. The random number y is obtained from a uniformly distributed random number x as $y=x^{\alpha}$. For α=1 (uniformly distributed random numbers) all digits are equally likely, P(d)=1/9. For comparison, the three dashed lines are the probabilities $\mathrm{Log}_{10}(1+1/d)$ according to Benford's law.

Quite interestingly, one could alternatively try to perform the average over α on the level of the original density ρ(y;α), and try to construct an (α-averaged) density. For simplicity, we focus here on this average density for 0<y<1 defined as

$$\rho(y) = \lim_{(A\to\infty)} (A)^{-1} \int_0^A d\alpha\, \rho_{>}(y;\alpha) \qquad \text{(A2)}$$

This evaluation of ρ(y) is trickier due to the required normalization $\int_0^\infty dy\, \rho(y) = 1$. We require this property as it was true for ρ(y;α) for any value of α. By inserting $\rho_{>}(y;\alpha)$ into Eq. (A2), we obtain $\rho(y) = 1/y\, C$, where the vanishing factor $C = \lim_{(A\to\infty)} (A)^{-1} \int_0^A d\alpha\, \alpha^{-1}\, y^{1/\alpha}$ does not depend on y. The fact that C does not depend on y can be seen if we calculate the derivative of C with respect to y. We obtain $dC/dy = \lim_{(A\to\infty)} (A)^{-1} \int_0^A d\alpha\, \alpha^{-2}\, y^{1/\alpha-1}$. Introducing a new integration variable z = A/α and assuming that A is arbitrarily large but finite, we obtain $dC/dy = \lim_{(A\to\infty)} (A)^{-2}\, y^{-1} \int_1^\infty dz$

$y^{z/A}$. The integral can be evaluated to lead to $dC/dy = \lim_{(A\to\infty)} (A)^{-2}\, 1/ [y \ln (y^{1/A})]\, y^{z/A} |_{z=1}^{\infty}$. Inserting the limits, we obtain $dC/dy = \lim_{(A\to\infty)} 1/ [A\, y \ln (y)]\, (y^{\infty/A}\, y - y^{1/A})$. As $0<y<1$ this expression approaches 0 as A increases. We have therefore shown that C is indeed does not depend on y. Because of the normalization, C must be equal to $C = 1/\int_0^1 dy\, y^{-1}$.

Next, we show that this particular density leads to the Benford distribution $P(d)= \log_{10} (1+1/d)$.

$$
\begin{aligned}
P(d) &= \Sigma_{n=-\infty}^{\infty} \int_c^e dy\, \rho(y) \\
&= C\, \Sigma_{n=-\infty}^{\infty} \int_c^e dy\, (1/y) \\
&= C\, \Sigma_{n=-\infty}^{-1} [\ln (e) - \ln(c)] \\
&= C\, \Sigma_{n=-\infty}^{-1} [\ln [(d+1)\, 10^n] - \ln [d\, 10^n]] \\
&= C\, \Sigma_{n=-\infty}^{-1}\, n\, [\ln [(d+1)10] - \ln [d10]] \\
&= C\, \Sigma_{n=-\infty}^{-1}\, n\, \ln(1+1/d) \\
&= [C\, \Sigma_{n=-\infty}^{-1}\, n/\log_{10}(e)]\, \log_{10} (1+1/d) \qquad \text{(A3)}
\end{aligned}
$$

The prefactor $C\, \Sigma_{n=-\infty}^{-1}\, n /\log_{10}(e)$ is obviously independent of d and because of the required normalization condition for P(d), $\Sigma_{d=1}^{9} P(d)=1$, for consistency we have $C\, \Sigma_{n=-\infty}^{-1}\, n/\log_{10}(e) = 1$.

Appendix B

Here we review how we can decompose any real 9×9 matrix M(i,j) into its canonical vectors.

$$M(i,j) \;=\; \Sigma_{n=1}^{9}\, w_n\, p(i;n)\, q(j;n) \tag{B1}$$

In this decomposition, we assume the L^1-normalization of the positive vectors $\Sigma_{n=1}^{9}\, p(i;n) = \Sigma_{n=1}^{9}\, q(i;n) = 1$ and that the canonical vectors are perpendicular to each other, $\Sigma_{i=1}^{9}\, p(i;n)\, p(j;n) = 0$ for $j \neq i$. As two auxiliary positive definite 9×9 matrices P and Q we construct

$$P(i,j) \;=\; \Sigma_{k=1}^{9}\, M(i,k)M(j,k) \tag{B2a}$$

$$Q(i,j) \;=\; \Sigma_{k=1}^{9}\, M(k,i)M(k,j) \tag{B2b}$$

If we insert the functional dependence of M from Eq. (B1) into these definitions and use the orthogonality, we obtain

$$P(i,j) \;=\; \Sigma_{n=1}^{9}\, w_n^{2}\, p(i;n)\, p(j;n) \tag{B3a}$$

$$Q(i,j) \;=\; \Sigma_{n=1}^{9}\, w_n^{2}\, q(i;n)\, q(j;n) \tag{B3b}$$

Next, we compute the eigenvalues and eigenvectors of these two matrices. Both have an identical spectrum given by λ_n with n=1,2,3 ... 9, but they can differ by their eigenvectors V_n and W_n

$$\Sigma_{j=1}^{9}\, P(i,j)\, V_n(j) \;=\; \lambda_n\, V_n(i) \tag{B4a}$$

$$\Sigma_{j=1}^{9}\, Q(i,j)\, W_n(j) \;=\; \lambda_n\, W_n(i) \tag{B4b}$$

If we L^1-normalize all eigenvectors, $\Sigma_{j=1}^{9}\, V_n(j) = \Sigma_{j=1}^{9}\, W_n(j) = 1$ and compare Eqs. (B3) and (B4), we find that $V_n(j) = p(j;n)$ and $W_n(j) = q(j;n)$ that $\lambda_n = w_n^{2}\, \Sigma_{j=1}^{9} p(j;n)^2\, \Sigma_{i=1}^{9} q(i;n)^2$. In other words, we can reconstruct the original (positive) weight factors as

$$w_n \;=\; \lambda_n^{-1/2}\, [\Sigma_{j=1}^{9} p(j;n)^2\, \Sigma_{i=1}^{9} q(i;n)^2]^{-1/2} \tag{B5}$$

References

[1] T.P. Hill, Proc. Am. Math. Soc., 123(3), 887 (1995).

[2] T.P. Hill, Statistical Science, 10, 354 (1995).

[3] M.J. Nigrini "Benford's law: applications for forensic accounting, auditing and fraud detection" (Wiley Publications, New Jersey, 2012).

[4] G.A. Gottwald and M. Nicol, Physica A 303, 387 (2013).

[5] S.J. Miller, "Benford's law: theory and applications" (Princeton University Press: Princeton and Oxford, 2015).

[6] A. Berger and T.P. Hill, "An introduction to Benford's law" (Princeton University Press: Princeton, NJ, 2015).

[7] M. Sambridge, H. Tkalčić and A. Jackson, Geo. Phys. Lett. 37, 22301 (2010).

[8] J.L.H Cáceres, J.L. Pollo García, C.M. Martínez Ortiz and L. García Domínguez, Electron J. Biomed 1, 27 (2008).

[9] J. Burke and E. Kincanon, Am. J. Phys. 59, 952 (1991).

[10] J-C. Pain, Phys. Rev E 77, 012102 (2008).

[11] S. Newcomb, Am. J. Math 4 (1), 39 (1881).

[12] F. Benford, Proc. Am. Phil. Soc. 78, No. 4, 551 (1938).

[13] W. R. Mebane, Jr.. "Election Forensics: The second-digit Benford's law test and recent American presidential elections." in R.M. Alvarez, T.E. Hall and S. D. Hyde, eds., "Election fraud: detecting and deterring electoral manipulation" (Brookings Press, Washington, DC, 2008).

[14] A very extended collection of articles that analyze the first-digit law can be found under http://www.benfordonline.net/

[15] T.A. Mir, "Citations to articles citing Benford's law: a Benford analysis", arXiv:1602.01205v2 [cs.DL] (2016).

[16] L. Shao and B.-Q. Ma, Physica A 389, 3109 (2010).

[17] B.P. Flannery, S. Teukolsky, W.H. Press, and W.T. Vetterling, "Numerical recipes: the art of scientific computing" 3rd ed. (Cambridge University Press, Boston, 2007).

[18] E. Schmidt, Math. Ann. 63, 433 (1906).

[19] H. Everett, Rev. Mod. Phys. 29, 454 (1957).

[20] R. Grobe, K. Rzazewski and J.H. Eberly, J. Phys. B 27, L503 (1994).

[21] W.C. Liu, J.H. Eberly, S.L. Haan and R. Grobe, Phys. Rev. Lett. 83, 520 (1999).