

Research



Cite this article: Roman S, Bertolotti F. 2022
A master equation for power laws. *R. Soc. Open
Sci.* **9**: 220531.
<https://doi.org/10.1098/rsos.220531>

Received: 22 April 2022
Accepted: 14 November 2022

Subject Category:
Mathematics

Subject Areas:
mathematical physics/applied mathematics/
statistical physics

Keywords:
power laws, Markov chains, urn model

Author for correspondence:
Sabin Roman
e-mail: sr911@cam.ac.uk

A master equation for power laws

Sabin Roman^{1,2} and Francesco Bertolotti³

¹Centre for the Study of Existential Risk, University of Cambridge, Cambridge, UK

²Odyssean Institute, London, UK

³Carlo Cattaneo University LIUC, Castellanza, Lombardia, Italy

SR, 0000-0002-7513-3479

We propose a new mechanism for generating power laws. Starting from a random walk, we first outline a simple derivation of the Fokker–Planck equation. By analogy, starting from a certain Markov chain, we derive a master equation for power laws that describes how the number of cascades changes over time (cascades are consecutive transitions that end when the initial state is reached). The partial differential equation has a closed form solution which gives an explicit dependence of the number of cascades on their size and on time. Furthermore, the power law solution has a natural cut-off, a feature often seen in empirical data. This is due to the finite size a cascade can have in a finite time horizon. The derivation of the equation provides a justification for an exponent equal to 2, which agrees well with several empirical distributions, including Richardson’s Law on the size and frequency of deadly conflicts. Nevertheless, the equation can be solved for any exponent value. In addition, we propose an urn model where the number of consecutive ball extractions follows a power law. In all cases, the power law is manifest over the entire range of cascade sizes, as shown through log–log plots in the frequency and rank distributions.

1. Introduction

A power law is a nonlinear relationship between two quantities x and y that can be modelled generically by the following formula: $y = ax^k$, where k and a are constants, respectively, the exponent of the power law, and the width of the scaling relationship. The power law is scale-invariant relationship $f(xc) \propto f(x)$ that holds for any value of c . Graphically, it implies that the curve describing the relationship between x and y maintains its shape under any possible dilatation. Moreover, power laws can be represented as straight lines on a log–log plot, called the signature of the power law.

The signature can be employed to analyse empirical data by comparing the distribution of the data on a log–log plot with the best fitting power law. Probably, this representation was first introduced by J. C. Wills in 1922 to plot the distribution of the

number of species in a genus [1]. Many lognormal relationships appear as power laws when plotted for small ranges. As a consequence, to assert that a relationship between two variables is a power law, it should hold for at least two orders of magnitude [2].

In the next section, we enumerate some representative examples of power law distributions from physics, biology, ecology and social sciences. Furthermore, we list the most well-known mechanisms for generating power laws. In the third section, we proceed to propose a Markov process model for generating power laws. First, we outline a derivation of the Fokker–Planck equation by starting from a random walk and taking a continuum limit from a discrete state space. This is a well-known approach for deriving the Fokker–Planck equation. We provide our own treatment of this procedure by changing the difference operator Δ acting on discrete functions to the differential operator d that acts on real variables. We then propose a Markov chain that displays a power law in its long-term equilibrium distribution. Based on this and in analogy with the Fokker–Planck equation derivation, we obtain a master equation for power laws that we simplify to an analytically tractable form. Surprisingly, the solution can be written in closed form. Afterwards, in the fourth section, we propose an urn model where the number of cascades (consecutive balls drawn) scales according to a power law. This improves on an urn model by Brunk [3] that displays a power law relationship only in a limited range. Finally, we summarize our results in the Conclusion.

2. Background

The main challenge regarding power law distributions is understanding their origin. Broadly speaking, there are two ways power laws can emerge: probability transformations and generative processes. Probability transformations derive power law distributions from other distributions. Generative processes are algorithms that create new distributions. Here, we provide more details on generative processes, because our own work is of this nature. To the best of our knowledge, there does not exist a comprehensive review of all the power law distribution generative models (even if [1] finds a large set), and it is not in the scope of this paper to provide one.

There are three main strategies to obtain a power law distribution by transforming an existing probability distribution: combining exponential distributions, inverting quantities or looking at the extreme values of distributions.

In the first case, the transformation consists of substituting the variable of an exponential distribution with another exponential distribution. The resulting distribution follows a power law, in which parameters depend on the original exponential distribution. The distribution of sizes of populations that grow exponentially (i.e. with infinite carrying capacity) and can suddenly become extinct at any time step with the same probability follows a power law that can be derived from this model [4].

The second case requires the inversion of the quantities of the probability distribution. If these quantities pass through zero, the transformation results in a power law distribution with exponent equal to 2, resulting from the derivation of an inverted variable. The most notable example of a power law distribution obtained in this way occurs in the paramagnetic phase of the Ising model [5,6].

The third way originates from the extreme value theorems which provide results on the asymptotic behaviour of the extreme realizations. The Pickands–Balkema–de Haan theorem [7,8] states that the conditional distribution of a random variable above a certain threshold tends to a Pareto distribution when the threshold tends to infinity. Therefore, it provides a rationale for the widespread observation of Pareto or power law behaviour, since it is the limiting behaviour of large events for a whole class of probability distributions [9]. Analogously, the Fisher–Tippett–Gnedenko theorem [10,11] describes the possible distributions that the maximum value can have; this can shed light in cases where a simple power law does not appear from the application of the Pickands–Balkema–de Haan theorem [12].

In addition, we identified in the literature at least four notable generative processes for power law distributions: phase transitions (along with self-organized criticality), random walks, sample-space reduction processes and the Yule processes.

The idea underlying phase transitions and critical phenomena is simple but powerful. In systems governed by a single length scale, the scale can diverge in some given conditions, giving birth to scale-free systems in which quantities are distributed as a power law. The precise point at which the length scale diverges is called a critical continuous point or phase transition. Nevertheless, it is unlikely that the parameters that regulate the phase transition of a real-world system happen to fall on that specific value. So, the existence of critical phenomena is not enough to explain the presence of power law distributions in many natural and social systems. But some systems appear to self-organize to lay close to critical points, independently from the initial conditions. This phenomenon, called self-organized criticality, can be

exemplified by the classic sandpile model [13], which generates avalanches whose sizes are distributed according to a power law. Other examples of how self-organized criticality generates power law distributions includes earthquakes [14], neuronal avalanches [15] and forest fires [1,16]. However, the emergence and applicability of self-organized criticality is debatable [17].

Random walks are a succession of randomly generated steps on a given space [18]. The statistical properties of random walks tend toward universal distributions when the steps are independent of each other, and their number grows unbounded [19]. As a consequence, they have the ability to generate power law distributions even when the underlying rules are very simple. One of the most famous examples (and probably the simplest) relates to the first-return time. It can be proved that, under specific conditions, the distributions of the first-return times follows a power law [1,19].

The sample-space reduction process models history-dependent systems, such as the formation of sentences, whose number of possible different meanings is progressively reduced the more words are added to it. Corominas-Murtra *et al.* demonstrate that a power law distribution can be generated from such a process [20]. More specifically, Zipf's Law necessarily emerges from the reduction over time of the number of possible states in which a system can be, as a consequence of symmetry breaking in random sampling processes [20] or from dependency structures in component systems [21]. The reduction space process explains some well-known domain-specific power law generation models, such as preferential attachment.

The Yule process is a generative model developed by G. Udny Yule to explain the power law distribution of the number of species in taxonomic groups [22]. It derives from simple rules. Firstly, starting from a pool of groups, the probability for a group to increase by one is proportional to the number of its elements. Secondly, at any speciation event, there is a possibility of generating a new species belonging to a brand-new taxonomic group. The model is a simplification of reality since it ignores extinction events. Nevertheless, it has been customized to explain the emergence of power laws in other systems such as city sizes [23] or paper citations [24].

Finally, other generative processes include highly optimized tolerance [25,26], the coherent noise model of biological extinction [27], the repeated fragmentation model of fixed length elements [1], the dynamic of times between records in a random process [28], the Hawkes processes [29] and out-from-criticality feedback [30]. A master equation for power laws has been previously obtained in the literature [31–35] but only considers stationary distributions or transient dynamics, while our results give a non-stationary solution [36]. Furthermore, our choice of transition rates has not been considered before.

To the best of our knowledge, the above examples and mechanisms for generating power laws do not generally address the question: how do power laws change (or are preserved) over time? Answering this question requires a way to specify the time evolution of the power law distribution. Furthermore, in many cases, empirical power laws have a cut-off point above which the relationship no longer holds [1], and this has not been captured by generative models, which are realistic only in specific value ranges. The next section provides a mechanism by which an equation dictating the time evolution of power laws can be obtained. In addition, the solution displays a natural cut-off beyond a certain time horizon.

3. Markov process models

Below we present an outline for a well-known derivation of the Fokker–Planck equation starting from a discrete Markov chain and taking a continuum limit. By analogy, we introduce a discrete Markov chain that has a power law stationary distribution and then we derive a master equation for power laws.

3.1. Fokker–Planck equation

The transition probabilities for a one-dimensional random walk are

$$\text{and } \left. \begin{aligned} P_{i,i+1} &= p \\ P_{i,i-1} &= 1 - p = q. \end{aligned} \right\} \quad (3.1)$$

The probability for x jumps to the right and $n - x$ jumps to the left is

$$\begin{aligned} P(x, n) &= \binom{n}{x} p^x q^{n-x} \\ &\sim \frac{1}{\sqrt{2\pi npq}} \exp\left[-\frac{(x - np)^2}{2npq}\right], \end{aligned} \quad (3.2)$$

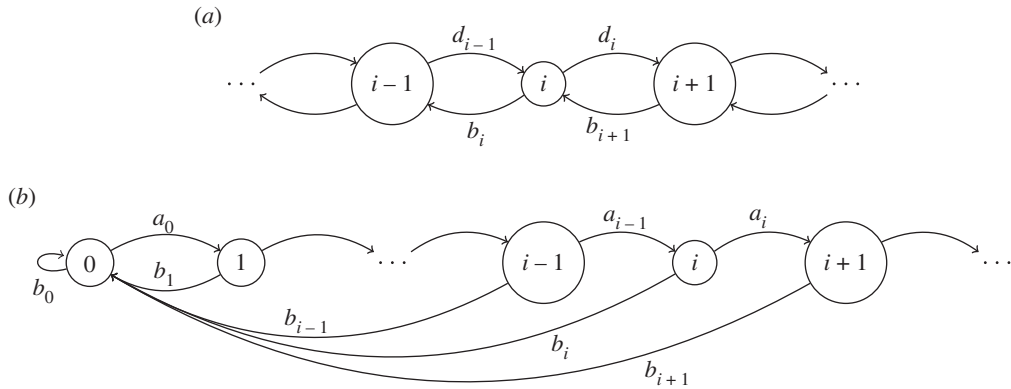


Figure 1. Discrete-state Markov chains for: (a) a continuous-time random walk, with multi-step transition probabilities satisfying equation (3.3), (b) a continuous-time process, with multi-step transition probabilities satisfying equation (3.8). The stationary distribution of the embedded Markov chain is a power law, as shown in derivation (3.7).

where in the limit of large enough n we can approximate the binomial distribution with a Gaussian. Let i, s be arbitrary states. For a continuous-time Markov chain we write $p_i(t) = P_{si}(0; t)$. For the Markov chain in figure 1a, we can write down the following equation for $p_i(t)$:

$$\begin{aligned} \frac{dp_i(t)}{dt} &= d_{i-1}p_{i-1} + b_{i+1}p_{i+1} - d_i p_i - b_i p_i \\ &= -(a_{i+1}p_{i+1} - a_i p_i) + d_{i-1}p_{i-1} - 2d_i p_i + d_{i+1}p_{i+1}, \end{aligned} \quad (3.3)$$

where b_i, d_i are the rates at which the process leaves state i to state $i-1$, respectively to state $i+1$. Furthermore, we define $a_i = d_i - b_i$. Using the forward and backward difference operators $\Delta x_n = x_{n+1} - x_n$ and $\nabla x_n = x_n - x_{n-1}$, the above equation can be written

$$\left. \begin{aligned} \frac{dp_i(t)}{dt} &= -\Delta(a_i p_i(t)) + \Delta \nabla(d_i p_i(t)) \\ \frac{\partial p(x, t)}{\partial t} &= -\frac{\partial}{\partial x}(a(x)p(x, t)) + \frac{\partial^2}{\partial x^2}(d(x)p(x, t)), \end{aligned} \right\} \quad (3.4)$$

which in the continuous limit gives the Fokker-Planck equation. If $a(x) = 0$ and $d(x) = D$ where D is constant, then the solution with initial condition $p(0, x) = \delta(x - x_0)$ is given by

$$p(x, t) = \frac{1}{\sqrt{4\pi Dt}} \exp\left[-\frac{(x - x_0)^2}{4Dt}\right]. \quad (3.5)$$

3.2. Power law equation

We consider the Markov chain with the following transition matrix:

$$\left. \begin{aligned} P_{i,i+1} &= \exp\left(-\frac{k}{i+1}\right) \\ P_{i,0} &= 1 - \exp\left(-\frac{k}{i+1}\right). \end{aligned} \right\} \quad (3.6)$$

and

The dominant contribution to the long-term equilibrium distribution is given by

$$\begin{aligned} \pi(x) &\simeq \prod_{i \leq x} P_{i,i+1} \\ &= \exp\left(-k \sum_i \frac{1}{i}\right) \\ &\sim \exp(-k \ln x) \\ &= x^{-k}. \end{aligned} \quad (3.7)$$

In figure 2a, we compare numerical solutions for the equilibrium distribution with the approximation (3.7), showing that the power law approximation is highly accurate. The Markov chain given by (3.6) is

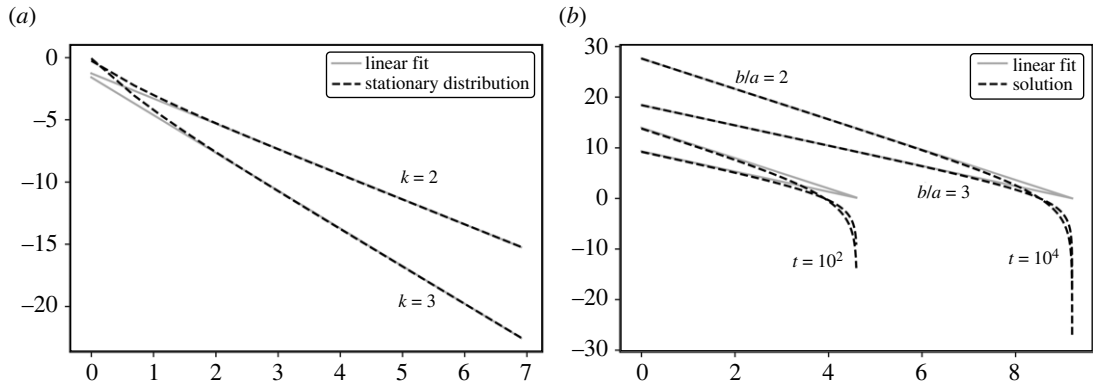


Figure 2. Log–log plots of (a) the stationary distribution (dotted black line) of the Markov chain with transition probabilities (3.6) for $k = 2, 3$, along with a linear fit (solid grey line), and (b) solution (3.12) (dotted black line) to equation (3.9) for $b/a = 2, 3$ up to times $t = 10^2, 10^4$, along with linear fits (solid grey line).

the embedded chain of the continuous-time Markov process in figure 1b. Using the same notation as in the previous section, the probability $p_i(t)$ satisfies the following differential equation:

$$\frac{dp_i(t)}{dt} = a_{i-1}p_{i-1} - a_i p_i - b_i p_i. \quad (3.8)$$

We can write (3.8) using the backward difference operator. If we let the space index go from discrete to continuous, such as in the previous section where $i \rightarrow x$, then equation (3.8) becomes

$$\left. \begin{aligned} \frac{dp_i(t)}{dt} &= -\nabla(a_i p_i(t)) - b_i p_i(t) \\ \downarrow \\ \frac{\partial p(x, t)}{\partial t} &= -\frac{\partial}{\partial x}(a(x)p(x, t)) - b(x)p(x, t) \end{aligned} \right\} \quad (3.9)$$

which gives

$$\frac{\partial p(x, t)}{\partial t} = -a(x) \frac{\partial p(x, t)}{\partial x} - (a'(x) + b(x))p(x, t), \quad (3.10)$$

where $a(x) = a e^{-k/x}$, $b(x) = b(1 - e^{-k/x})$ and a, b are constants. By Taylor expanding up to $1/x$, we get $a(x) \simeq a - (ak/x)$ and $a'(x) + b(x) \simeq (bk/x)$. Without loss of generality, we can relabel bk as b and the equation reduces to

$$\frac{\partial p(x, t)}{\partial t} = -a \frac{\partial p(x, t)}{\partial x} - \frac{b}{x} p(x, t). \quad (3.11)$$

The term containing a/x can be neglected because the probability $p(x, t)$ depends on a power of x (see below), and so each term on the right-hand side of (3.11) is of the same order in x .

Let $c(x, t)$ be the unnormalized solution of equation (3.11). The solution $c(x, t)$ has a closed-form expression for $x > 0$ and $x < at$,

$$c(x, t) = c_0 \left(\frac{at}{x} - 1 \right)^{b/a}, \quad (3.12)$$

where $c(x, 0) = c_0$. In figure 2b, we plot the solution (3.12) for different values of b/a and see a power law behaviour over a range of scales that increases over time. The tail of the solution falls off abruptly just as we see in many real-world examples of power law distributions [1].

Equation (3.11) posits that cascades (consecutive transitions) occur over time in inverse proportion to their size. Furthermore, there is a constant rate a at which they can increase in size. For a finite time horizon t , there is a maximum size that a cascade can reach, namely at , as can be seen in the solution (3.12).

We can get additional insights from equation (3.10) if we consider the second order Taylor expansion of $a'(x) + b(x) \simeq (bk/x) + ((2ak - bk^2)/2x^2)$. If $k=2$ a/b then the $1/x^2$ term cancels, and this implies an exponent of 2 in the solution (3.12). This is not necessary in deriving equation (3.11) but can serve as a rationale for the experimental observation of power laws with exponent 2.

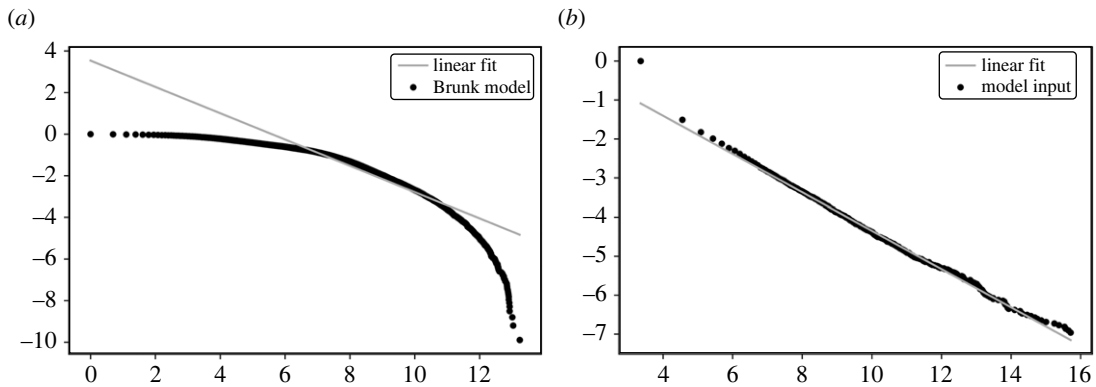


Figure 3. Log–log plots of the rank distribution of (a) cascade sizes (black dots) in Brunk’s model [3], showing a narrow range where a power law is observed, along with a linear fit (solid grey line), and (b) cascade sizes (black dots) in the model we propose with exponent $k = 1.5$, along with a linear fit (solid grey line). The line of best fit has slope 0.49, in good agreement with the theoretical value of $k - 1 = 0.5$.

4. Urn model

Building on the insights of the previous section, we propose an urn model that displays power law behaviour. We can approximate the probability (3.8) to move from i to $i + 1$ state as

$$\exp\left(-\frac{k}{i+1}\right) \simeq 1 - \frac{k}{i+1} \quad (4.1)$$

for large enough i . Hence, if an adequate value of k is chosen, then we can expect to see a power law emerging even in a simple model.

Brunk [3] proposes a model claiming to exhibit cascades that occur with frequencies inversely proportional to the cascade size, and provides graphs showing that a power law distribution emerges. We attempted to reproduce the results of the paper, but we obtained an exponential distribution instead of a power law. A power law relationship is seen to occur in a narrow range of scales that the graphs in the article use (figure 3a). Furthermore, we can analytically prove that the distribution that emerges is an exponential distribution and not a power law. The article [3] has been cited in recent years, and we consider it important to highlight the limitations of the model.

Assume the urn contains N balls in total, with K balls black and the rest white. The model has the following steps:

- (i) A ball is drawn, if it is black then the next ball is drawn, and the process is repeated until a white ball is drawn. This constitutes a cascade within the model. Black balls are drawn without replacement.
- (ii) Each time a white ball is drawn the number of black balls is increased by a fraction g , while the number of white balls stay the same. If a white ball is drawn, then we repeat from step (i).

The probability that a cascade of size n occurs, i.e. n black balls are drawn followed by a white ball, is

$$\begin{aligned} P(S = n) &= \frac{K}{N} \cdot \frac{K-1}{N-1} \cdot \dots \cdot \frac{K-n+1}{N-n+1} \cdot \frac{N-K}{N-n} \\ &= \frac{K!}{N!} \cdot \frac{(N-n)!}{(K-n)!} \cdot \frac{N-K}{N-n} \\ &\approx \frac{K^{K+1/2}}{N^{N+1/2}} \cdot \frac{(N-n)^{N-n+1/2}}{(K-n)^{K-n+1/2}} \cdot \frac{N-n-(K-n)}{N-n} \\ &= \left(1 - \frac{n}{N}\right)^{N+1/2} \left(1 - \frac{n}{K}\right)^{-(K+1/2)} \left(\frac{K-n}{N-n}\right)^n \left(1 - \frac{K-n}{N-n}\right). \end{aligned} \quad (4.2)$$

If $K \approx N$ then the first two terms in the last row of (4.2) cancel to a large extent. If we take the logarithm of (4.2), we get

$$\begin{aligned} \log P(S = n) &\approx n \log \left(\frac{K - n}{N - n} \right) + \log \left(1 - \frac{K - n}{N - n} \right) \\ &\approx n \log(f) + \log(1 - f), \end{aligned} \quad (4.3)$$

where $f = (K - n)/(N - n) \approx K/N$ is approximately the fraction of black balls in the urn. Equation (4.3) shows that probability of a cascade is exponentially distributed, at least as long as $n \ll N$. Furthermore, as K increases then $f \rightarrow 1$, which implies a uniform distribution.

We propose a different urn model. Let N_0 , K_0 and W be the initial number of balls, respectively, black and white ones. The procedure of the model is as follows:

- (i) Black balls are drawn with replacement. The number of black balls is increased by W/k after each draw. We continue to draw until a white ball is drawn.
- (ii) The process repeats with resetting the number of balls to their original values $N = N_0$, $K = K_0$.

The rank distribution of the cascade sizes is plotted in figure 3b for $k = 1.5$. As we can see, the graph is very well fit by a straight line, indicating that a genuine power law emerges across the entire range of cascade sizes.

At the i th extraction we have the total number of balls N_i , of black balls K_i and white balls W . Then

$$N_i = K_i + W. \quad (4.4)$$

The urn model is designed such that at each extraction the ratio of K_i and N_i is approximately given by

$$\frac{K_i}{N_i} = e^{-k/(i+1)}. \quad (4.5)$$

Using the previous two equations we obtain that

$$\begin{aligned} K_i &= \frac{e^{-k/(i+1)}}{1 - e^{-k/(i+1)}} W \\ &\simeq \left(\frac{i+1}{k} - 1 \right) W. \end{aligned} \quad (4.6)$$

The number of black balls is increased after each draw by

$$K_{i+1} - K_i \simeq \frac{W}{k}. \quad (4.7)$$

If a constant C number of black balls is added after each draw, then $k = W/C$.

5. Discussion

The relevance of this study relies upon two aspects: the increasing importance of understanding statistical laws in complex systems due to the availability of larger datasets [37], and the presence of power law distributions in various application domains [4,38,39]. The applicability of the models proposed in §3.2 is restricted to phenomena where time plays a key role in increasing cascade sizes. The more time passes from the initial observation, the greater the chances of seeing a larger cascade (which would place it at the right-end tail of the distribution at the time of observation). These phenomena can appear in a wide variety of disciplinary fields.

The frequency versus the amplitude of earthquakes [1] and avalanches [40,41] follows a power law, as well as the distribution of the peak gamma-ray intensity of solar flares [42], and the size of moon craters per surface area [43]. The largest observed cascade in these examples depends on time, with more extreme examples being discovered the longer the observation record is.

The social sciences also offer numerous examples of power law distributions [44] that naturally evolve in time (but whose time dependence has not necessarily been quantified), such as the range of time between two deaths in serial killers' behavioural patterns [45], in narrative structure [46] or in the budget distribution of movies [47]. In economics, they mainly derive from aggregation and rich-get-richer processes. The existence of power laws in the distribution of wealth and income are known at least from the nineteenth century [48]. Notably, the work of Pareto was the first one to discover the presence of a power law distribution in a social system [38]. Power law distributions also appear in markets. For example, the frequency of firm size,

measured by the number of employees, is distributed according to a power law [49]. Specifically, the exponent of this power law is approximately 1.06, which suggests there exists a ‘Zip’s Law’ for firms [38].

Western popular music markets present power law distributions in the lifetime of albums [50]. Stock markets exhibited power law distributions since their very beginning [51]. It has been shown that the number of trades per day, the size of price movements when a large volume of shares are bought or sold, and the number of shares traded per time period are power law distributed with exponents 3, 0.5 and 1.5, respectively [52–54]. Inter-trading times also exhibit scaling properties consistent with power laws [55].

Another field in which power law distributions are pervasive is urban studies, where there are many scaling relationships between quantifiable city properties [56], such as between the number of gas stations and population levels [57], and the distribution of the population of cities in a given geographical area [58]. This latter distribution appears in different geographical areas [59] and depends on the granularity of the sampling [60]. Power law distributions also characterize digital infrastructure. The network analysis of Internet topology shows that at the end of the 1990s the degrees of the nodes were already power law distributed [61,62]. Also, the number of connections to a server in a single day, by a specific subsection of Internet users, follows a power law [63].

Furthermore, the academic system generates power laws. As first noted by Price, the number of citations of papers follows a power law distribution [64]. Similar features in academic systems can be observed in recent times, even in specific disciplinary areas. For instance, the cumulative number of citations over time for papers dealing with protein kinases is distributed as a power law. This is a consequence of a phenomenon called the Harlow–Knapp (H-K) effect, which is the propensity of the biomedical and pharmaceutical research communities to concentrate their research on a tiny fraction of the proteome [65,66].

In history, the well-known Richardson’s Law states that the frequency of wars’ sizes (measured in casualties) are power law distributed. This law is considered one of the few robust statistical regularities in studies of political conflict [67]. Nevertheless, some recent studies seem to cast doubt on Richardson’s Law, or at least suggest having a greater caution regarding it [68]. Another example from history regards the Roman Empire [69], for which the survival time of its emperors is distributed according to a power law [70].

Certain power law relationships do not depend on the observation period and the proposed models would not apply. For example, in physics, the well-known Stefan–Boltzmann Law describes a power law relationship between the total amount of energy radiated per time from a black body due to its temperature [71,72]. Furthermore, in particle physics and astrophysics there are numerous examples of power laws, from the Tully–Fisher relationship between a galaxy’s luminosity and its rate of rotation [73], to the proportionality of the spin and mass of hadrons [74] and the density of eigenvalues of the Dirac operator in certain theories [75,76]. Other examples come from biology, where many works describe a scaling relationship between the basal metabolic rate and the body mass of animals [77–79], such as the well-known Kleiber’s Law, which indicates that the basal metabolism of mammals increases according to $m^{3/4}$ where m is the overall body mass [80].

Another well-known example (with no time dependence) is Zipf’s Law, which states that the rank frequency distribution of words in any sufficiently long text is distributed as a power law [81]. This law is independent of the language and holds even for artificial languages such as Esperanto [82]. While Zipf investigates specifically the distribution of words, a generalization of this law was later developed by Mandelbrot [83].

We end the discussion with a cautionary note on the discovery of power law in data and models. As is well noted in the literature [1], linear relationships in log–log plots in the frequency distributions of observations is insufficient to robustly establish that a power law holds. A more reliable test is to check if the linear relationship holds in log–log plots of rank distributions. As we have shown, the model proposed by Brunk [3] does not show power law behaviour. Another example where a power law is claimed to be observed is in the emergence of a scale-free network in systems with bounded rationality [84]. As other work has shown [85], the topology is not scale-free but rather of a core–periphery type.

6. Conclusion

In this paper, we have proposed a new mechanism to generate power laws based on Markov models. By analogy with the derivation of the Fokker–Planck equation, we have obtained a master equation for power laws. The result is supported by the fact that the underlying embedded Markov chain has an equilibrium distribution that is a power law and because a simplified version of the equation admits a closed form solution that is a power law.

We proposed three models: a discrete-time and discrete-space Markov chain with transition probabilities (3.6), a continuous-time and discrete-space Markov chain shown in figure 2b and a Markov process (continuum in both time and space), with master equation given by (3.9). A simplified form of (3.9) is equation (3.11), which is valid in the limit of large cascade sizes. Further considerations indicate that an exponent of $k = 2$ gives additional cancellations independent of cascade size. This value of k is consistent with exponents observed in Richardson's Law [68] and in the net worth of Americans [1]. Nevertheless, equation (3.11) has a closed-form solution for any exponent value. In addition, the equation provides the time dynamics for the power law relationship along with a natural cut-off size depending on the time horizon.

The stationary solutions for the general equation are given by a first-order differential equation, whose solutions have been explored elsewhere [32]. However, a stationary solution to the power law equation might not always be well defined, as our time-dependent solution (3.12) for the simplified equation (3.11) shows (the time derivative is non-zero). The insight we get from the exact time-dependent solution (3.5) of the Fokker-Planck equation is that the standard deviation of spatial displacements is proportional to the square root of time, which is characteristic of Brownian motion. Similarly, the time-dependent solution (3.12) gives the largest cascade size we can expect to see is linearly proportional to time.

Based on the insights from the Markov process, we propose a simple urn model that illustrates power law behaviour over the entire range of cascade sizes. The model only considers a constant addition of balls over time and despite the simplicity of the mechanism, a robust power law is observed to emerge. Our results are in contrast to a prior model by Brunk [3] which we prove does not show a genuine power law distribution. Finally, we discuss some possible applications of the proposed model to phenomena where power laws are observed and where the duration of the period of observation is important. As far as we know, all these contributions are novel to the literature.

Data accessibility. This article has no additional data.

Authors' contributions. S.R.: conceptualization, formal analysis, investigation, methodology, project administration, visualization, writing—original draft, writing—review and editing; F.B.: data curation, validation, writing—original draft, writing—review and editing.

All authors gave final approval for publication and agreed to be held accountable for the work performed therein.

Conflict of interest declaration. We declare we have no competing interests.

Funding. This work was supported by Grantham Foundation for the Protection of the Environment. We also thank the World University Network (WUN) for early support of this work in 2014.

References

- Newman ME. 2005 Power laws, Pareto distributions and Zipf's law. *Contemp. Phys.* **46**, 323–351. (doi:10.1080/00107510500052444)
- Stumpf MP, Porter MA. 2012 Critical truths about power laws. *Science* **335**, 665–666. (doi:10.1126/science.1216142)
- Brunk GG. 2000 Understanding self-organized criticality as a statistical process. *Complexity* **5**, 26–33. (doi:10.1002/(SICI)1099-0526(200001/02)5:3<26::AID-CPLX6>3.0.CO;2-C)
- Reed WJ. 2001 The Pareto, Zipf and other power laws. *Econ. Lett.* **74**, 15–19. (doi:10.1016/S0165-1765(01)00524-9)
- Jan N, Moseley L, Ray T, Stauffer D. 1999 Is the fossil record indicative of a critical system? *Adv. Complex Syst.* **2**, 137–141. (doi:10.1142/S0219525999000084)
- Krasnytska M, Berche B, Holovatch Y, Kenna R. 2021 Generalized Ising model on a scale-free network: an interplay of power laws. *Entropy* **23**, 1175. (doi:10.3390/e23091175)
- Balkema AA, de Haan L. 1994 Residual life time at great age. *Ann. Probab.* **2**, 792–804. (doi:10.1214/aop/1176996548)
- Pickands J. 1975 Statistical inference using extreme order statistics. *Ann. Stat.* **3**, 119–131. (doi:10.1214/aos/1176343003)
- Alfarano S, Lux T. 2010 Extreme value theory as a theoretical background for power law behavior. Kiel Working Paper no. 1648, pp. 1–10.
- Fisher RA, Tippett LH. 1928 Limiting forms of the frequency distribution of the largest or smallest member of a sample. *Math. Proc. Camb. Philos. Soc.* **24**, 180–190. (doi:10.1017/S0305004100015681)
- Gnedenko B. 1943 Sur la distribution limite du terme maximum d'une serie aleatoire. *Ann. Math.* **44**, 423–453. (doi:10.2307/1968974)
- Sornette D. 2006 *Critical phenomena in natural sciences. Chaos, fractals, selforganization and disorder: concepts and tools*. Berlin, Germany: Springer.
- Bak P, Tang C, Wiesenfeld K. 1987 Self-organized criticality: an explanation of the 1/f noise. *Phys. Rev. Lett.* **59**, 381–384. (doi:10.1103/PhysRevLett.59.381)
- Olami Z, Feder HJS, Christensen K. 1992 Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes. *Phys. Rev. Lett.* **68**, 1244–1247. (doi:10.1103/PhysRevLett.68.1244)
- Beggs JM, Plenz D. 2003 Neuronal avalanches in neocortical circuits. *J. Neurosci.* **23**, 11167–11177. (doi:10.1523/JNEUROSCI.23-35-11167.2003)
- McKenzie D, Kennedy MC. 2012 Power laws reveal phase transitions in landscape controls of fire regimes. *Nat. Commun.* **3**, 726. (doi:10.1038/ncomms1731)
- Frigg R. 2003 Self-organised criticality—what it is and what it isn't. *Stud. Hist. Philos. Sci. A* **34**, 613–632. (doi:10.1016/S0039-3681(03)00046-3)
- Hughes BD. 1995 *Random walks and random environments: random walks*, vol. 1. Oxford, UK: Clarendon Press.
- Vignat C, Plastino A. 2006 Power-law random walks. *Phys. Rev. E—Stat. Nonlinear, Soft Matter Phys.* **74**, 051124. (doi:10.1103/PhysRevE.74.051124)
- Corominas-Murtra B, Hanel R, Thurner S. 2015 Understanding scaling through history-dependent processes with collapsing sample

- space. *Proc. Natl Acad. Sci. USA* **112**, 5348–5353. (doi:10.1073/pnas.1420946112)
21. Mazzolini A, Grilli J, De Lazzari E, Osella M, Lagomarsino MC, Gherardi M. 2018 Zipf and Heaps laws from dependency structures in component systems. *Phys. Rev. E* **98**, 012315. (doi:10.1103/PhysRevE.98.012315)
 22. Yule GU. 1925 A mathematical theory of evolution based on the conclusions of Dr. J. C. Willis, F.R.S. *J. R. Stat. Soc.* **88**, 433–436. (doi:10.2307/2341419)
 23. Simon HA. 1955 On a class of skew distribution functions. *Biometrika* **42**, 425–440. (doi:10.1093/biomet/42.3-4.425)
 24. Krapivsky PL, Redner S, Leyvraz F. 2000 Connectivity of growing random networks. *Phys. Rev. Lett.* **85**, 4629. (doi:10.1103/PhysRevLett.85.4629)
 25. Carlson JM, Doyle J. 1999 Highly optimized tolerance: a mechanism for power laws in designed systems. *Phys. Rev. E—Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics* **60**, 1412.
 26. Carlson JM, Doyle J. 2000 Highly optimized tolerance: robustness and design in complex systems. *Phys. Rev. Lett.* **84**, 2529. (doi:10.1103/PhysRevLett.84.2529)
 27. Sneppen K, Newman ME. 1997 Coherent noise, scale invariance and intermittency in large systems. *Physica D* **110**, 209–222. (doi:10.1016/S0167-2789(97)00128-0)
 28. Sibani P, Littlewood PB. 1993 Slow dynamics from noise adaptation. *Phys. Rev. Lett.* **71**, 1482. (doi:10.1103/PhysRevLett.71.1482)
 29. Kanazawa K, Sornette D. 2020 Nonuniversal power law distribution of intensities of the self-excited Hawkes process: a field-theoretical approach. *Phys. Rev. Lett.* **2**, 137–141. (doi:10.1103/PhysRevLett.125.138301)
 30. Schulman LS. 2021 Apparent power laws can occur without criticality. *Entropy* **23**, 1486. (doi:10.3390/e23111486)
 31. Biró TS, Neda Z. 2017 Dynamical stationarity as a result of sustained random growth. *Physical Review E* **95**, 032130. (doi:10.1103/PhysRevE.95.032130)
 32. Biró TS, Neda Z. 2018 Unidirectional random growth with resetting. *Phys. A: Stat. Mech. Appl.* **499**, 335–361. (doi:10.1016/j.physa.2018.02.078)
 33. Biró TS, Neda Z, Telcs A. 2019 Entropic divergence and entropy related to nonlinear master equations. *Entropy* **21**, 993. (doi:10.3390/e21100993)
 34. Gere I, Kelemen S, Tóth G, Biró TS, Neda Z. 2021 Wealth distribution in modern societies: collected data and a master equation approach. *Phys. A: Stat. Mech. Appl.* **581**, 126194. (doi:10.1016/j.physa.2021.126194)
 35. Biró TS, Csillag L, Neda Z. 2021 Transient dynamics in the random growth and reset model. *Entropy* **23**, 306. (doi:10.3390/e23030306)
 36. Inácio I, Velinho J. 2022 Comments on mathematical aspects of the Biró–Neda model. *Mathematics* **10**, 644. (doi:10.3390/math10040644)
 37. Gerlach M, Altmann EG. 2019 Testing statistical laws in complex systems. *Phys. Rev. Lett.* **122**, 168301. (doi:10.1103/PhysRevLett.122.168301)
 38. Gabaix X. 2016 Power laws in economics: an introduction. *J. Econ. Perspect.* **30**, 185–206. (doi:10.1257/jep.30.1.185)
 39. Pinto CM, Mendes Lopes A, Machado JA. 2012 A review of power laws in real life phenomena. *Commun. Nonlinear Sci. Numer. Simul.* **17**, 3558–3578. (doi:10.1016/j.cnsns.2012.01.013)
 40. Naisbitt A, Forster R, Birkeland K, Harrison W. 2008 Avalanche frequency and magnitude: using power law exponents to investigate snow avalanche size proportions through space and time. In *Int. Snow Science Workshop, Whistler, British Columbia, Canada*.
 41. Polizzi S, Pérez-Reche FJ, Arneodo A, Argoul F. 2021 Power-law and log-normal avalanche size statistics in random growth processes. *Phys. Rev. E* **104**, L052101. (doi:10.1103/PhysRevE.104.L052101)
 42. Lu ET, Hamilton RJ. 1991 Avalanches and the distribution of solar flares. *Astrophys. J.* **380**, L89–L92. (doi:10.1086/186180)
 43. Neukum G, Ivanov B. 1994 Crater size distributions and impact probabilities on Earth from lunar, terrestrial planets, and asteroid cratering data. In *Hazards due to comets and asteroids*, pp. 359–416. Tucson, AZ: University of Arizona Press.
 44. Turcotte DL, Rundle JB. 2002 Self-organized complexity in the physical, biological, and social sciences. *Proc. Natl Acad. Sci. USA* **99**, 2463–2465. (doi:10.1073/pnas.012579399)
 45. Simkin MV, Roychowdhury VP. 2014 Stochastic modeling of a serial killer. *J. Theor. Biol.* **355**, 111–116. (doi:10.1016/j.jtbi.2014.03.039)
 46. Gessey-Jones T, Connaughton C, Dunbar R, Kenna R, MacCarron P, O’Conchobhair C, Yose J. 2020 Narrative structure of *A Song of Ice and Fire* creates a fictional world with realistic measures of social complexity. *Proc. Natl Acad. Sci. USA* **117**, 28 582–28 588. (doi:10.1073/pnas.2006465117)
 47. Bertolotti F, Roman S. 2022 Risk sensitive scheduling strategies of production studios on the US movie market: an agent-based simulation. *Intell. Artif.* **16**, 81–92.
 48. Pareto V. 1896 *Écrits sur la courbe de la répartition de la richesse*. Geneva, Switzerland: Librairie Droz.
 49. Axtell RL. 2001 Zipf distribution of U.S. firm sizes. *Science* **293**, 1818–1820. (doi:10.1126/science.1062081)
 50. Schneider L, Gros C. 2019 Five decades of US, UK, German and Dutch music charts show that cultural processes are accelerating. *R. Soc. Open Sci.* **6**, 190944. (doi:10.1098/rsos.190944)
 51. Ribeiro AS, Pinheiro FL, Santos FC, Polónia A, Pacheco JM. 2018 Structural and temporal patterns of the first global trading market. *R. Soc. Open Sci.* **5**, 180577. (doi:10.1098/rsos.180577)
 52. Bouchaud JP, Farmer JD, Lillo F. 2009 How markets slowly digest changes in supply and demand. In *Handbook of financial markets: dynamics and evolution*, pp. 57–160. Amsterdam, The Netherlands: North Holland.
 53. Kyle AS, Obizhaeva AA. 2013 Market microstructure invariance: theory and empirical tests. See http://www.smallake.kr/wp-content/uploads/2016/03/Kyle_Obizhaeva_INWARIANCE_79_BERKELEY.pdf.
 54. Plerou V, Gopikrishnan P, Stanley HE. 2005 Quantifying fluctuations in market liquidity: analysis of the bid-ask spread. *Phys. Rev. E—Stat. Nonlinear Soft Matter Phys.* **71**, 046131. (doi:10.1103/PhysRevE.71.046131)
 55. Ivanov PC, Yuen A, Podobnik B, Lee Y. 2004 Common scaling patterns in intertrade times of US stocks. *Phys. Rev. E* **69**, 056107. (doi:10.1103/PhysRevE.69.056107)
 56. Wu L, Gong C, Yan X. 2019 Taylor’s power law and its decomposition in urban facilities. *R. Soc. Open Sci.* **6**, 180770. (doi:10.1098/rsos.180770)
 57. Kühnert C, Helbing D, West GB. 2006 Scaling laws in urban supply networks. *Physica A* **363**, 96–103. (doi:10.1016/j.physa.2006.01.058)
 58. Gabaix X, Ioannides YM. 2004 The evolution of city size distributions. In *Handbook of regional and urban economics*, vol. 4, pp. 2341–2378. Amsterdam, The Netherlands: North Holland.
 59. Rozenfeld HD, Rybski D, Gabaix X, Makse HA. 2011 The area and population of cities: new insights from a different perspective on cities. *Am. Econ. Rev.* **101**, 2205–25. (doi:10.1257/aer.101.5.2205)
 60. Mori T, Smith TE, Hsu WT. 2020 Common power laws for cities and spatial fractal structures. *Proc. Natl Acad. Sci. USA* **117**, 6469–6475. (doi:10.1073/pnas.1913014117)
 61. Barabási AL, Albert R. 1999 Emergence of scaling in random networks. *Science* **286**, 509–512.
 62. Faloutsos M, Faloutsos P, Faloutsos C. 1999 On power-law relationships of the internet topology. *Comput. Commun. Rev.* **29**, 251–262. (doi:10.1145/316194.316229)
 63. Adamic L, Huberman BA. 2005 The nature of markets in the world wide web. *SSRN Electron. J.* (doi:10.2139/ssrn.166108)
 64. De Solla Price DJ. 1965 Networks of scientific papers. *Science* **149**, 510–515. (doi:10.1126/science.149.3683.510)
 65. Fedorov O, Müller S, Knapp S. 2010 The (un)targeted cancer kinome. *Nat. Chem. Biol.* **6**, 166–169. (doi:10.1038/nchembio.297)
 66. Grueneberg DA et al. 2008 Kinase requirements in human cells: I. Comparing kinase requirements across various cell types. *Proc. Natl Acad. Sci. USA* **105**, 16 472–16 477. (doi:10.1073/pnas.0808019105)
 67. Clauset A, Wiegand FW. 2010 A generalized aggregation-disintegration model for the frequency of severe terrorist attacks. *J. Conflict Resolut.* **54**, 179–197. (doi:10.1177/0022002709352452)
 68. Zwetsloot R. 2018 Testing Richardson’s law: a (cautionary) note on power laws in violence data. *SSRN Electron. J.* (doi:10.2139/ssrn.3112804)
 69. Roman S, Palmer E. 2019 The growth and decline of the Western Roman Empire: quantifying the dynamics of army size, territory, and coinage. *Clodynamics* **10**, 1–23.
 70. Ramos PL, Costa LF, Louzada F, Rodrigues FA. 2021 Power laws in the Roman Empire: a survival analysis. *R. Soc. Open Sci.* **8**, 210850. (doi:10.1098/rsos.210850)

71. Boltzmann L. 1884 Ableitung des Stefan'schen Gesetzes, betreffend die Abhängigkeit der Wärmestrahlung von der Temperatur aus der electromagnetischen Lichttheorie. *Ann. Phys. Chem.* **22**, 291–294.
72. Stefan J. 1879 Über die beziehung zwischen der wärmestrahlung und der temperatur. *Wien. Ber. II* **79**, 391–428.
73. Tully RB, Fisher JR. 1977 A new method of determining distances to galaxies. *Astron. Astrophys.* **54**, 661–673.
74. Chew GF, Frautschi SC. 1962 Regge trajectories and the principle of maximum strength for strong interactions. *Phys. Rev. Lett.* **8**, 41–44.
75. Cheng A, Hasenfratz A, Petropoulos G, Schaich D. 2013 Scale-dependent mass anomalous dimension from Dirac eigenmodes. *J. High Energy Phys.* **2013**, 1–23. (doi:10.1007/JHEP07(2013)061)
76. Del Debbio L, Lucini B, Pica C, Patella A, Rago A, Roman S. 2013 Large-volume results in SU (2) with adjoint fermions. (<https://arxiv.org/abs/1311.5597>)
77. Benedict FG. 1938 *Vital energetics: a study in comparative basal metabolism*. Washington, DC: Carnegie Institute Washington Publication.
78. Brody S, Lardy HA. 1946 Bioenergetics and growth. *J. Phys. Chem.* **50**, 168–169. (doi:10.1021/j150446a008)
79. Kleiber M. 1961 *The fire of life: an introduction to animal energetics*. New York, NY: John Wiley & Sons, Inc.
80. Niklas KJ, Kutschera U. 2015 Kleiber's law: how the *Fire of Life* ignited debate, fueled theory, and neglected plants as model organisms. *Plant Signal. Behav.* **10**, e1036216. (doi:10.1080/15592324.2015.1036216)
81. Zipf GK. 1949 *Human behavior and the principle of least effort*. Cambridge, MA: Addison-Wesley.
82. Manaris B, Pellicoro L, Pothering G, Hodges H. 2006 Investigating Esperanto's statistical proportions relative to other languages using neural networks and Zipf's law. In *AIA Int. Conf. on Artificial Intelligence and Applications, Innsbruck, Austria, 13–16 February*.
83. Mandelbrot B. 1968 Information theory and psycholinguistics: a theory of word frequencies. In *Readings in mathematical social science*, pp. 350–368. Cambridge, MA: MIT Press.
84. Kasthurirathna D, Piraveenan M. 2015 Emergence of scale-free characteristics in socio-ecological systems with bounded rationality. *Sci. Rep.* **5**, 1–16. (doi:10.1038/srep10448)
85. Roman S, Brede M. 2017 Topology-dependent rationality and quantal response equilibria in structured populations. *Phys. Rev. E* **95**, 052310. (doi:10.1103/PhysRevE.95.052310)