The Transition Towards Immortality: Non-linear Autocatalytic Growth of Citations to Scientific Papers

Michael Golosovsky & Sorin Solomon

🐴 Springer

Springer

The Transition Towards Immortality: Non-linear Autocatalytic Growth of Citations to Scientific Papers

Michael Golosovsky · Sorin Solomon

Abstract We discuss microscopic mechanisms of complex network growth, with the special emphasis of how these mechanisms can be evaluated from the measurements on real networks. As an example we consider the network of citations to scientific papers. Contrary to common belief that its growth is determined by the *linear* preferential attachment, our microscopic measurements show that it is driven by the *nonlinear autocatalytic growth*. This invalidates the scale-free hypothesis for the citation network. The nonlinearity is responsible for a dramatic dynamical phase transition: while the citation lifetime of majority of papers is 6–10 years, the highly-cited papers have practically infinite lifetime.

1 General Introduction

A lot of empirical evidence for the power-law degree distribution in natural networks has been amassed during last decade. This led to the conjecture that these networks are scale-free. It is widely believed that the growth of the scale-free networks is driven by the cumulative advantage mechanism [1] which is commonly known as the preferential attachment [2, 3]. This mechanism assumes that Δk, the number of links acquired by a node during a short time interval Δt is determined by the number of already acquired links k,

$$\Delta k = A(k + k_0). \tag{1}$$

Here, k is the node degree, k_0 is the "initial attractivity" and A is the attachment rate (aging function) which is time-dependent. Equation (1) yields the power-law degree distribution, $P(k) \propto 1/k^\gamma$, which is generally considered as a fingerprint of a scale-free network. The linear preferential attachment (Eq. (1)) is believed to be one of the most important microscopic mechanisms that generates the scale-free complex networks which are so ubiquitous in nature.

M. Golosovsky (✉) · S. Solomon
The Racah Institute of Physics, The Hebrew University of Jerusalem, Jerusalem 91904, Israel
e-mail: michael.golosovsky@mail.huji.ac.il

This statement is often reversed and the power-law degree distribution in a growing network is considered as an evidence for the linear preferential attachment. The parameter k_0 is estimated from the exponent of the degree distribution [2]:

$$\gamma = 2 + \frac{k_0}{m}, \tag{2}$$

where m is the mean degree. This approach meets several difficulties. First of all it yields unrealistically high $k_0 \approx m$. Second and most important- the validity of the power-law approximation for degree distribution in complex networks has been contested. Indeed, since the node degree is a discrete and non-negative number, the scale-free power-law function cannot provide a good fit for the nodes with small degree. At best, it can fit only the fat tail of the distribution. However, several recent studies showed that the degree distribution in complex networks can deviate from the power-law dependence even in the fat tail [4–8].

Krapivsky and Redner [9] showed that the deviation necessarily occurs if the attachment kernel is nonlinear,

$$\Delta k = A(k + k_0)^\alpha. \tag{3}$$

In particular, for sublinear attachment kernel, $\alpha < 1$, the network is characterized by the stretched exponential degree distribution; while for the superlinear kernel, $\alpha > 1$, the network organizes into a "winner takes all" configuration [9, 10]. While for linear attachment kernel the network achieves stationary degree distribution, for the nonlinear case the degree distribution is nonstationary. In the sequel we call the dynamics governed by Eq. (3) as the "nonlinear autocatalytic growth" [11, 12] and reserve the term "preferential attachment" for Eq. (1) that generates the power-law degree distribution.

Although the deviation of α from unity can be small it affects dramatically the network structure. To what extent the growth mechanism of real networks deviates from the linear preferential attachment (Eq. (1)) is an important question. Recent experimental studies [4, 13–24] that measured microscopic growth of complex networks, came up with the conclusion that α is close to unity, in such a way that the growth mechanism is nearly linear (see Table 1). However, to which extent α deviates from unity remained an open question until now. The above studies could hardly measure this deviation due to time-dependence of α, finite precision limited by the size of their databases and, most important—due to uncertainty arising from the use of different methodologies. In particular, Ref. [24] applied four different methods to measure attachment kernel in the network of the US patent-to-patent citations and found different exponents ranging from 1.12 to 1.38.

The goal of this study is the high precision measurement of the microscopic growth rate of a complex network and the determination of the attachment exponent α. Following the accepted practice [4, 13, 15, 18], as an object of our research we chose one of the best-documented complex networks: citations to scientific papers. Here, the papers are nodes and citations to these papers are links. We performed high-statistics and time-resolved study of the citation dynamics of a very large and homogeneous set of papers. In what follows we compare two methods of measuring the microscopic growth rate of this network: averaging (histogram) and cumulation. We found that the former method is quite reliable and yields superlinear attachment kernel, $\alpha \approx 1.25$, while the latter method is prone to quantization errors. We came to conclusion that the microscopic growth mechanism of the citation network follows *nonlinear* autocatalytic growth (Eq. (3)).

We elaborate on a dramatic consequence of nonlinearity: if one considers a citations dynamics governed by the superlinear attachment kernel, one is led to conclusion that this network contains a subset of the papers that will be cited forever. Thus we witness a dynamical

Table 1 Testing preferential attachment in real networks

Network	Ref.	Attachment exponent α	Method
Citations of scientific papers	Jeong et al. [13]	0.95	cumulation
	Eom & Fortunato [15]	1	cumulation
	Redner [4]	0.9–1.05	running average
	Wang et al. [18]	1	running average
	Golosovsky & Solomon [29]	1–1.25 (*grows with time*)	histogram
Citations of US patents	Csardi et al. [17]	1.2	histogram
	Sheridan et al. [24]	1.23–1.27	Metropolis-Hastings
	Valverde et al. [16]	1–1.25 (*grows with time*)	cumulation
Scientific collaboration	Jeong et al. [13]	0.8	cumulation
	Tomassini & Luthi [19]	0.76	cumulation
	Newman [20]	1	histogram
Movie actors	Jeong et al. [13]	1	cumulation
	Eom et al. [14]	1	cumulation
Wikipedia	Capocci et al. [21]	0.76	histogram
Google	Jeong et al. [13]	1.05	cumulation
Internet	Eom et al. [14]	1	cumulation
Internet dictionary	Herdagdelen et al. [22]	1	histogram
Protein networks	Eisenberg & Levanon [23]	1	cumulation

phase transition in which citation lifetime of a paper diverges to infinity. Our measurements provide experimental evidence for such runaway papers that have practically infinite citation lifetime.

2 Methodology

To assess the microscopic growth mechanism of the citation network we focused on one discipline—Physics. We considered a cohort of papers published in the same year T_{publ} and measured the number of citations garnered by each paper in every subsequent year T_{cit}. To this end we used the Thomson-Reuters ISI Web of Science, chose 82 leading Physics journals, excluded review articles, comments, editorial, etc., and analyzed citation history of 40,195 original research papers published in these journals in $T_{publ} = 1984$ (this covers ~ 95 % of the Physics papers published in this year). The cumulative citation distributions for this data set were demonstrated elsewhere [8]. Figure 1 shows some aggregate characteristics of this set: the mean number of citations and the fraction of uncited papers.

In what follows we focus on two variables: (a) $k_{i,t}$—the cumulative number of citations, i.e. the total number of citations accumulated by a paper i in the period between T_{publ} and T_{cit}; and (b) $\Delta k_{i,t+\Delta t}$—the number of additional citations gained by the same paper in a short time window between T_{cit} and $T_{cit} + \Delta t$. Here, $\Delta t = 1$ year and $t = T_{publ} - T_{cit} + 1$ (if $T_{publ} = T_{cit}$ then $t = 1$, in such a way that $k_{i,1}$ measures the number of citations during the year when the paper was published). Figure 2 shows $\Delta k_{i,t+1}$ versus $k_{i,t}$. (Specifically, $k_{i,6}$ is the number of citations garnered by a paper i from 1984 to 1989 while $\Delta k_{i,7}$ is the number of

Fig. 1 Time dependence of the fraction of uncited papers P_0 and of the mean number of citations m for 40195 Physics papers published in 1984. t is the number of years after publication, whereas the publication year corresponds to $t = 1$. The *continuous lines* are guide to the eye. While the number of uncited papers saturates after 15 years, the mean number of citations does not saturate even after 25 years
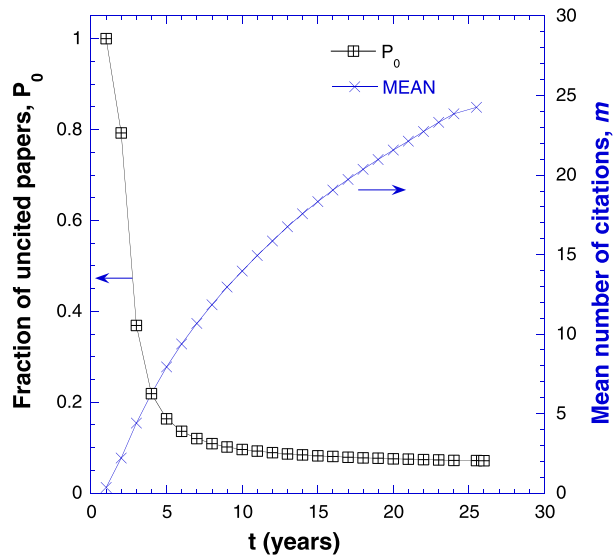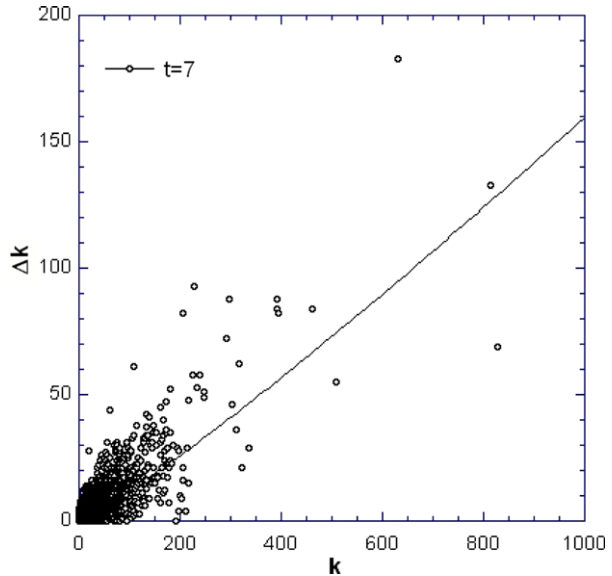

Fig. 2 The scatter plot of the number of additional citations $\Delta k_{i,7}$, garnered by each paper i during seventh year after publication. The *horizontal axis* shows $k_{i,6}$—the total number of citations garnered by the same paper during six previous years. The *solid line* displays approximation by Eq. (4) with $\alpha = 1.13, k_0 = 1, A = 0.065$

citations garnered by the same paper in 1990.) While the trend of increasing $\Delta k_{i,t+1}$ versus $k_{i,t}$ is clearly visible, the fluctuations are so strong that Fig. 2 does not provide an obvious proof of the validity of Eq. (1).

This is not unexpected since the actual number of newly acquired citations is a stochastic variable. We define $\lambda_i(t) = \overline{\Delta k_{i,t}}/\Delta t$ which is the average citation rate over the ensemble of the nodes with the same $k_{i,t}$. The autocatalytic growth model actually claims that $\lambda_i = A(k_i + k_0)^\alpha$, in such a way that

$$\Delta k_{i,t} = A(k_i + k_0)^\alpha \Delta t + \sigma \, dW(t) \tag{4}$$

Fig. 3 Mean number of additional citations, $\lambda(k) = \overline{\Delta k_i}$, as a function of the number of previous citations $k(t)$; t is the number of years after publication. To include uncited papers ($k = 0$) the *horizontal axis* displays $k + 1$ instead of k. Each set of points corresponds to a certain citing year. The *straight dashed line* shows linear approximation $\overline{\Delta k} \propto (k + k_0)$ where $k_0 = 1$. The data deviate upwards from this linear dependence, especially at $t = 15$–24. The *continuous lines* show better, superlinear fits, $\lambda = A(k + k_0)^\alpha$ where A, α and k_0 are fitting parameters. The superlinear dependences fit the highly-cited papers ($k > 100$) and uncited papers ($k = 0$) as well

where $\sigma\,dW(t)$ is a random variable with zero mean and σ^2 variance (for brevity we replace thereon $t + 1$ by t). In contrast to $\Delta k_{i,t}$ which is a discrete variable, $\lambda_i(t)$ is a continuous one. To verify whether the noisy data, such as those shown in Fig. 2, are generated by the growth law suggested by Eq. (4), there have been developed two methods: averaging (histogram) and cumulation. We processed our data using both these methods and obtained conflicting results. In what follows we compare these two methods and develop a control tool to check their internal consistency.

3 Comparison Between Different Methods to Measure the Microscopic Growth Law of Citation Network
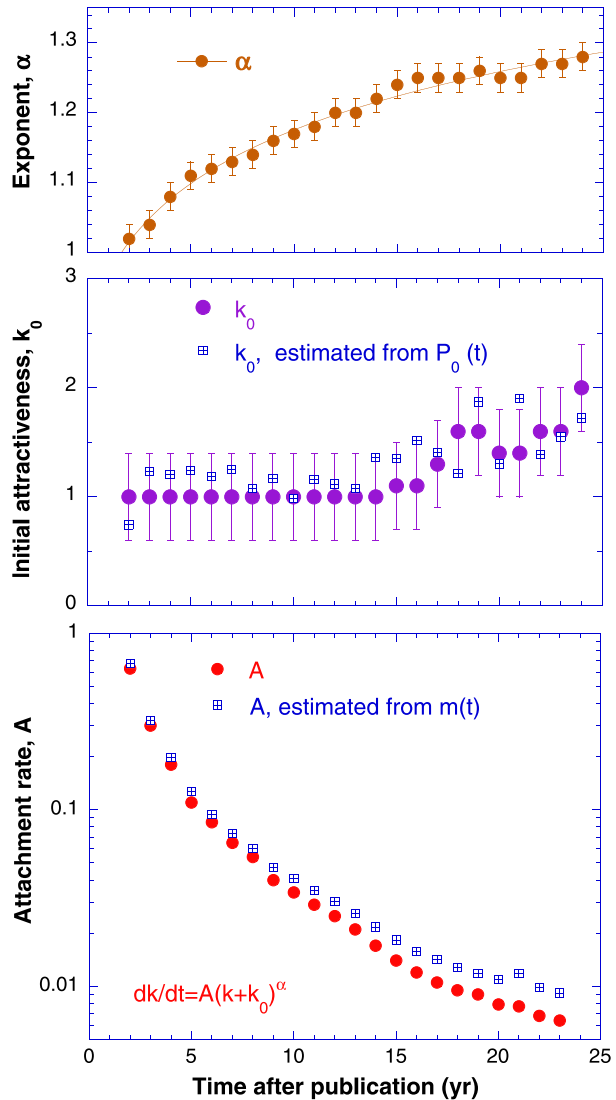
3.1 Histogram (Averaging) Method

To infer the microscopic growth law from the noisy data such as those shown in Fig. 2, one bins the data, finds the mean $\lambda = \overline{\Delta k_i}$ for each bin, and compares the resulting histogram to the prediction of Eq. (4). This approach was first used by Newman [20] to verify the linear preferential attachment hypothesis in real networks. References [17, 21, 22] followed this approach as well, while Refs. [4, 18] used a very similar moving average procedure.

To process our data in such a way we chose a certain citing year T_{cit}, grouped all papers into ~ 40 logarithmically-spaced bins, each bin containing the papers with close $k(t)$, found the mean number of citations λ for each bin and plotted it versus k. Figure 3 shows such $\lambda(k)$ dependences. In particular, the black circles indicate the results of the averaging procedure applied to the data of Fig. 2. The $\lambda(k)$ dependences are fairly well fitted by Eq. (4).

Figure 4 shows time dependence of the fitting parameters α, k_0, A. The exponent α gradually increases with time from $\alpha = 1$ to $\alpha = 1.28$, indicating linear attachment kernel for "young" papers and superlinear attachment kernel for "old" papers. The initial attractiveness $k_0 \approx 1.1$ is almost time-independent and is surprisingly close to ad hoc assumption of de Solla Price [1]. The time dependence of the attachment rate A can be approximated by

Fig. 4 Time dependence of the parameters of Eq. (4).
(**a**) Exponent α. The *continuous line* is a guide to the eye.
(**b**) Initial attractivity k_0. (**c**) Rate constant A. The *blue squares* in (**b**) and (**c**) show the estimates of k_0 and A based on Eqs. (6), (7) and Fig. 3, correspondingly. The consistency of k_0 and A obtained by two methods [*circles vs squares*] validates the *superlinear preferential attachment, $\alpha > 1$*
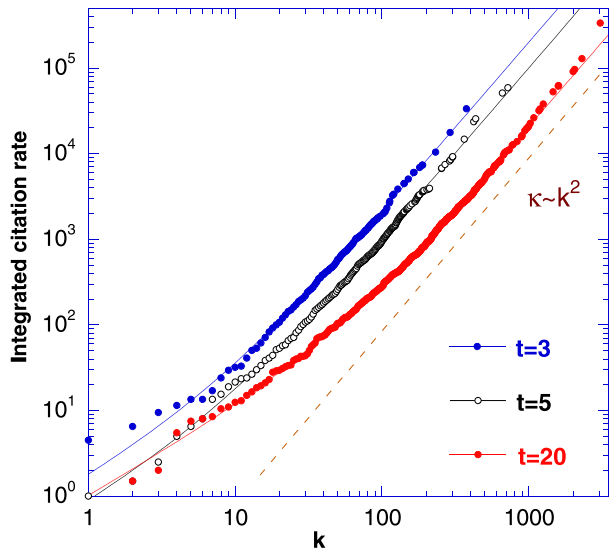

the empirical power-law dependence, $A = 3.3/(t + 0.3)^2$. A similar power-law dependence can be inferred from the US patent citation data of Ref. [17].

3.2 Cumulation Method

Jeong, Neda, and Barabasi [13] were the first to measure the growth rate of evolving networks used the cumulation method which quickly became the most popular tool to assess the preferential attachment in real networks [14–16, 19, 23, 24]. This method consists in calculation of the kernel $\kappa(k) = \int_0^k \Delta k' dk'$ where k' is the total number of citations garnered by a paper by year T_{cit} and $\Delta k'$ is the number of citations accrued by this paper during time window between T_{cit} and $T_{cit} + \Delta t$ where Δt is usually 1 year. The integration is performed over all papers that garnered $k' \leq k$ citations by year T_{cit}. The key assumption behind this

Fig. 5 Integrated number of additional citations $\kappa(k) = \int_0^k \Delta k' dk'$, as a function of the number of previous citations k; t is the number of years after publication. We used the same raw data as those shown in Fig. 2 and applied trapezoidal numerical integration routine implemented by MATLAB. The *dashed line* shows quadratic dependence $\kappa \propto k^2$ as expected for the linear preferential attachment. The data at high k follow this dependence as if they were generated by the linear preferential attachment, $\alpha \approx 1$. The *continuous lines* show fit given by Eq. (5) with $\alpha = 1$ and A and k_0 as fitting parameters

scheme is that the fluctuations in Δk are averaged out and the resulting integral is the same as if Eq. (4) were integrated directly over k at fixed t i.e.,

$$\kappa(k) = \frac{A}{\alpha + 1} \left[(k + k_0)^{\alpha+1} - k_0^{\alpha+1} \right]. \tag{5}$$

We applied this cumulation method to our data. Figure 5 shows the results. The fluctuations have been dramatically reduced, as expected. Equation (5) fits well the data for high k, while for low k the fit is less satisfactory. Figure 6 shows the fitting parameters. We found that the deviation of the exponent α from unity is within the experimental uncertainty, $\Delta\alpha = \pm 0.05$. Therefore, to find A and k_0 we set $\alpha = 1$ in our fitting procedure.
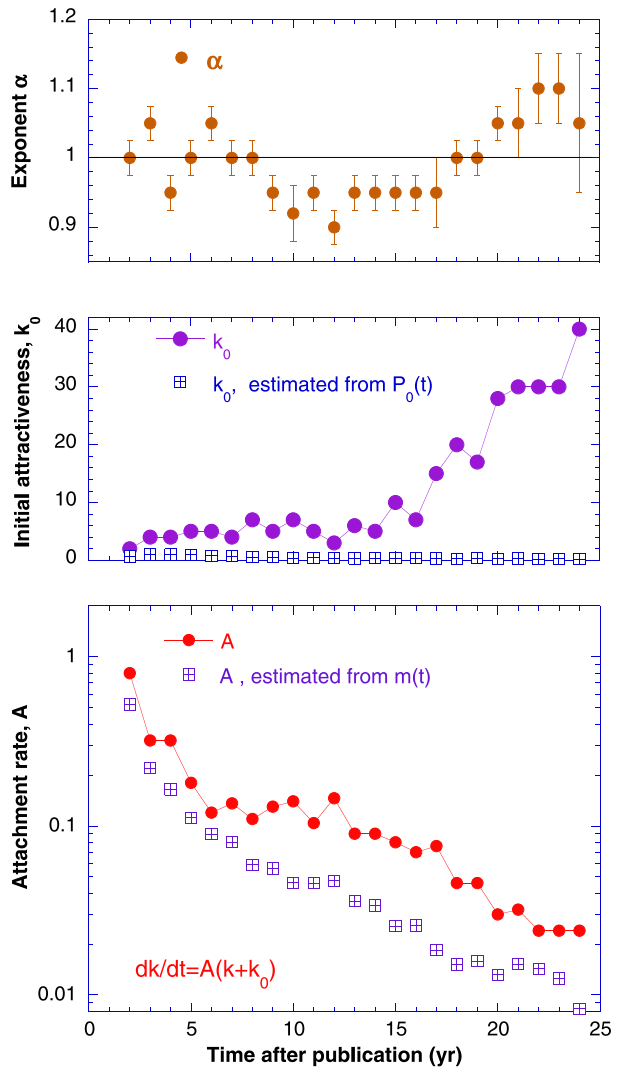
Figure 6 shows that the fitting parameters found in such a way are notably different from those found by the averaging method (Fig. 4). Most important—the exponent α is close to unity while that found from the averaging method is higher than unity. The initial attractivity k_0 is high and increases with time, while that found from the averaging method is close to unity and almost time-independent. The attachment rate A exceeds that found from the averaging method, especially at long times. The discrepancy between the two methods calls for some control tool. In what follows we develop such tool and use it to decide which method: averaging or cumulation is more reliable.

3.3 Control Tool

We consider here an additional tool to estimate the microscopic growth parameters of a growing network. We have developed this indirect method to check the internal consistency of the histogram and cumulation methods. This control method is based on two assumptions: (i) the microscopic growth law given by Eq. (4) is valid for all papers including uncited ones, and (ii) the exponent α is known. Then, the microscopic parameters A and k_0 (see Eq. (4)) may be estimated from the dynamics of the macroscopic parameters: the mean number of citations m, and the fraction of uncited papers P_0.

The mean, $m = \overline{k_i(t)}$, is the average number of citations garnered by a paper during the period between T_{publ} and $T_{publ} + t$. The averaging here is performed over all papers. Differ-

Fig. 6 Time dependence of the parameters of Eq. (4) as found using numerical integration (Eq. (5)). (**a**) Exponent α. (**b**) Initial attractivity k_0. (**c**) The rate constant A. The *blue squares* in (**b**) and (**c**) show, correspondingly, the estimates of k_0 and A based on Eqs. (6), (7) and the data of Fig. 3. The results obtained by two methods [*circles vs squares*] strongly differ. This casts doubt on the validity of the cumulation method as applied to citations and especially on its claim of the linear preferential attachment, $\alpha = 1$

entiation with respect to time yields the average number of additional citations garnered by a paper between t and $t + \Delta t$, namely, $\frac{dm}{dt} \Delta t = \overline{\Delta k_i}$. We average Eq. (4) over all papers and for $\Delta t = 1$ we find $\frac{dm}{dt} = A\overline{(k + k_0)^\alpha} \approx A(m + k_0)^\alpha$ (the last approximation holds because α is close to unity). This yields the rate constant

$$A \approx \frac{dm}{dt} \frac{1}{(m + k_0)^\alpha}. \tag{6}$$

For $k = 0$ Eq. (4) reduces to $\lambda_0 = Ak_0^\alpha$ where λ_0 is the average citation rate of previously uncited papers. The latter can be recast through the fraction of uncited papers P_0 as follows,

$\lambda_0 \approx \frac{1}{P_0} \frac{dP_0}{dt}$.[1] Equation (6) yields then

$$k_0 \approx \left(\frac{1}{AP_0} \frac{dP_0}{dt} \right)^{\frac{1}{\alpha}}. \tag{7}$$

We solve Eqs. (6), (7) for known α and find A and k_0. We expect that the parameters A and k_0 obtained by this control method are consistent with those found directly.

For the histogram (averaging) method, the A and k_0 found from Eqs. (6), (7) are indeed consistent with those found by the direct procedure (Fig. 4). The difference in k_0 is within the measurement uncertainty, while a small difference in A can be traced to the Jensen's inequality, $\overline{x^\alpha} > \overline{x}^\alpha$ for $\alpha > 1$. However, for the cumulation procedure, the A and k_0 found from Eqs. (6), (7) are inconsistent with those found directly (Fig. 5): A is substantially lower and k_0 is also much smaller than those found directly. This inconsistency calls for a deeper consideration of the validity of the cumulation method as applied to citations.

While the cumulation method works well for noisy continuous data with Gaussian fluctuations, its applicability to citations is problematic. Since the additional citations Δk are discrete and non-negative, their fluctuations around the mean are non-symmetrical and their magnitude is on the order of the mean (see Fig. 2). It appears that the standard numerical integration procedure as implemented in MATLAB does not work well for discrete, wildly fluctuating data that have non-symmetrical distribution around the mean. The straightforward application of the numerical integration procedure (cumulation) for quantifying dynamics of growing networks is thus ineffective. This method shall be specially tailored for the discrete variables with the non-Gaussian and strongly skewed fluctuation spectrum.

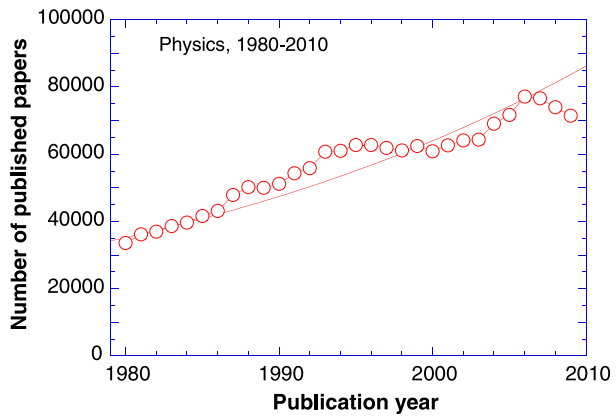## 4 The Effect of the Exponential Growth of Publications on Citation Dynamics

Most theoretical studies consider networks that grow linearly in time. In fact, they define a "network time" in such a way that new nodes are added to network at constant rate. It should be noted that citation networks grow exponentially with time. In what follows we define the network time for the citation network and recast our results in terms of the network time.

Figure 7 shows the annual growth of the number of original research papers and reviews in 82 leading Physical journals as covered by the ISI Thomson-Reuters Web of Science (excluding editorials, conference proceedings, etc.). The network growth is close to exponential, hence the difference between the physical time and network time is essential.

We define the "network year Δt^*" in such a way that the number of papers published during this time interval is equal to the number of papers published in 1984. (This is approximately equivalent to time rescaling, $t^* = \frac{e^{0.032t} - 1}{0.032}$ where t is the physical time.) To determine microscopic parameters of the citation dynamics we use Eq. (4) where $\Delta t = 1$

---

[1]Consider a cohort of papers published in the same year. After a couple of years, when the general interest to this cohort already decayed, the annual number of citations gained by previously uncited papers is either 0 or 1. Therefore, the mean annual number of additional citations gained by previously uncited papers is $\overline{\Delta k_i} \approx \Delta N_0/N_0$ where N_0 is the number of uncited papers and $\Delta N_0 = N_0(t) - N_0(t+1)$ is the number of uncited papers that got their first citation during recent year $t+1$. If the total number of papers in the dataset is N, then $\overline{\Delta k_i} \approx \Delta P_0/P_0$ where $P_0 = N_0/N$ is the fraction of uncited papers.

Fig. 7 The annual number of published Physics research papers (not including conference proceedings). The solid line shows exponential approximation corresponding to 1.3 % annual growth

year is replaced by Δt^*. The data shown in Fig. 3 are unaffected by this transformation although now they refer to network time which is different from physical time. The parameters α and k_0 remain the same. The new rate constant A^* is determined from the slope of the dependences shown in Fig. 3 divided by Δt^* instead of Δt.

Figure 8 shows some citation parameters in this time frame. The saturation exhibited by m and α is clearly visible. This should be compared to Figs. 1, 4a where the growth of m and α is more close to logarithmical. The attachment rate $A^*(t^*)$ turns out to be almost the same as the $A(t)$ dependence. This coincidence may be occasional.
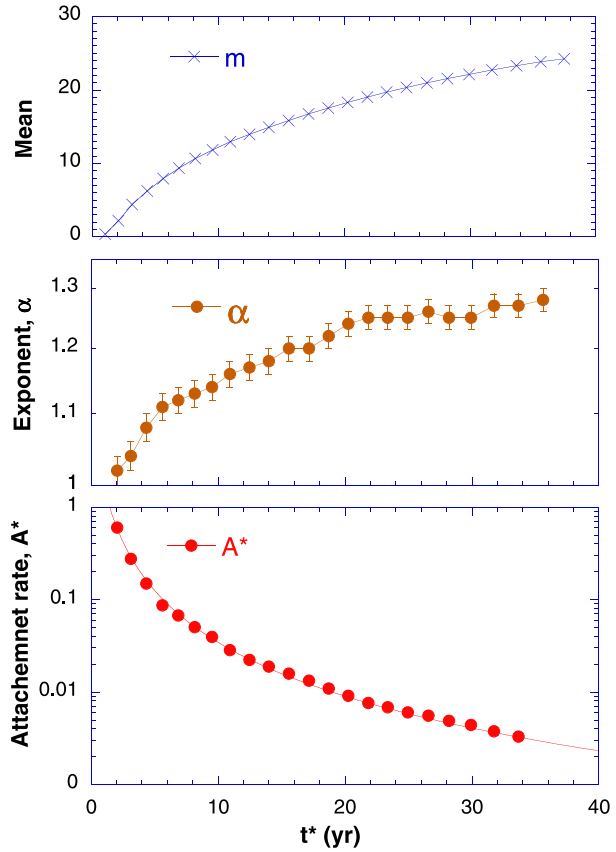
We conclude that the microscopic parameters of the growing citation network obtained by the averaging (histogram) method are most reliable. The results indicate that the growth of the citation network follows Eq. (4) with time-dependent exponent α which gradually increases from 1 to 1.28. By the way, the studies of US patent-to-patent citations by the histogram method yielded similar values $\alpha \sim 1.2$–1.27 [16, 17, 24]. We conclude that the citation network undergoes the nonlinear autocatalytic growth with the superlinear attachment kernel. Although the nonlinearity is weak, it leads to far-reaching consequences which we analyze below.

5 Divergence of the Citation Lifetime—Dynamical Phase Transition Towards Immortality

In what follows we analyze the consequences of the *nonlinear* growth mechanism of the citation network. In particular, we demonstrate that the nonlinearity is responsible for the enormous spread of citation lifetimes of scientific papers. To show this we consider citation dynamics of individual papers and distinguish between the initial period of $t_0 = 2$–3 years when a paper makes an immediate impact and the subsequent period when the citation dynamics of this paper to some extent is built on its initial success. The time dependence of the total citation count of a paper during this later period can be crudely estimated by integrating Eq. (4) with respect to time as if k were a continuous variable. In what follows we focus only on the papers with $k \gg 1$ (moderately- and highly-cited papers), in such a way that the term k_0 in Eq. (4) can be neglected. We also neglect for a moment the stochastic component of k. Assuming time-independent α we integrate Eq. (4) and find

$$k(t) = \frac{\widetilde{k}_0}{(1 - \delta \widetilde{k}_0^{\delta} \int_{t_0}^{t} A\,dt)^{1/\delta}}. \tag{8}$$

Fig. 8 Microscopic parameters of the citation dynamics versus network time t^*. The latter is defined in such a way that the citation network grows linearly with t^*. (**a**) m, the mean number of citations. (**b**) α, exponent of the attachment kernel. (**c**) A^*, the rate constant. The continuous line shows empirical approximation $A^* = 3.8/(t^* + 0.5)^2$

Here, $\widetilde{k}_0 = k(t_0)$ stands for the number of citations garnered by a paper during initial period t_0 (it shouldn't be mixed with k_0 that appears in Eq. (4)), and $\delta = \alpha - 1$. For $\delta \ll 1$ Eq. (8) reduces to a more transparent form

$$k(t) = \widetilde{k}_0 e^{\widetilde{k}_0^{\delta} \int_{t_0}^{t} A dt}. \tag{9}$$

The analysis of this equation we start from the case $\delta = 0$ that corresponds to the linear growth. In this case the $k(t)$ dependence can be factorized, $k(t) = \widetilde{k}_0 e^{\int_{t_0}^{t} A dt}$, where the $\widetilde{k}_0$ sets the scale and the factor $e^{\int_{t_0}^{t} A dt}$ sets the time dependence of the citation count. Therefore, if the growth mechanism were linear then the citation dynamics of all papers would follow the universal dependence $e^{\int_{t_0}^{t} A dt}$ which is independent of $\widetilde{k}_0$. This is in contrast to the nonlinear growth mechanism, $\delta \neq 0$, for which Eq. (9) can not be factorized. In this case both the scale and the time dependence of the total citation count depend on $\widetilde{k}_0$.

To provide the experimental evidence for such behavior we consider "citation age" $\langle t \rangle$— a nonparametric measure of the paper longevity introduced by Redner [4]. It is nothing else but the mean age of the papers that cite a given paper i,

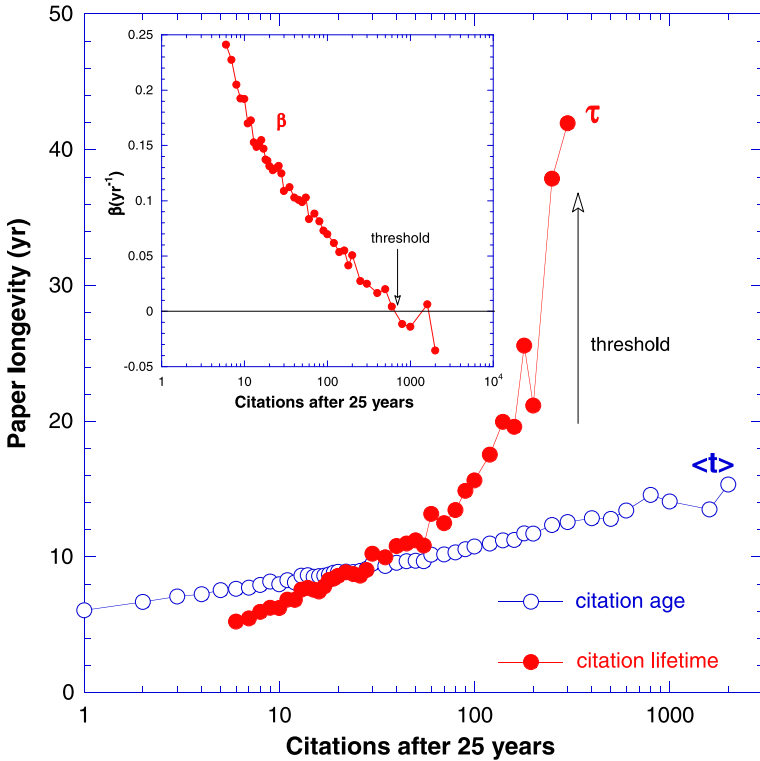$$\langle t_i \rangle = \frac{\int_0^t t_i dk_i}{k_i(t)} \tag{10}$$

Fig. 9 Citation age $\langle t \rangle$ (Eq. (10)) and citation lifetime τ (Eq. (11)) versus $k(t = 25)$—the number of citations garnered by a paper after 25 years. The data were binned and each data point represents the average over one bin. Note divergence of τ at the threshold of $k(t) \approx 600$. The *inset* shows citation rate β which changes its sign and becomes positive for $k(t) > 600$. The growth of $\langle t \rangle$ and τ with k is a signature of the *nonlinear autocatalytic process*

where t is the number of years after publication. In the extreme case when the citations grow linearly in time, $\langle t_i \rangle = t/2$. Citation age exceeding $t/2$ indicates accelerating growth while citation age below $t/2$ indicates some kind of saturation. Figure 9 shows that $\langle t \rangle$ increases with the number of citations as expected for the nonlinear growth mechanism, and eventually achieves the critical value of $t/2 = 12.5$ years. This means that there is appreciable number of papers whose citation dynamics didn't come to saturation and they are actively cited even 25 years after publication!

Another way to illustrate such exceptional behavior is to approximate Eq. (8) by the exponential dependence

$$k(t) = \frac{K}{\beta}\left[1 - e^{-\beta(t-\Delta)}\right] \tag{11}$$

where K is some scale factor, Δ is the (small) delay between the publication of the paper and the onset of citations, and β characterizes the citation rate. The latter is negative when $k(t)$ accelerates with time and positive when $k(t)$ comes to saturation. In this latter case $\tau = 1/\beta$ has the meaning of citation lifetime which is related to citation age (Eq. (10)) as follows: for the exponential dynamics and in the long-time limit $\tau + \Delta \approx \langle t \rangle$.

We measured $k(t)$ for all papers in our dataset, approximated it using Eq. (11) and found microscopic parameters β and Δ. Since these microscopic parameters strongly fluctuate, we binned all dataset into 40 logarithmically-spaced bins and considered the average over the papers in each bin. The results are shown in Fig. 9. The citation rate changes sign and becomes positive for highly-cited papers, indicating acceleration. This change of sign occurs at the same threshold where citation age becomes equal to $t/2$ (Eq. (10)). The citation lifetime is $\tau \sim 5$–6 years for low-cited papers, for moderately- and highly-cited papers citation lifetime increases and even diverges, as it is predicted by Eq. (8).

The divergence of the citation lifetime is a direct consequence of the nonlinear autocatalytic growth and it demonstrates the tendency of citation network to develop a few hubs that attract the majority of citations. In the cumulative citation distribution these hubs appear as "runaways" [8]. These most highly-cited papers have all chances to achieve infinite citation lifetime. This observation extends the well-known adage "the rich get richer" to "the rich live longer".

6 Discussion

6.1 Why Is the Growth Mechanism of Citation Networks so Close to Linear?

When viewed from the perspective of network dynamics, the preferential attachment mechanism does not favor any particular value of the attachment exponent. Therefore, the ubiquity of linear or nearly linear preferential attachment seems enigmatic. However, if we consider network dynamics from the perspective of a single node, the ubiquity of nearly linear preferential attachment appears naturally.

Indeed, in the context of citations, the linear preferential attachment means that citation dynamics of the papers published in the same year has the same functional dependence and differs only in scale (we totally neglect here the stochastic character of the citation process). The difference between citation numbers of these papers is due to initial conditions, namely the number of citations that the papers garnered during first 2–3 years after publication. This is related to the number of readers which is determined by the journal's circulation. Since the majority of readers are graduate students who tend to copy once prepared reference list in all their publications, the citation lifetime of a paper that some Ph.D. student came across, is the duration of his Ph.D. stay, namely 3–5 years. Therefore, the initial impact of a paper on research groups that undertook to cite it, usually continues for 1–2 generations of the Ph.D. students, namely for 6–10 years (see Fig. 9).

If the above scenario were true for all papers, then the growth of the citation network would follow linear preferential attachment and the citation lifetime of all papers would be more or less the same. Figure 9 shows that while the citation lifetime of the vast majority of papers is indeed 6–10 years, there are quite a few papers that have much longer lifetime. We believe that these are the papers that induce "chain" reaction or cascade. Namely, the researchers can pick up such paper not by reading the journal where it was published but through the impact of this paper on other research groups. In this case the paper starts it citation career in a new research group and its citation lifetime increases by another 6–10 years. Such process of spreading the ideas is similar to epidemiological process [25, 26] and to the copying mechanism [27, 28]. It seems that the papers whose impact propagates through the cascade process are responsible for the nonlinear growth of the citation network. The fraction of such papers in the whole pool of papers determines the degree of deviation of the attachment exponent from unity. The fact that this deviation is small, indicates that only

a small fraction of papers ignites the chain reaction or cascade. Our measurements [29] indicate that these are the papers that garnered at least $\sim$ 50–70 citations at some moment in their citation career.

This cascade mechanism is specific for the citation network and it does not necessarily occurs in other networks. Therefore, the growth mechanism of the complex networks other than citation network (Table 1) can still follow linear preferential attachment.

7 Conclusions

The dynamics of citation network is driven by the nonlinear autocatalytic growth with the attachment exponent $\alpha \sim$ 1.2–1.3. The small but appreciable deviation of the growth process from linearity leads to a dramatic dynamical phase transition: papers that exceed at some stage a certain number of citations become practically immortal: their citation lifetime diverges. In the language of epidemiology these papers become endemic.

References

1. de Solla Price, D.: A general theory of bibliometric and other cumulative advantage processes. J. Am. Soc. Inf. Sci. **27**, 292 (1976)
2. Albert, R., Barabasi, A.L.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**, 47 (2002)
3. Newman, M.E.J.: The structure and function of complex networks. SIAM Rev. **45**, 167 (2003)
4. Redner, S.: Citation statistics from 110 years of physical review. Phys. Today **58**, 49 (2005)
5. Stringer, M.J., Sales-Pardo, M., Amaral, L.A.N.: Effectiveness of journal ranking schemes as a tool for locating information. PLoS ONE **3**, e1683 (2008)
6. Radicchi, F., Fortunato, S., Castellano, C.: Universality of citation distributions: toward an objective measure of scientific impact. Proc. Natl. Acad. Sci. USA **105**, 17268 (2008)
7. Petersen, A.M., Wang, F., Stanley, H.E.: Methods for measuring the citations and productivity of scientists across time and discipline. Phys. Rev. E **81**, 036114 (2010)
8. Golosovsky, M., Solomon, S.: Runaway events dominate the heavy tail of citation distributions. Eur. Phys. J. **205**, 303 (2012)
9. Krapivsky, P.L., Redner, S.: Organization of growing random networks. Phys. Rev. E **63**, 066123 (2001)
10. Dorogovtsev, S.N., Mendes, J.F.F.: Evolution of networks. Adv. Phys. **51**, 1079 (2002)
11. Malcai, O., Biham, O., Solomon, S.: Power-law distributions and Lévy-stable intermittent fluctuations in stochastic systems of many autocatalytic elements. Phys. Rev. E **60**, 1299 (1998)
12. Blank, A., Solomon, S.: Power laws in cities population, financial markets and Internet sites (scaling in systems with a variable number of components). Physica A **287**, 279 (2000)
13. Jeong, H., Neda, Z., Barabasi, A.L.: Measuring preferential attachment in evolving networks. Europhys. Lett. **61**, 567 (2003)
14. Eom, Y.-H., Jeon, C., Jeong, H., Kahng, B.: Evolution of weighted scale-free networks in empirical data. Phys. Rev. E **77**, 056105 (2008)
15. Eom, Y.-H., Fortunato, S.: Characterizing and modeling citation dynamics. PLoS ONE **6**, e24926 (2011)
16. Valverde, S., Solé, R.V., Bedau, M.A., Packard, N.: Topology and evolution of technology innovation networks. Phys. Rev. E **76**, 056118 (2007)
17. Csardi, G., Strandburg, K.J., Zalanyi, L., Tobochnik, J., Erdi, P.: Modeling innovation by a kinetic description of the patent citation system. Physica A **374**, 783 (2007)
18. Wang, M., Yu, G., Yu, D.: Measuring the preferential attachment mechanism in citation networks. Physica A **387**, 4692 (2008)
19. Tomassini, M., Luthi, L.: Empirical analysis of the evolution of a scientific collaboration network. Physica A **385**, 750 (2007)
20. Newman, M.E.J.: Clustering and preferential attachment in growing networks. Phys. Rev. E **64**, 025102 (2001)

21. Capocci, A., Servedio, V.D.P., Colaiori, F., Buriol, L.S., Donato, D., Leonardi, S., Caldarelli, G.: Preferential attachment in the growth of social networks: the Internet encyclopedia Wikipedia. Phys. Rev. E **74**, 036116 (2006)

22. Herdagdelen, A., Aygun, E., Bingol, H.: A formal treatment of generalized preferential attachment and its empirical validation. Europhys. Lett. **78**, 60007 (2007)

23. Eisenberg, E., Levanon, E.Y.: Preferential attachment in the protein network evolution. Phys. Rev. Lett. **91**, 138701 (2003)

24. Sheridan, P., Yagahara, Y., Shimodaira, H.: Measuring preferential attachment in growing networks with missing-timelines using Markov chain Monte Carlo. Physica A **391**, 5031 (2012)

25. Goffman, W., Newill, V.A.: Generalization of epidemic theory: an application to the transmission of ideas. Nature **4953**, 225 (1964)

26. Bettencourt, L.M.A., Cintron-Arias, A., Kaiser, D.I., Castillo-Chavez, C.: The power of a good idea: quantitative modeling of the spread of ideas from epidemiological models. Physica A **364**, 513 (2006)

27. Krapivsky, P.L., Redner, S.: Network growth by copying. Phys. Rev. E **71**, 036118 (2005)

28. Simkin, M.V., Roychowdhury, V.P.: A mathematical theory of citing. J. Am. Soc. Inf. Sci. Technol. **58**, 1661 (2007)

29. Golosovsky, M., Solomon, S.: Stochastic dynamical model of a growing citation network based on a self-exciting point process. Phys. Rev. Lett. **109**, 098701 (2012)