RESEARCH

Open Access

Towards a simple mathematical theory of citation distributions

Yurij L. Katchanov[*]

*Correspondence:
yurij.katchanov@gmail.com
National Research University
Higher School of Economics
(HSE), 20 Myasnitskaya Ulitsa,
101000 Moscow, Russia

Abstract

The paper is written with the assumption that the purpose of a mathematical theory of citation is to explain bibliometric regularities at the level of mathematical formalism. A mathematical formalism is proposed for the appearance of power law distributions in social citation systems. The principal contributions of this paper are an axiomatic characterization of citation distributions in terms of the Ekeland variational principle and a mathematical exploration of the power law nature of citation distributions. Apart from its inherent value in providing a better understanding of the mathematical underpinnings of bibliometric models, such an approach can be used to derive a citation distribution from first principles.

Keywords: Bibliometrics, Citation distributions, Power law distribution, Wakeby distribution

Mathematics Subject Classification: 91D30, 91D99

Background

Scholars have been investigating their own citation practice for too long. Bibliometrics already forced considerable changes in citation practice Michels and Schmoch (2014). Because the overwhelming majority of bibliometric studies focus on the citation statistics of scientific papers (see, e.g., Adler et al. 2009; Albarrán and Ruiz-Castillo 2011; De Battisti and Salini 2013; Nicolaisen 2007; Yang and Han 2015), special attention is devoted to citation distributions (see, inter alia, Radicchi and Castellano 2012; Ruiz-Castillo 2012; Sangwal 2014; Thelwall and Wilson 2014; Vieira and Gomes 2010). However, the fundamentals of the citation distribution (or CD for convenience) are far from being well established and the universal law of CD is still unknown (we do not go into details, and refer the reader to Bornmann and Daniel 2009; Eom and Fortunato 2011; Peterson et al. 2010; Radicchi et al. 2008; Ruiz-Castillo 2013; Waltman et al. 2012). Furthermore, existing bibliometric models of CDs place little or no emphasis on characteristics of the mathematical formalism itself (cf. Egghe 1998; Simkin and Roychowdhury 2007; Zhang 2013).

A mathematical theory of the CDs does not considers social citation system in its actuality. (We prefer to abbreviate social citation system to SCS; for the definition of SCS the reader is referred to, e.g., Fujigaki 1998; Rodríguez-Ruiz 2009; Rousseau and Ye 2012.)

This task is completely left to scientometricists. The mathematical theory of CDs is used to investigate a mathematical substitute instead of a real process. For this mathematical substitute, the term mathematical structure has been introduced.

The objective of scientometrics is to bridge a gap between our insights of science and our knowledge of science Mingers and Leydesdorff ([2015](#)). A mathematical theory of citation can appear as an attempt to understand the structures that constitute the bases of scientometric models. To "understand" here means to bring a bibliometric structure into congruence with a mathematical structure. The purpose of a mathematical theory is fulfilled if it provides a structure of thought objects that allows us to relate bibliometric data sets and interpret the state of affairs in science by making mathematical deductions. A scientometric model attempts to create a heuristic explanation of an empirical data set. In contrast, a mathematical theory of citation is not concerned with bibliometric data per se and strives to construct a clear and coherent framework that accurately expresses some scientometric propositions in mathematical language. In this way, opportunities emerge for applying sophisticated mathematical concepts to bibliometric phenomena. The difference between a bibliometric model and a mathematical theory of citation is more apparent than real because, although the concepts of bibliometrics can be analyzed in terms of mathematics, they cannot be eliminated in favor of the latter without losing the understanding gained by bibliometrics. In particular, a firm foundation for a mathematical theory of citation can be obtained only phenomenologically by comparing the consequences of basic mathematical statements to bibliometric data.

Motivation

We will study the axioms on which a mathematical description of SCS can be based. The author risks asserting that a mathematical theory appears to be a systematic reformulation of the problem of cumulative CDs on a purely mathematical basis. That is the main intent of this paper. Before we proceed with the analysis, we remark that there are no strong arguments leading from the bibliometric facts to the axioms. However, as we hope to show below, one can obtain additional conceptual information (relating to SCS) that is not readily available from a conventional bibliometric model by means of the axioms.

Purpose

The purpose of the research reported in this article is to provide a simple and coherent presentation of CDs based on the Ekeland variational principle. We stress the elementary variational principle governing the state of SCS and have also attempted to provide enough technical detail to create a basis for potential future studies.

Methodology

The paper addresses the construction of structural hypotheses for "how SCS works" rather than statistical inferences from bibliometric data. We accept that the continuous reproduction of a scientific inequality is a conceptual basis for almost all SCSs (cf. Bourdieu [2004](#)). An emphasis is placed on the role of the variational principle as a valid approach for describing the local behavior of an continuous SCS. We consider an SCS to obey the following scheme. Suppose an SCS is a sufficiently smooth "motion" to ensure

the consistency and the integrity of citations. In phase space, this condition is equivalent to a variational principle that produces the Euler equation for the weak form of a CD. This variational principle asserts that, for an appropriate functional, one can add a small perturbation to make it attain a minimum.

Preliminaries

A mathematical theory of CDs cannot make sense of bibliometric models of CDs. However, this theory can make sense of mathematical models; therefore, it stipulate that bibliometrics must be presented in mathematical terms. We will call a function $z \mapsto N(z)$ giving the number n of scientific papers which have been cited a total of z times a *citation distribution* (CD). By construction, we define the event ω as the value of z, $\omega := z$. Under mild assumptions, it can be assumed (somewhat non-rigorously) that, with respect to the Lebesgue measure, for any Borel set B,

$$\big(\zeta(\omega) = \omega\big)\big(\Omega = [z_{\min}, z_{\max}]\big) : \mathsf{P}(\zeta \in B) \quad = \int_B f(z)dz \propto \int_B N(z)dz,$$

ζ being a corresponding random variable (RV for short), defined on a certain probability space $(\Omega, \mathcal{B}, \mathsf{P})$. Because of this result, without any great error one can also view the quantity $N(z)$ as the probability density function (or, in abbreviated form, PDF) $f(z)$ for finding the citation process at the point z in phase space (see also Redner 1998; Gupta et al. 2005), ignoring, for now, the objection that z and n assume only non-negative integer values. We do not consider the case where supp $f(\cdot)$ is discrete separately. Although this is not mathematically rigorous, it is often useful to identify the CD $N(\cdot)$ with the PDF $f(\cdot)$ because a discrete SCS might be too complex to allow analytical results to be obtained.

All empirical CDs are different, but many have some statistical properties in common. In broad terms, power distributions of the form

$$(z \gg z_{\min}) : f(z) \propto l(z)z^{-a} \tag{1}$$

are frequently accepted without question. In the expression (1), $z_{\min}$ is a threshold value, and by $l(z)$, we denote a slowly varying function (for the precise definition, see Borovkov 2013) such that for any fixed $k > 0$, the expression $\lim \frac{l(kz)}{l(z)}$, as $z \to \infty$, is equal to 1. More concretely, owing to complexities arising from the intricate citation dynamics of papers, the age distribution of references, the role of scientific journals, etc., the CDs are quite complicated in detail. However, to a reasonable approximation, a CD can be represented (in the long-time limit of the observation period) by the relation (1) (among an abundant literature, we refer to Albarrán et al. 2011; Brzezinski 2015; Egghe 2005; Radicchi and Castellano 2015; Redner 2005; Wallace et al. 2009). The systematic study of CDs' deviations from power laws is not the subject of this paper. However, the literature on this topic is currently growing; the reader can see, e.g., Golosovsky and Solomon (2012), Golosovsky and Solomon (2014); Radicchi et al. 2012), Thelwall and Wilson (2014), Wang et al. (2013) and Yao et al. (2014).

Through the paper, we adopt the following notation: V is a real separable reflexive Banach space equipped with a norm $\|\cdot\|$, and $V^\star$ is its topological dual endowed with the natural norm $\|\cdot\|_\star$. The duality mapping between V and $V^\star$ is denoted by $\langle \cdot, \cdot \rangle$. In

addition, recall that for a nontrivial function $\phi(\cdot) : V \to \mathbb{R} \cup \{+\infty\}$ with the effective domain

$$\operatorname{dom} \phi := \{v \in V : \phi(v) < +\infty\},$$

the subset of $V^\star$

$$(\forall v \in V) : \partial\phi(v_0) = \{v^\star \in V^\star : \phi(v) - \phi(v_0) \geq \langle v^\star, v - v_0 \rangle\}$$

is said to be the subdifferential of $\phi(\cdot)$ at v_0.

In the language of $\mathsf{P}(Z \in B)$, the PDF $f(\cdot)$ is (almost everywhere) given by formal differentiating; as a result of this, a rather simple interpretation of $f(\cdot)$ can be given in the framework of Sobolev spaces $H^k(\mathbb{R})$. (For the definitions and properties of Sobolev spaces, see Maz'ya 2011.)

Unless specified, in the following, I is an open interval in $\mathbb{R}$. For technical reasons, $H^k(I)$ is a good example of the space of RVs ζ such that

$$\left(\forall \zeta \in H^k(I)\right) : \int_I |\zeta|^k \mathsf{P}(dz) < \infty.$$

In addition to the probabilistic treatment, one can say that an SCS acting on some function $\varphi(\cdot)$ yields an RV ζ. In other words, we can also state that an SCS allow us to bring one and only one well-defined RV $\zeta \in V$ into correspondence with each function $\varphi(\cdot) \in V^\star$.

Results

Because $\varphi(\cdot)$ and ζ are so fundamental in this paper, it may seem strange that we have not explicitly defined them in formal mathematical terms. As with other primitive objects of the mathematical theory, the most one can do is to give the implicit definitions by postulating the properties that hold for $\varphi(\cdot)$ and ζ.

We shall attempt now to shed some light upon the relation between the function $\varphi(\cdot)$ and the quantity ζ. Pick any $\varepsilon > 0$; then, the partial order of Bishop − Phelps on $V \times \mathbb{R}$ can be defined as follows (cf. Johnson and Lindenstrauss 2001):

$$[v_1, v_1^\star] \preceq [v_2, v_2^\star] \iff \langle v_1^\star, v_1 \rangle + \varepsilon \|v_1 - v_2\| \leq \langle v_2^\star, v_2 \rangle. \tag{2}$$

For a nonempty closed convex subset $M \subset V \times \mathbb{R}$ which is bounded below, in sense that,

$$(v^\star \in V^\star) : \inf_V \left\{ \langle v^*, v \rangle \in \mathbb{R} : \exists v \in V, \ [v, v^\star] \in M \right\} > -\infty,$$

there is a minimal element $[v_*, v_*^\star]$ in the partial order (2), according to the classical theorem of Bishop − Phelps (see, e.g., Deville and Ghoussoub 2001). In this connection, it is evident that the map $\varphi \mapsto \zeta$ is nothing but the Riesz − Fréchet isomorphism from $V^\star$ onto V. This means that we have characterization via

$$(\zeta, v)_V = \langle \varphi, v \rangle_{V^\star, V},$$

where $(\cdot, \cdot)_V$ indicates the scalar product in V. Based on this assertion, we will collect the basic properties of $\varphi(\cdot)$ and ζ in the following axioms:

A₁ A function $\varphi(\cdot)$ for $\forall \nu \in (V, \| \cdot \|)$ is a proper (dom $\varphi \neq \emptyset$), lower semicontinuous ($\varphi(\nu) \leq \liminf_{n \to \infty} \varphi(\nu_n)$ if $\nu_n \to \nu$), convex, and bounded below ($\inf_V \varphi > -\infty$) function from $(V, \| \cdot \|)$ to $\mathbb{R}_+$, satisfying the following condition:

$$(\forall \nu \in V)(\forall c \in \mathbb{R}) : \varphi(u + \nu) = \varphi(u) + \varphi(\nu) \Leftrightarrow \nu = |c|u.$$

A₂ Among all admissible ζ, the quantity ζ_* which actually describes a given CD, is assigned in such a way that the function $\varphi(\cdot)$ reaches its minimum. The axiom **A₁** hinges on a corollary Aubin and Ekeland (2006), p. 262 of the Ekeland variational principle Ekeland (1974). The term "Ekeland variational principles" refers here essentially to a result stating that the function $\varphi(\nu)$ possesses arbitrarily small perturbations

$$(\nu \in \text{dom } \varphi)(\varepsilon > 0) : \varphi(\nu) \leq \inf_V \varphi + \varepsilon,$$

such that the perturbed function

$$\big(\forall \nu \in V \backslash \{\nu_\varepsilon\}\big) : \nu \mapsto \varphi(\nu) + \varepsilon \|\nu - \nu_\varepsilon\|,$$

will have an absolute (and even strict) minimum Ioffe and Tikhomirov (1997). More precisely, there exists $\nu_\varepsilon \in \text{dom } \varphi$ and $\nu_\varepsilon^\star \in \partial \varphi(\nu_\varepsilon)$ such that

$$\big(\forall \nu \in V \backslash \{\nu_\varepsilon\}\big) : \|\nu - \nu_\varepsilon\| \leq \varphi(\nu) - \varphi(\nu_\varepsilon),$$
$$\big(\partial \varphi(\nu_\varepsilon) \neq \emptyset\big) : \big\|\nu_\varepsilon^\star\big\|_\star \leq \varepsilon.$$

Loosely speaking, if ν_ε is at least as good as ν, then ν_ε is almost the same function for which the minimum of $\varphi(\cdot)$ is almost achieved. At the same time, it is important to stress that the Ekeland variational principle does not guarantee that $\varphi(\cdot)$ attains its minimum.

The essential features of our approach include the use of Sobolev spaces. Consider now the Gelfand triple

$$V = \widetilde{H}(I) \hookrightarrow L^2(I) \hookrightarrow \widetilde{H}^{-1}(I) = V^\star,$$

where

$$\widetilde{H}(I) := \Big\{ \nu, \nu_\varepsilon \in H^1(I) : \nu - \nu_\varepsilon \in H_0^1(I) \Big\}$$

is a closed affine subspace of the second-order Sobolev space $H^1(I)$ (which is also a separable Hilbert space) and the embeddings are dense, continuous, and compact. Given a function ν_ε, we set

$$\Big(\forall \nu, \nu_\varepsilon \in H^1(I)\Big) : \tilde{\nu} := \nu - \nu_\varepsilon.$$

From now on, we will use the transformed functions $\tilde{\nu}$, but for convenience drop the tilde.

The space $V := \widetilde{H}(I)$ is endowed with the usual scalar product

$$(\forall u, v \in V) : (u, v)_V = \int_I \left(uv + u'v' \right).$$

We see that the associated norm

$$\|v\|_V^2 = \int_I \left(v'^2 + v^2 \right)$$

satisfies the axiom $\mathbf{A_1}$. Before moving on to consider ζ, however, it is tempting to slightly generalize the definition of V. Modelization of an complex real SCS requires us to introduce some additional constructions. The differential operator K,

$$\left(0 < \alpha \leq p(\cdot) \in C^1(\overline{I}) \right) \left(0 \leq q(\cdot) \in C(\overline{I}) \right)$$
$$\left(\forall v \in \widetilde{H}(I) \right) : v := -\left(p(\cdot)v' \right)' + q(\cdot)v,$$

is an isomorphic map $K : V \leftrightarrows V^\star$. We will work on the so-called energetic space $H_E(I)$ such that

$$V_E = H_E(I) \hookrightarrow \widetilde{H}(I) \hookrightarrow L^2(I) \hookrightarrow \widetilde{H}^{-1}(I) \hookrightarrow H_E^{-1}(I) = V_E^\star.$$

Hereafter we use the notation $E := V_E = H_E(I)$. To avoid overloading our presentation, we refer the reader to Zeidler (1999) for details, proofs and explanations; the interested reader can compare this approach to the one described in Kristály et al. (2010).

Thanks to the notation introduced above, the energetic space E is equipped with the scalar product given by

$$(\forall \tilde{u}, v \in E) : (u, v)_E = \int_I \left(pu'v' + quv \right). \tag{3}$$

The induced energetic norm can be written as

$$\left(\forall v \in E \right) : \|v\|_E^2 = \int_I \left(pv'^2 + qv^2 \right).$$

By definition, put

$$\left(\forall v \in E \right) : \varphi(v) = \|v\|_E^2. \tag{4}$$

That is, φ may be interpreted as measuring the average value of the weak derivatives. Recall from Attouch et al. (2014) that the statements listed below are equivalent.

S_1 There exists a unique $\zeta \in E$ such that

$$\left(\forall v \in E \right) : \left(\zeta, v \right)_E = 0. \tag{5}$$

The formula (5) reads as the "weak" Euler equation in the current setting.

S_2 ζ is obtained by

$$\min_{v \in E} \left\{ \frac{1}{2} \varphi(v) \right\}. \tag{6}$$

It is well known from the theory for variational problems in Sobolev spaces that any local minimizer of $\varphi(v)$ in the E topology is also a local minimizer of $\varphi(v)$ in the C^1 topology, and it follows in the standard way that for the quantity ζ, we have

$$\zeta = \lambda(c_1 \cosh \eta + c_2 \sinh \eta), \tag{7}$$

where

$$\eta = \int \left(\frac{q(z)}{p(z)} \right)^{\frac{1}{2}} dz, \tag{8}$$

$$\lambda = \left(p(z)q(z) \right)^{-\frac{1}{4}}, \tag{9}$$

and c_1, c_2 are constants.

To arrive at specific, relevant RV ζ one has to make an assumption in addition to the axioms $\mathbf{A_1}$ and $\mathbf{A_2}$. Now let us choose the function $z \mapsto \eta$ defined in Eq. (4) according to the formula (1), i.e. in the form of a slowly varying function (see Borovkov 2013). Therefore, we have

$$\lim_{z \to \infty} \eta(z) = o(\ln z).$$

We then introduce the function $\eta(z)$ given by

$$\eta(z) \propto \ln z. \tag{10}$$

To set up the problem, we eliminate the factor λ from Eq. (7) by rescaling the quantity ζ

$$(\lambda)^{-1}\zeta \to \zeta. \tag{11}$$

In an appropriate normalization, instead of the original RV ζ occurring in Eq. (7), we now have to deal with the renormalized RV ζ. Substituting expressions (11) and (12) into Eq. (8), we obtain

$$\zeta \propto \kappa_1 z^{-\delta} + \kappa_2 z^{\beta}. \tag{12}$$

At least for all practical purposes, it is possible to represent the relation (12) by means of a standard uniform RV U

$$\zeta \propto \kappa_1 U^{-\delta} + \kappa_2 U^{\beta}. \tag{13}$$

We will treat Eq. (13) in a broader sense — as the Wakeby distribution (WD) (for more details, see Katchanov and Markova 2015; the WD features may be found in Hosking and Wallis 2005). Introducing the continuous parameters β, γ, δ, which are called shape parameters in statistics, the continuous location parameter ξ, and the continuous scale parameter α, we may cast Eq. (13) into the following general WD form Johnson et al. (2010), p. 44–46

$$\zeta = \xi + \frac{\alpha}{\beta}\left(1 - (1 - U)^{\beta}\right) - \frac{\gamma}{\delta}\left(1 - (1 - U)^{-\delta}\right). \tag{14}$$

In a special, but very important case, when $\alpha = 0$ or $\gamma = 0$, the WD in Eq. (14) reduces to the generalized Pareto distribution (GPD)

$$\zeta = \xi - \frac{\gamma}{\delta}\left(1 - (1 - U)^{-\delta}\right). \tag{15}$$

To furnish a concrete illustration, we collected the sample of articles and reviews published in journals that put in print more than 100 documents per year and were indexed in Journal Citation Reports 2003 Science edition (Thomson Reuters). All data were downloaded from Web of Science (WoS, updated on August 8, 2013), with a 10-year time window. The number of citations z was counted as the total number of times a paper appears as a reference of a more recently published paper indexed in the Web of Science Core Collection. There are 31, 097, 160 citations among 1, 062, 961 papers. The WD, the Weibull, and the lognormal distribution were fitted these bibliometric data using the 'lmomco' R package. Goodness-of-fit was done based on Kolmogorov – Smirnov's statistic D and Anderson – Darling's statistic A. The values of the test statistics are reported in Table 1 (see also Figs. 1, 2, 3, 4, 5 and 6). Comparing the obtained values and goodness-of-fit statistics given in the Table 1, it will be seen that the WD offers a higher level of accuracy than the other probability distributions considered. We conclude that the CD obtained turns out to be similar to the WD.

Discussion

One of the most exciting and fruitful applications of mathematical methods in the natural sciences is the variational principle. The substantive aim of the present paper is the derivation of a variational principle, which makes it possible to interpret the empirical regularities of the CDs as a logical necessity. Starting from the famous Ekeland variational principle, we show that the derivation of the CDs given in this paper might be considered a step in indicated above direction. Using the variational principle (6) in the energetic space E together with empirical evidences about the existence of the slowly varying functions representing the right tail of the CDs allows us to introduce the WD (and the GPD) naturally.
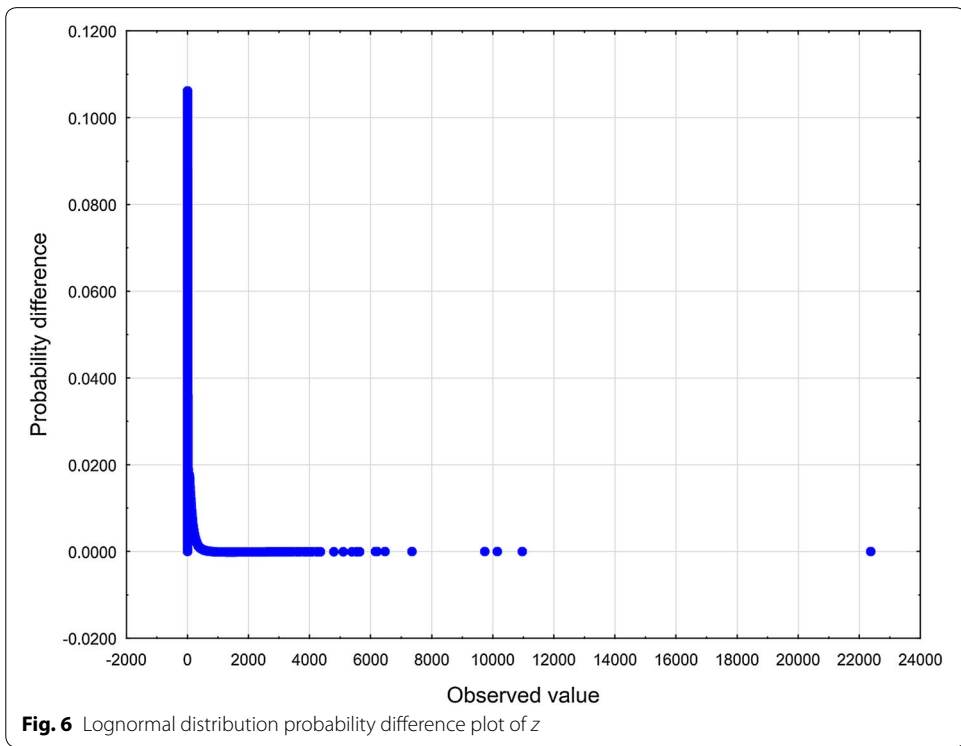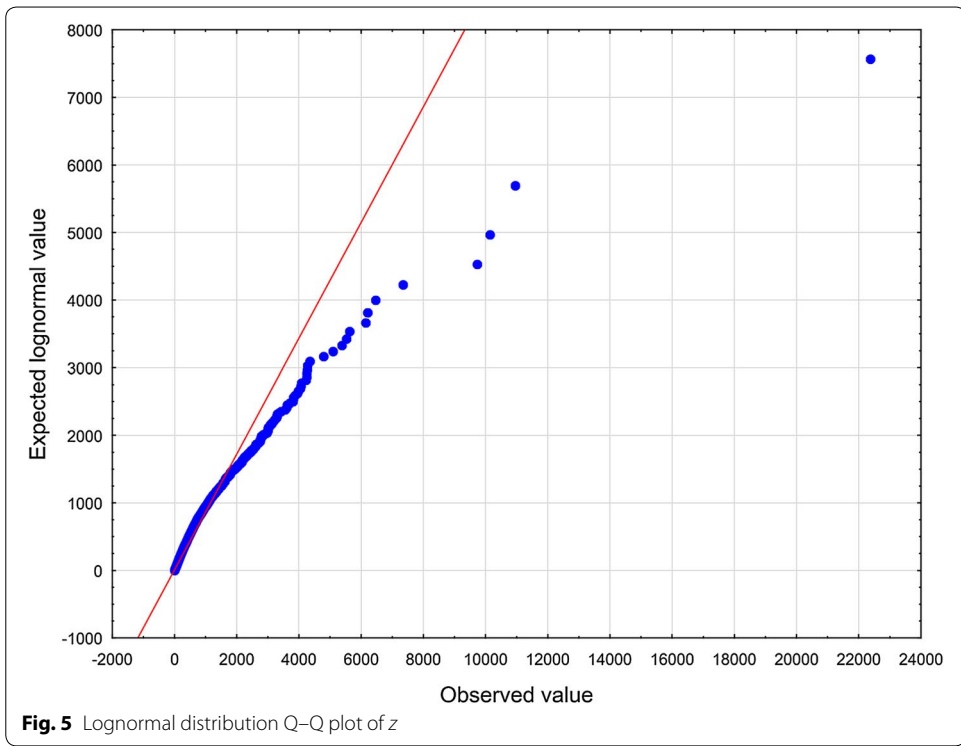
Let us stress that modest mathematical means concerning some simple facts of functional analysis yield a simple mathematical theory of CDs from which, as its consequence, concrete CDs are immediate derived. It is remarkable that a first-principles

Table 1 Goodness of fit—summary

Distribution	Kolmogorov–Smirnov $\alpha = 0.1$, Crit. val. 0.0013 Statistic D	Anderson–Darling $\alpha = 0.1$, Crit. val. 1.929 Statistic A
WD	0.0406	1449.0
Lognormal	0.1062	1.258E + 5
Weibull	0.1262	1.319E + 5

Fig. 1 WD Q–Q plot of *z*

Fig. 2 WD probability difference plot of *z*

Fig. 3 Weibull distribution Q–Q plot of *z*

Fig. 4 Weibull distribution probability difference plot of *z*

Fig. 5 Lognormal distribution Q–Q plot of *z*

Fig. 6 Lognormal distribution probability difference plot of *z*
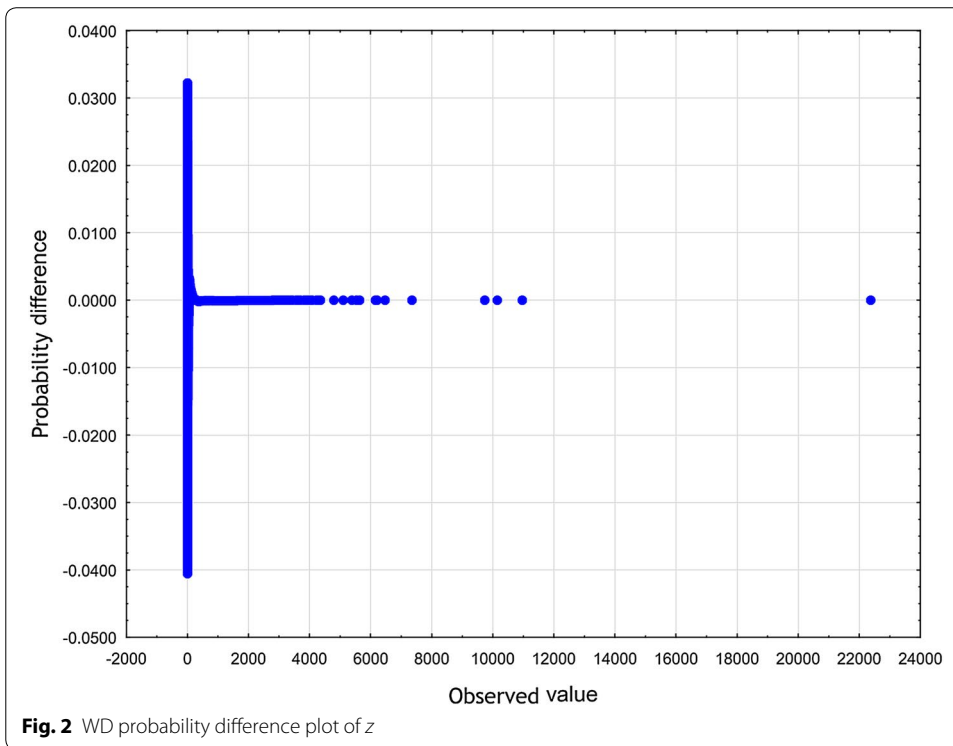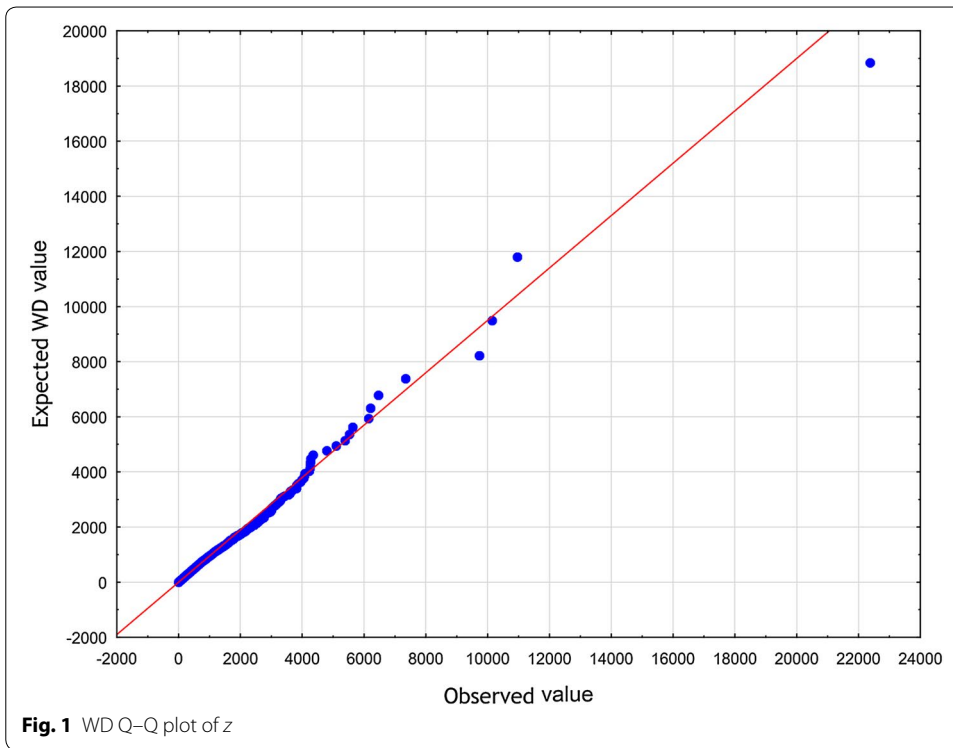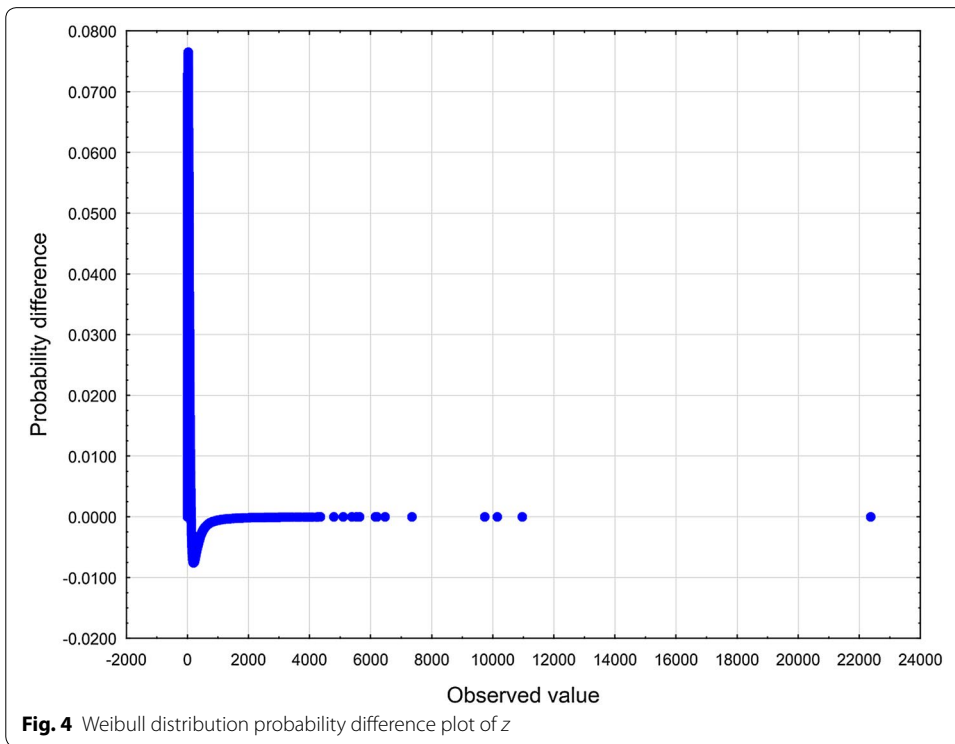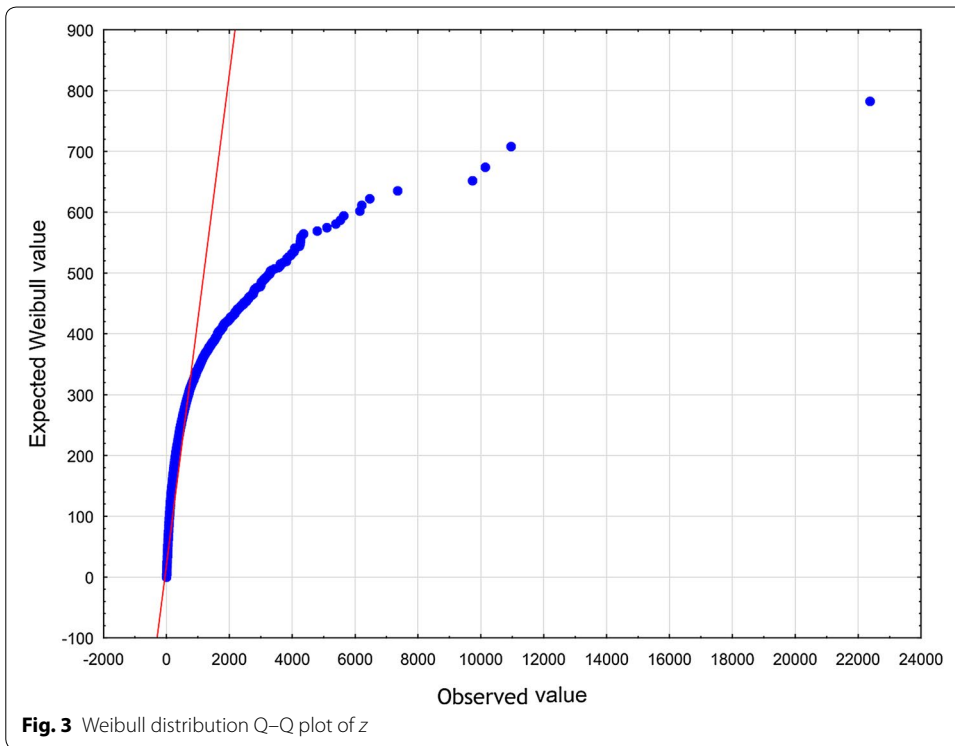
derivation of the CDs (e.g., GPD) in a bibliometric model is possible at the price of uncontrollable assumptions, which are justified a posteriori. On the contrary, in our derivation it is only assumed that Eq. (8) is relevant. This is, of course, more satisfactory. However, note that there are no proper bibliometric reasons for which the Sobolev spaces are preferred over any other, and, therefore there are also no reasons to give the vague bibliometric meaning of the consistency and the integrity of citations the mathematical form of Ekeland's variational principle.

One must bear in mind that our result refers to properties of some "pure mathematical structure". Like any mathematical result, the Eq. (13) cannot give a completely accurate description of a empirical CD. Moreover, in the mathematical theory of CDs, "by construction", we have no direct knowledge of the statistical parameters. Thus, we can only measure the parameters that index the CDs, not compute them from the axioms.

Conclusions

In summary, the approach suggested here allows an interpretation of the Ekeland variational principle in terms of the standard uniform RV, which may have some interest. It is shown that in a sufficiently "smooth" SCS a power-law tail of the static CD can appear. However, there are no grounds to consider this a mathematical model underlying bibliometric theory. At the same time, the present study may be instructive beyond the specific research site and can contribute to a mathematical theory of CDs building.

References
Adler R, Ewing J, Taylor P et al (2009) Citation statistics. Stat Sci 24(1):1. doi:10.1214/09-STS285
Albarrán P, Crespo JA, Ortuño I, Ruiz-Castillo J (2011) The skewness of science in 219 sub-fields and a number of aggregates. Scientometrics 88(2):385–397. doi:10.1007/s11192-011-0407-9
Albarrán P, Ruiz-Castillo J (2011) References made and citations received by scientific articles. J Am Soc Inf Sci Technol 62(1):40–49. doi:10.1002/asi.21448
Attouch H, Buttazzo G, Michaille G (2014) Chapter 6: Variational problems: Some classical examples. In: Variational Analysis in Sobolev and BV Spaces: Applications to PDEs and Optimization. MOS-SIAM Series on Optimization. Society for Industrial and Applied Mathematics: Mathematical Programming Society, Philadelphia, pp 219–284. doi:10.1137/1.9781611973488.ch6
Aubin J-P, Ekeland I (2006) Applied Nonlinear Analysis. Dover Books on Mathematics Series. Dover Publications, Mineola. URL: http://store.doverpublications.com/0486453243.html
Bornmann L, Daniel H-D (2009) Universality of citation distributions—a validation of Radicchi et al.'s relative indicator cfc/c_0 at the micro level using data from chemistry. J Am Soc Inform Sci Technol 60(8):1664–1670. doi:10.1002/asi.21076
Borovkov AA (2013) The basic properties of regularly varying functions and subexponential distributions. In: Probability Theory. Universitext, Springer, London, pp 665–685. doi:10.1007/978-1-4471-5201-9
Bourdieu P (2004) Science of Science and Reflexivity. University of Chicago Press, Chicago. URL: http://www.press.uchicago.edu/ucp/books/book/chicago/S/bo3630402.html
Brzezinski M (2015) Power laws in citation distributions: evidence from Scopus. Scientometrics 103(1):213–228. doi:10.1007/s11192-014-1524-z
De Battisti F, Salini S (2013) Robust analysis of bibliometric data. Stat Methods Appl 22(2):269–283. doi:10.1007/s10260-012-0217-0

Deville R, Ghoussoub N (2001) Perturbed minimization principles and applications. In: Johnson WB, Lindenstrauss J (eds) Handbook of the geometry of Banach Spaces. vol 1, Elsevier Science B.V., Amsterdam, pp 393–435. doi:10.1016/S1874-5849(01)80012-7

Egghe L (1998) Mathematical theories of citation. Scientometrics 43(1):57–62. doi:10.1007/BF02458394

Egghe L (2005) Power Laws in the Information Production Process: Lotkaian Informetrics. Elsevier / Academic Press, Kidlington, Oxfordshire. doi:10.1108/S1876-0562(2005)0000005004

Ekeland I (1974) On the variational principle. J Math Anal Appl 47(2):324–353. doi:10.1016/0022-247X(74)90025-0

Eom Y-H, Fortunato S (2011) Characterizing and modeling citation dynamics. PLoS One 6(9):24926. doi:10.1371/journal.pone.0024926

Fujigaki Y (1998) The citation system: citation networks as repeatedly focusing on difference, continuous re-evaluation, and as persistent knowledge accumulation. Scientometrics 43(1):77–85. doi:10.1007/BF02458397

Golosovsky M, Solomon S (2014) Uncovering the dynamics of citations of scientific papers. ArXiv e-prints. /hyperimage-http://arixiv.org/abs/1410.0343arXiv:1410.0343

Golosovsky M, Solomon S (2012) Runaway events dominate the heavy tail of citation distributions. Eur Phys J Spec Top 205(1):303–311. doi:10.1140/epjst/e2012-01576-4

Gupta HM, Campanha JR, Pesce RAG (2005) Power-law distributions for the citation index of scientific publications and scientists. Braz J Phys 35:981–986. doi:10.1590/S0103-97332005000600012

Hosking JRM, Wallis JR (2005) Regional Frequency Analysis: an Approach Based on *L*-moments. Cambridge University Press, Cambridge. URL:http://www.cambridge.org/9780521430456

Ioffe AD, Tikhomirov VM (1997) Some remarks on variational principles. Math Notes 61(2):248–253. doi:10.1007/BF02355736

Johnson WB, Lindenstrauss J (2001) Basic concepts in the geometry of Banach spaces. In: Johnson WB, Lindenstrauss J (eds) Handbook of the geometry of Banach spaces, vol 1, Elsevier Science B.V., Amsterdam, pp 1–84. doi:10.1016/S1874-5849(01)80003-6

Johnson NL, Kotz S, Balakrishnan N (2010) Continuous univariate distributions, vol 1, 3rd edn., Wiley Series in Probability and Statistics SeriesJohn Wiley & Sons Incorporated, New York

Katchanov YL, Markova YV (2015) On a heuristic point of view concerning the citation distribution: introducing the Wakeby distribution. SpringerPlus 4:94. doi:10.1186/s40064-015-0821-1

Kristály A, Rădulescu VD, Varga C (2010) Elliptic systems of gradient type. In: Variational Principles in Mathematical Physics, Geometry, and Economics: Qualitative Analysis of Nonlinear Equations and Unilateral Problems. Encyclopedia of Mathematics and its Applications, vol 136. Cambridge University Press, Cambridge, pp 117–145. URL:http://www.cambridge.org/9780521117821

Maz'ya V (2011) Basic properties of Sobolev spaces. In: Sobolev spaces with spplications to elliptic partial differential equations. Grundlehren der mathematischen Wissenschaften, vol 342. Springer, Berlin, pp 1–121 . doi:10.1007/978-3-642-15564-2-1

Michels C, Schmoch U (2014) Impact of bibliometric studies on the publication behaviour of authors. Scientometrics 98(1):369–385. doi:10.1007/s11192-013-1015-7

Mingers J, Leydesdorff L (2015) A review of theory and practice in scientometrics. Eur J Oper Res 246(1):1–19. doi:10.1016/j.ejor.2015.04.002

Nicolaisen J (2007) Citation analysis. Ann Rev Inform Sci Technol 41(1):609–641. doi:10.1002/aris.2007.1440410120

Peterson GJ, Pressé S, Dill KA (2010) Nonuniversal power law scaling in the probability distribution of scientific citations. Proc Natl Acad Sci 107(37):16023–16027. doi:10.1073/pnas.1010757107

Radicchi F, Fortunato S, Castellano C (2008) Universality of citation distributions: toward an objective measure of scientific impact. Proc Natl Acad Sci 105(45):17268–17272. doi:10.1073/pnas.0806977105

Radicchi F, Castellano C (2015) Understanding the scientific enterprise: citation analysis, data and modeling. In: Gonçalves B, Perra N (eds) Social Phenomena. Computational Social Sciences. Springer, Cham, pp 135–151. doi:10.1007/978-3-319-14011-7-8

Radicchi F, Castellano C (2012) Testing the fairness of citation indicators for comparison across scientific domains: The case of fractional citation counts. J Informetr 6(1):121–130. doi:10.1016/j.joi.2011.09.002

Radicchi F, Fortunato S, Vespignani A (2012) Citation networks. In: Scharnhorst A, Börner K, van den Besselaar P (eds) Models of Science Dynamics. Understanding Complex Systems. Springer, Berlin, pp 233–257. doi:10.1007/978-3-642-23068-4-7

Redner S (1998) How popular is your paper? An empirical study of the citation distribution. Eur Phys J B Condens Matter Complex Syst 4(2):131–134. doi:10.1007/s100510050359

Redner S (2005) Citation statistics from 110 years of Physical Review. Phys Today 85(6):49–54. doi:10.1063/1.1996475

Rodríguez-Ruiz O (2009) The citation indexes and the quantification of knowledge. J Educ Adm 47(2):250–266. doi:10.1108/09578230910941075

Rousseau R, Ye FY (2012) Basic independence axioms for the publication-citation system. J Scientometr Res 1(1):22–27. doi:10.5530/jscires.2012.1.6

Ruiz-Castillo J (2012) The evaluation of citation distributions. SERIEs 3(1–2):291–310. doi:10.1007/s13209-011-0074-3

Ruiz-Castillo J (2013) The role of statistics in establishing the similarity of citation distributions in a static and a dynamic context. Scientometrics 96(1):173–181. doi:10.1007/s11192-013-0954-3

Sangwal K (2014) Distributions of citations of papers of individual authors publishing in different scientific disciplines: Application of Langmuir-type function. J Informetr 8(4):972–984. doi:10.1016/j.joi.2014.09.009

Simkin MV, Roychowdhury VP (2007) A mathematical theory of citing. J Am Soc Inf Sci Technol 58(11):1661–1673. doi:10.1002/asi.20653

Thelwall M, Wilson P (2014) Distributions for cited articles from individual subjects and years. J Informetr 8(4):824–839. doi:10.1016/j.joi.2014.08.001

Vieira ES, Gomes JANF (2010) Citations to scientific articles: Its distribution and dependence on the article features. J Informetr 4(1):1–13. doi:10.1016/j.joi.2009.06.002

Wallace ML, Larivière V, Gingras Y (2009) Modeling a century of citation distributions. J Informetr 3(4):296–303. doi:10.1016/j.joi.2009.03.010

Waltman L, van Eck NJ, van Raan AFJ (2012) Universality of citation distributions revisited. J Am Soc Inf Sci Technol 63(1):72–77. doi:10.1002/asi.21671

Wang X-W, Zhang L-J, Yang G-H, Xu X-J (2013) Modeling citation networks based on vigorousness and dormancy. Modern Phys Lett B 27(22):1350155. doi:10.1142/S0217984913501558

Yang S, Han R (2015) Breadth and depth of citation distribution. Inform Process Manag 51(2):130–140. doi:10.1016/j.ipm.2014.12.003

Yao Z, Peng X-L, Zhang L-J, Xu X-J (2014) Modeling nonuniversal citation distributions: the role of scientific journals. J Stat Mech Theor Exp 2014(4):04029. doi:10.1088/1742-5468/2014/04/P04029

Zeidler E (1999) Self-adjoint operators, the Friedrichs extension, and the partial differential equations of mathematical physics. In: Applied functional analysis. Applied Mathematical sciences, vol 108. Springer, New York, pp 253–424. doi:10.1007/978-1-4612-0815-0-5

Zhang C-T (2013) A novel triangle mapping technique to study the *h*-index based citation distribution. Sci Rep 3:1023. doi:10.1038/srep01023