Journal of Informetrics 9 (2015) 466–476

Contents lists available at ScienceDirect

Journal of Informetrics

journal homepage: www.elsevier.com/locate/joi

On the growth dynamics of citations of articles by some Nobel Prize winners

Keshra Sangwal *

Department of Applied Physics, Lublin University of Technology, ul. Nadbystrzycka 38, 20-618 Lublin, Poland

ARTICLE INFO

Article history:
Received 10 November 2014
Received in revised form 24 March 2015
Accepted 24 March 2015
Available online 16 May 2015

Keywords:
Citation analysis
Citation growth curves
Unified Avrami–Weibull function
Power-law relation

ABSTRACT

The citation behavior of Nobel Prize winning articles in physics published by selected Chinese Americans, discussed before (Liu, Y. X. & Rousseau, R. (2014). *Journal of American Society for Information Science and Technology, 65(2)*, 281–289), is analyzed using a unified Avrami–Weibull equation based on concepts of formation of citation nuclei instantaneously and progressively similar to the concepts involved in the theories of overall crystallization of solid phase from a closed liquid system of fixed volume. It was found that: (1) initial concave and convex curvatures of plots of cumulative citations $L(t)$ of individual articles against citation time t are associated with the generation of citations by instantaneous and progressive citation nucleation, respectively, (2) the time constant Θ and exponent q in the unified relation are indicators that distinguish between the $L(t)$ plots with initial concave and convex curvatures for individual articles, (3) in cases of $L(t)$ plots with initial convex curvature, the data may be described by the unified relation with $q > 1$ (i.e. when nuclei are formed progressively) and/or by power-law relation, and (4) in some cases two citation regions of an $L(t)$ plots follow different nucleation mechanisms or the same mechanism with different values of the parameters of its equation.

1. Introduction

It is well known that the cumulative number $L(t)$ of various items (e.g., articles, journals and authors) in a database or citations of the publication output of an author initially increases rapidly with time, followed by a maximum level attained after a certain time, exhibiting an S-shaped curve (Avramescu, 1979; Bharathi, 2011; Egghe, Ravichandra Rao, & Rousseau, 1995; Gupta, 1990; Liu & Rousseau, 2014; Price, 1963; Sangwal, 2011a, 2012a). Various models have been developed and applied over years to describe this growth behavior of items (Price, 1963; Egghe & Ravichandra Rao, 1992; Sangwal, 2011a, 2012b; Wong & Goh, 2010). Among the different equations of various models, power-law, exponential and logistic functions are commonly used.

The nature of the S-shaped curve of the $L(t)$ data mentioned above is characterized by initial and ultimate curvatures which are convex and concave, respectively. However, using citation data of articles written by selected Chinese American Nobel prize winners in physics, Liu and Rousseau (2014) reported three types of cumulative citation $L(t)$ plots: (1) *normal* S-shaped plots with initial convex curvature followed by concave curvature, (2) *inverse* S-shaped plots with initial concave curvature followed by convex curvature, and (3) linear plots of $L(t)$ data. The above authors pointed out that these differences

* Tel.: +48 81 5384 504; fax: +48 81 5384 731.
E-mail address: k.sangwal@pollub.pl

http://dx.doi.org/10.1016/j.joi.2015.03.004
1751-1577/

in the $L(t)$ curves are representative of different types of interactions between old ideas and new insights and are associated with different diffusion and integration processes of scientific information necessary for the development of science. The authors explained different curvatures of the $L(t)$ plots phenomenologically following the concepts of position of an object, its instantaneous velocity and instantaneous acceleration used in physics. Liu and Rousseau (2012, 2013) have previously used similar concepts to interpret abrupt changes in the growth velocities of citations in the $L(t)$ plots in terms of important scientific discovery and breakthrough.

In terms of the absolute number of items per unit time (e.g. citations per year of an author; also called citation frequency and citation velocity), usually defined as $\Delta L = [L(t) - L(t-1)]$ when t is taken in years Y, initially they increase, then, after going through a maximum value, slowly decrease and finally attain a zero value with increasing time (e.g., see: Chuang & Ho, 2014; Ho & Kahn, 2014; Sangwal, 2011b, 2012a,b, 2013a, 2013b). This phenomenon of slowly decreasing growth of items with time is usually called obsolescence (Aversa, 1985; Avramescu, 1979; Egghe, 1993; Egghe et al., 1995; Gupta, 1990), aging (Egghe et al., 1995) or decay. In the case of citations of individual articles, typical curves of ΔL against citation time t are of the following types (Avramescu, 1979): (a) initially much-praised articles, (b) recognized basic work, (c) scarcely reflected work, (d) well received but later erroneously qualified work, and (e) general work. According to this classification, citation velocity curves with steep positive slope are characteristic of initially much-praised articles or articles that recognized basic work. This classification of growth of citation velocity curves has been used recently by Ho and Kahn (2014) in the discussion of top-cited single-author papers.

The decay process of items is usually explained (Avramescu, 1979; Gupta, 1990; Egghe et al., 1995; Sangwal, 2013a, 2013b) by using decreasing exponential functions. For the dependence of cumulative citations L of individual articles on time t, Avramescu (1979) proposed the empirical function

$$L(t) = L_{\max}[\exp(-\lambda t) - \exp(-m\lambda t)], \tag{1}$$

where $L_{\max}$ is the citation amplitude, λ is the age decrement time constant, and $m > 1$ is the initial increment. Sangwal (2011a) suggested that the combined growth and decay $L(t)$ curve of citations can be described as the product of fractions α_1 and α_2 of citations corresponding to growth and decay processes in the form of the empirical relation

$$\alpha(t) = \alpha_1\alpha_2 = \frac{L_1(t)}{L_{\max 1}} \times \frac{L_2(t)}{L_{\max 2}} = \frac{L(t)}{L} = \left[1 - \exp\left\{-\left(\frac{t}{\Theta_1}\right)^{q_1}\right\}\right] \times \left[1 - \exp\left\{-\left(\frac{t}{\Theta_2}\right)^{-q_2}\right\}\right], \tag{2}$$

where $L(t)$ is the number of citations at time t, $L_{\max}$ is the maximum number of possible citations, Θ is the time constant, the constant $q > 1$, and the lower indices 1 and 2 denote growth and decay processes, respectively. The time constants Θ_1 and Θ_2 and the exponents q_1 and q_2 are given by Eq. (A3). In Eq. (2) the first and the second terms in the square brackets denote the growth and decay processes, respectively. The first term follows from the progressive nucleation mechanism (PNM) advanced originally to describe overall crystallization kinetics (see below). Eq. (2) predicts a maximum value of $\alpha(t)$ at a particular value of t/Θ, but the values of α^* and $(t/\Theta)^*$ corresponding to the maximum in the $\alpha(t)$ plot are determined by the relative values of q_1 and q_2.

Recently, the process of growth and decay of yearly citations has been studied for 18 top-cited single-author papers (Chuang & Ho, 2014) and 28 top-cited single- and multi-author reviews (Ho & Kahn, 2014). Chuang and Ho (2014) observed that the citation period of the top-cited single-author 17 papers selected for the study varies in a wide range between 3 and 60 years, and the time for an article to reach citation peak varies among articles, ranging between 10 and 30 years. Ho and Kahn (2014) reported that the citation period of top-cited single- and multi-author reviews varies between 3 and 30 years whereas the time for the reviews to attain citation peaks lies between 4 and 25 years. However, Sangwal (2012b) found the citation period of 4 top-cited papers of selected Polish professors varies between 8 and 15 years, and the time for these articles to attain citation peaks is between about 4 and 10 years. It seems that, in general, both citation period and time to citation peaks for individual papers are related to their cumulative citations L and specific features such as scientific discipline and type of papers (e.g. review or original paper).

It should be mentioned that different exponential functions used for the growth of items as well as the exponential functions employed to describe their decaying behavior contain empirical constants to which it is difficult to assign any physical significance. For example, in Eq. (1) it remains unclear what factors determine the values of parameters λ and m for individual articles. Eq. (2) is also empirical because it represents simultaneous occurrence of growth and decay of cumulative citations $L_1(t)$ and $L_2(t)$, and not change ΔL in cumulative citations L per unit time.

In order to explain the growth behavior of cumulative $L(t)$ citations and $\Delta L(t)$ citations per year of individual papers of selected Polish professors, Sangwal (2012b) proposed progressive nucleation mechanism (PNM) based on overall crystallization of solid phase in a closed liquid system of fixed volume. It was found that the PNM describes: (a) the time dependence of cumulative citations L of the papers, and (b) the growth and decay of change ΔL in their cumulative citations L per year exhibiting normal and anomalous citation behavior. In the latter case, the trends of $\Delta L(t)$ curves are characterized by satisfactory and unsatisfactory fit of the $\Delta L(t)$ data in the entire citation period, respectively, by the PNM relation (see Eq. (3) and Appendix). In the case of normal citation behavior the citation period for the papers was observed to be less than 15 years and even 6–8 years in several cases. Anomalous citation behavior was observed for papers with citation periods exceeding about 15 years but there were regions of citations in the plots of ΔL against t in which the citation data could be described by the PNM relation. Normal and anomalous citation behaviors were attributed to the occurrence and non-occurrence of stationary nucleation of citations for the papers, respectively. The PNM relation has also been applied to analyze the growth of cumulative citations of individual authors (Sangwal, 2011a), cumulative articles in three randomly selected databases in humanities, social sciences and science and technology (Sangwal, 2011a), cumulative journals, articles and authors in malaria research (Sangwal, 2011a), and cumulative papers on Czochralski technique of crystal growth (Sangwal, 2013a).

The aim of the present paper is to analyze the citation behavior of Nobel Prize winning articles in physics published by selected Chinese Americans using a unified Avrami–Weibull equation based on concepts of formation and growth of citation nuclei similar to the concepts of overall crystallization of solid phase from a closed liquid system of fixed volume. The growth behavior of citations of the selected Nobel Prize winning articles and the relevant citation data have previously been described and discussed by Liu and Rousseau (2014).

2. Basic concepts and equations for growth behavior of citations

The process of growth of items (such as articles of an author, citations of individual paper of an author or authors working in a research field) with time t is a typical example of information production process. Every information production process has a source of generating items with time. The source and the generating items comprise an individual, isolated system confined to its own production space. The process is similar to that of overall crystallization of solid phase from its melt or of solute from its supersaturated solution of fixed crystallization volume V. This means that, in the case of an individual paper receiving citations (i.e. an individual paper-citations system) for example, the abstract (or imaginary) space available for citations is fixed (i.e. citation volume V) and that all citations occupying this space are similar.

In analogy with the fraction $\alpha = V_c(t)/V$ of solid phase crystallized at time t, where $V_c(t)$ is the volume of the crystallized solid and V is the total volume for crystallization, we define the fraction α of cumulative number $L(t)$ of citations received at time t by an individual paper capable of receiving a total of L_{max} citations as: $\alpha = L(t)/L_{max}$. Then the cumulative number $L(t)$ of citations at time t received by the individual paper by instantaneous and progressive nucleation may be given by the unified Avrami–Weibull equation (cf. Appendix)

$$L(t) = L_{max}\left[1 - \exp\left\{-\left(\frac{t}{\Theta}\right)^q\right\}\right] \quad (3)$$

where L_{max} is the maximum number of citations that the individual paper can receive, Θ is a time constant, and the exponent $q > 0$. As mentioned in the Appendix, depending on the values of the constant ν and the dimensionality d for overall crystallization (see Eqs. (A2) and (A3)), the values of the exponent q in the ranges $0 < q < 1.5$ and $1 < q < 4$ correspond to citation processes controlled by volume diffusion and interface transfer, respectively. The lowest values of 0 and 1 and the highest values of 1.5 and 4 for q correspond to the dimensionality $d = 0$ and 3, respectively, when the citation process occurs by instantaneous and progressive nucleation mechanisms.

According to Eq. (3) $L(t)$ increases with an increase in t for $0 < q < 1$ and approaches a saturation value L_{max}, which is attained at a lower t with an increase in q. In this case, the value of $L(t)$ steadily increases with increasing t, exhibiting a concave curve such that its curvature increases with decreasing q. In contrast to the previous situation, for $q > 1$, with an increase in t the $L(t)$ curve is convex initially, followed by practically a linear dependence in a wide range of t and later exhibiting a concave curvature. In both cases, the number $L(t)$ of cumulative citations finally approaches a saturation value L_{max} at a relatively high t when $t/\Theta \to \infty$. The convex curvature of these so-called S-shaped curves in the initial stage increases with increasing q.

Sangwal (2012b) has previously used Eq. (3) to describe the dependence of cumulative citations L of individual articles of authors on citation duration t. In the cases studied by him, it was found that $q > 1$ (i.e. convex curves), implying that citations occurred by progressive nucleation mode.

When $(t/\Theta)^q << 1$, Eq. (3) reduces to the power-law relation

$$L(t) = L_{max}\left(\frac{t}{\Theta}\right)^q = At^q, \quad (4)$$

where the constant $A = L_{max}/\Theta^q$. Power-law relation (4) also predicts concave and convex $L(t)$ curves for $0 < q < 1$ and $q > 1$, respectively, and a linear dependence for $q = 1$. However, in contrast to unified relation (3), this relation does not predict a saturation L_{max} value for any t.

Differentiation of Eq. (3) with respect to t gives the change in the generation of $L(t)$ citations with time t in the form

$$\frac{dL(t)}{dt} = L_{max}\frac{q}{\Theta^q}t^{q-1}\exp\left\{-\left(\frac{t}{\Theta}\right)^q\right\}. \quad (5)$$

Eq. (3) explains the dependence of $L(t)$ on t whereas Eq. (5) the dependence of the number of citations per unit time on time t. However, the behavior of $dL(t)/dt$ depends on the value of q. For $q = 1$ the value of $dL(t)/dt$ steadily decreases with an increase in t, but for $q > 1$ the value of $dL(t)/dt$ increases first and then decreases with increasing t, exhibiting a maximum value of $dL(t)/dt$ at a particular t_c.

The value of t_c when $dL(t)/dt$ attains its maximum value may be obtained by maximizing Eq. (5), i.e.

$$\frac{d^2L(t)}{dt^2} = L_{\max}\frac{q}{\Theta^q}t^{q-1}\exp\left\{-\left(\frac{t}{\Theta}\right)^q\right\}\left[\frac{q-1}{t}-\frac{q}{\Theta^q}t^{q-1}\right] = \frac{dL(t)}{dt}\left[\frac{q-1}{t}-\frac{q}{\Theta^q}t^{q-1}\right] = 0, \tag{6}$$

which gives

$$t_c = \left(\frac{q-1}{q}\right)^{1/q}\Theta. \tag{7}$$

Substitution of t_c from Eq. (7) in Eq. (5) gives the maximum value of $\Delta L(t)/\Delta t$ in the form

$$\left(\frac{dL(t)}{dt}\right)_c = L_{\max}\frac{q}{\Theta}\left(\frac{q-1}{q}\right)^{(q-1)/q}\exp\left[-\left(\frac{q-1}{q}\right)\right]. \tag{8}$$

According to the above equations t_c and $(dL(t)/dt)_c$ depend on the parameters q and Θ. Eq. (5) also gives the maximum citation period t_{max} for an article, which can be obtained by taking $(dL(t)/dt) = 0$ (i.e. $L(t_{max}) = L_{max}$ when the change ΔL in citations per year is zero). Then one obtains $t_{max} = \infty$.

When $(t/\Theta)^q < < 1$, Eq. (5) reduces to the power-law relation

$$\frac{dL(t)}{dt} = \frac{qL_{\max}}{\Theta^q}t^{q-1}. \tag{9}$$

Eq. (9) can also be obtained from Eq. (4). Note that according to Eqs. (4) and (9) both $L(t)$ and $dL(t)/dt$ increase with increasing t for $q > 1$. However, with increasing t cumulative citations $L(t)$ increase whereas $dL(t)/dt$ decreases for $0 < q < 1$. In this case $t_c = 0$ (see Eq. (6)) and, irrespective of the citation duration of an article, no maximum value of $(dL(t)/dt)_c$ is expected.

It should be mentioned that cumulative citations $L(t)$ in Eqs. (3) and (4) represent the position of citations $L(t)$ at time t, $dL(t)/dt$ in Eqs. (5) and (9) is the instantaneous citation velocity v, whereas d^2L/dt^2 in Eq. (6) is the instantaneous citation acceleration $a = dv/dt = d^2L/dt^2$. In informetric studies, instead of instantaneous velocity $dL(t)/dt$ of items, the term growth rate, defined as change $\Delta L(t)$ in cumulative number $L(t)$ of items per year, is usually used (e.g., see: Egghe et al., 1995; Gupta, Kumar, Sangam, & Karisiddappa, 2002). The concept of instantaneous velocity v and instantaneous acceleration a has been used before (Sangwal, 2012a, 2013c).

In statistics, the fraction $\alpha = L(t)/L_{max}$, given above by Eq. (3), is called population (or cumulative) density function F whereas its differential form $d(L(t)/L_{max})/dt$, given by Eq. (5), is called probability density function $f = dF/dt$. Note that for $q = 2$ the differential form $d(L(t)/L_{max})/dt$ of Avrami–Weibull function (3), expressed by Eq. (5), reduces to Rayleigh distribution function with $\Theta = 2\,\mu/\pi^{1/2}$, where μ is the simple mean (Nadarajan & Kotz, 2007).

3. Citation data for analysis

For the analysis cumulative citation data of Nobel Prize winning articles of the following 5 authors were used: (1) Lee and Yang (1956), (2) Aubert et al. (1974) with S.C.C. Ting as one of the coauthors, (3) Tsui, Stormer, and Gossard (1982), (4) Ashkin, Dziedzic, Bjorkholm, and Chu (1986) with S. Chu as one of the coauthors, and (5) Kao and Hockham (1966). As judged by the award of Nobel Prize for the authors, it can be argued that the ideas of the above articles are original and accepted by the scientific community. However, to assess whether unaccepted ideas can also give concave curves, cumulative citation data of the paper by Fleischmann and Pons (1989), abbreviated here as F-S, on unaccepted idea of cold fusion are also analyzed. The bibliometric data on the growth of cumulative citations $L(t)$ with citation duration t are reported by Liu and Rousseau (2014).

The data of the growth behavior of cumulative number $L(t)$ of citations to different papers and their cumulative citations $\Delta L(t)$ per year were analyzed using unified Avrami–Weibull function (3) and power-law function (4) in the forms

$$L(t) = C + L_{\max}\left[1-\exp\left\{-\left(\frac{t-t_0}{\Theta}\right)^q\right\}\right], \tag{10}$$

and

$$L(t) = C + A(t-t_0)^q, \tag{11}$$

respectively. In these equations C is the number of citations at a reference time t_0 (i.e. when $t = t_0$, $L(t_0) = C$), and the time

$$t = Y - Y_0. \tag{12}$$

where Y is the year of the cumulative number $L(t)$ of citations and Y_0 is the actual publication or extrapolated year when $L(t_0) = C$. For the analysis of the data nonlinear least-squares fitting, involving chi-square residual of the data was carried out with commercially available "Origin 9.1" package. This package yields values of the fitting parameters of an equation, their standard deviations and the corresponding goodness-of-the-fit parameter R^2. While fitting the $L(t)$ data by the above functions, the number of citations C received by an article in the first year Y_0 were taken as the reference data.

Table 1
Values of parameters of Eq. (10) for different authors.[a]

Authors	Y range	Fig., curve	C (cites)	L_{max} (cites)	Y_0 (year)	Θ (years)	q	R^2
Lee–Yang	<1970	1a, solid	2	425.2 ± 18.2	1956	3.35 ± 0.41	0.741 ± 0.071	0.9911
	>1970	1a, solid	414	10 000	1970	211 ± 713	1.601 ± 0.159	0.9882
	<1967		2	367.6 ± 13.0	1956	2.64 ± 0.23	0.863 ± 0.080	0.9928
	>1967		387	10 000	1967	218 ± 687	1.619 ± 0.156	0.9880
Ting	<1995	1a, solid	6	891.5 ± 6.1	1974	3.69 ± 0.07	0.691 ± 0.020	0.9984
	>1995	1a, solid	894	10 000	1995	320 ± 11 526	1.463 ± 0.252	0.9880
Tsui	Entire	1b, solid	2	10 000	1982	143 ± 126	1.061 ± 0.0463	0.9984
Chu	Entire	1b, solid	1	8616 ± 5700	1986	36.8 ± 9.3	3.215 ± 0.156	0.9981
Kao	Entire	1c, solid	1	249.7 ± 19.7	1966	31.49 ± 3.53	1.274 ± 0.067	0.9919
	<1995	1c, dotted	1	155.8 ± 3.9	1966	17.12 ± 0.52	1.760 ± 0.056	0.9974
	>1995	1c, dotted	143	88.0 ± 12.4	1995	13 ± 36 ± 2.11	1.524 ± 0.104	0.9961
F-P	Entire	1c, solid	155	628.4 ± 8.4	1989	5.31 ± 0.21	0.690 ± 0.017	0.9989

[a] C is the initial number of citations received by an article in the year Y_0.

An alternative procedure to obtain the values of q and Θ from the $L(t)$ data for an article is to use Eq. (3) rewritten in the logarithmic form

$$\ln[-\ln(1-L/L_{max})] = -q\ln\Theta + q\ln t, \quad (13)$$

with suitably chosen values of L_{max} such that a linear plot of $\ln[-\ln(1-L/L_{max})]$ against $\ln t$ is obtained in the $\ln t$ range as large as possible. The plot of $\ln[-\ln(1-L/L_{max})]$ against $\ln t$ gives a slope equal to q and an intercept equal to $-q\ln\Theta$. From the slope q and the intercept $-q\ln\Theta$, one can calculate Θ.

4. Results and discussion

Fig. 1 shows the growth behavior of cumulative citations $L(t)$ to articles by different authors with citation year Y. In Fig. 1a plots of the $L(t)$ data for Lee–Yang and Ting are drawn for citation regions below and above L_{max} according to Avrami–Weibull function (10) with the best-fit parameters given in Table 1. The values of L_{max} are indicated by downward arrows in the figure. In Fig. 1b plots of the $L(t)$ data for Tsui and Chu are drawn according to Eq. (10). In this case the $L(t)$ data for the two articles can be described satisfactorily in the entire citation duration by this relation, with the best-fit values of their parameters given in Table 1. In Fig. 1c plots of the $L(t)$ data for Kao and F-S are drawn according to relation (10) with the best-fit values of its parameters given in Table 1. However, it can be seen that, as in the case of the $L(t)$ data for Lee–Yang, the $L(t)$ data in Fig. 1c for Kao can be represented reasonably well by Eq. (10) in two independent $L(t)$ regions, below and above L_{max} indicated by the arrow corresponding to 1995, with the best-fit values of the parameters listed in Table 1. As seen from the best fit of the $L(t)$ data for Lee–Yang and Kao, the goodness-of-the-fit parameter R^2 for the fit in the two regions of the citation duration $L(t)$ data improves considerably in comparison with that for the fit of the $L(t)$ data in the entire citation duration. The values of the R^2 parameter in the two regions are somewhat influenced by the choice of the transition L_{max} (see fit for Lee–Yang in Table 1).

It was found that the $L(t)$ data in several cases can equally be described by power-law relation (11). The best-fit parameters for the articles analyzed in this study are given in Table 2, whereas the best-fit plots according to power-law relation (11) for the $L(t)$ data for the articles by different authors are shown in Fig. 1a–c, respectively. It may be noted from Tables 1 and 2 that only in the case of Tsui and Chu the values of the R^2 parameter obtained by power-law relation (11) are comparable with those obtained by unified Avrami–Weibull relation (10); see Fig. 1b. The values of the exponent q in these cases are also comparable. In contrast to this, the values of R^2 for the best fit of the $L(t)$ data of Lee–Yang, Ting, Kao and F-P are much higher according to Avrami–Weibull relation (10) than those for the best fit according to power-law relation (11). This observation implies that in these case the Avrami–Weibull relation describes the $L(t)$ data better than the power-law relation.

Table 2
Values of parameters of Eq. (11) for different authors.

Authors	Y range	Fig., curve	C (cites)	A (cites)	Y_0 (year)	q	R^2
Lee–Yang	<1967		2	168.8 ± 11.3	1956	0.358 ± 0.034	0.9720
	>1967		387	1.76 ± 0.73	1967	1.591 ± 0.117	0.9090
	>1963		387	1.76 ± 0.59	1963	1.566 ± 0.094	0.9330
Ting	<1995		6	459.2 ± 19.9	1974	0.231 ± 0.017	0.9655
	>1995		894	2.18 ± 21.67	1995	1.458 ± 3.983	−42.71
Tsui	Entire	1b, dash	2	58.16 ± 0.18	1982	1	0.9989
Chu	Entire	1b, dash	1	0.136 ± 0.020	1986	3.011 ± 0.047	0.9981
Kao	Entire	1c, dash	1	7.28 ± 0.64	1966	0.884 ± 0.025	0.9838
F-P	Entire	1c, dash	155	220.1 ± 17.7	1989	0.334 ± 0.031	0.9291

[a] C is the initial number of citations received by an article in the year Y_0.

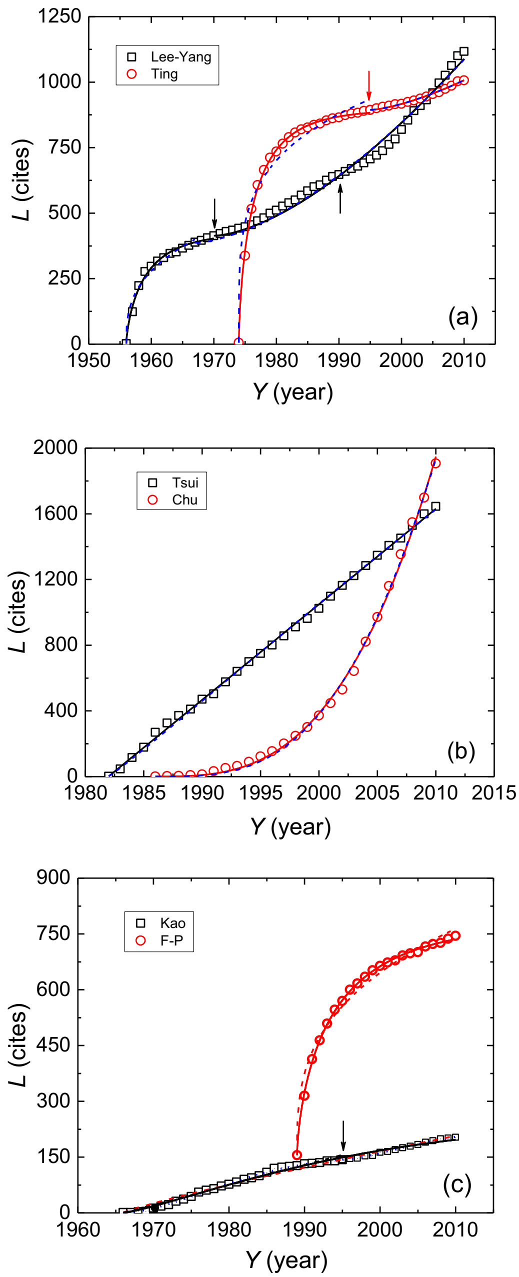

Fig. 1. Growth of cumulative citations $L(t)$ to articles by different authors with citation year Y: (a) Lee–Yang and Ting, (b) Tsui and Chu, and (c) Kao and F-S. Solid and dashed curves represent best fit of the $L(t)$ data according to Avrami–Weibull relation (11) and power-law relation (12), respectively. Solid and dotted curves in (a) and (c) are drawn for citation regions below and above L_{max}, respectively. In the plots values of L_{max} are indicated by arrows separating independent $L(t)$ regions in which growth of citation may be described by Avrami–Weibull relation (3). Values of the parameters of the relations are given in Table 1.

Fig. 2 illustrates the dependence of citations $\Delta L(t)$ per year to articles by different authors with citation year Y. Arrows in Fig. 2a indicate minimum values of ΔL for Lee–Yang and Ting. As expected from the data according to the values of $q < 1$ obtained by the Avrami–Weibull relation, with an increase in the citation time there is a steep decrease in $\Delta L(t)$ in both cases in the initial stage before the first arrow, followed by its slow increase. In the case of Lee–Yang, ΔL first increases beyond the first arrow, followed by a decrease up to the second arrow at about 1990, showing a maximum at about 1980. After the second arrow, the value of $\Delta L(t)$ increases again up to 2010. In the case of Ting also, one observes an increase after the minimum value of $\Delta L(t)$ at about 1995, indicated by an arrow. Plots of the $\Delta L(t)$ data for Tsui and Chu in Fig. 2b are drawn according to relation (5) with the best-fit values of the parameters given in Table 1. As seen from the figure, despite large fluctuations in its value, the value of $\Delta L(t)$ for Tsui essentially remains constant at 58 cites/year in the entire citation duration

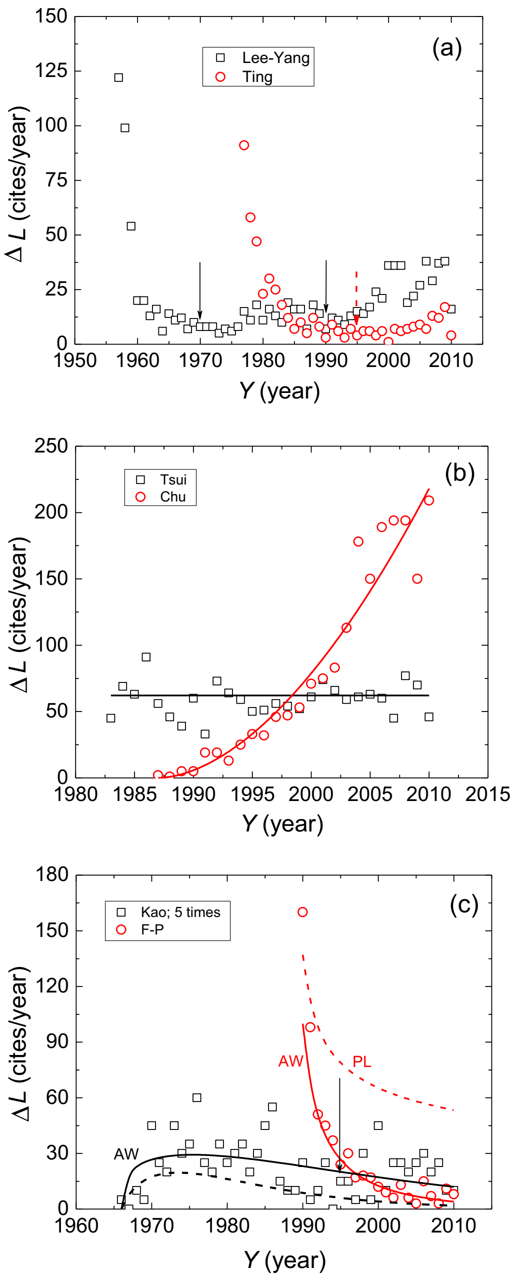

Fig. 2. Dependence of citations per year $\Delta L(t)$ to articles by different authors with citation year Y: (a) Lee–Yang and Ting, (b) Tsui and Chu, and (c) Kao and F-S. Arrows in (a) indicate minimum values of ΔL for Lee–Yang and Ting. Best-fit plots for $\Delta L(t)$ data for Tsui and Chu in (b) are drawn according to relation (5). In (c) solid and dashed curves for $\Delta L(t)$ data for Kao are drawn according to relation (5), with the best-fit parameters obtained from data in the entire and the initial regions, respectively, whereas solid and dashed curves for $\Delta L(t)$ data for F-S are drawn according to Avrami–Weibull (AW) and power-law (PL) relations (5) and (9), respectively. Values of best-fit parameters used to draw the curves are taken from Table 1. See text for details.

but that for Chu increases steadily with citation time t. In Fig. 2c the value of $\Delta L(t)$ for Kao initially increases and then, after going through a maximum at about 1975, decreases up to about 1995 and approaches a constant value. This trend is similar to that of $\Delta L(t)$ for Lee–Yang after 1970. However, the behavior of $\Delta L(t)$ for F-S is somewhat similar to that observed for Ting and Lee–Yang in the initial stages of their citations. In Fig. 2c the curves for the $\Delta L(t)$ data for Kao are drawn with two sets of the values of the best-fit constants of Avrami–Weibull relation (10) corresponding to the entire citation duration and the period up to 1995, whereas solid and dashed curves for the $\Delta L(t)$ data for F-S are drawn according to Eqs. (5) and (9) with the values of the constants obtained from best fit of the $L(t)$ data by Avrami–Weibull and power-law relations (10) and (11), respectively, given in Table 1.

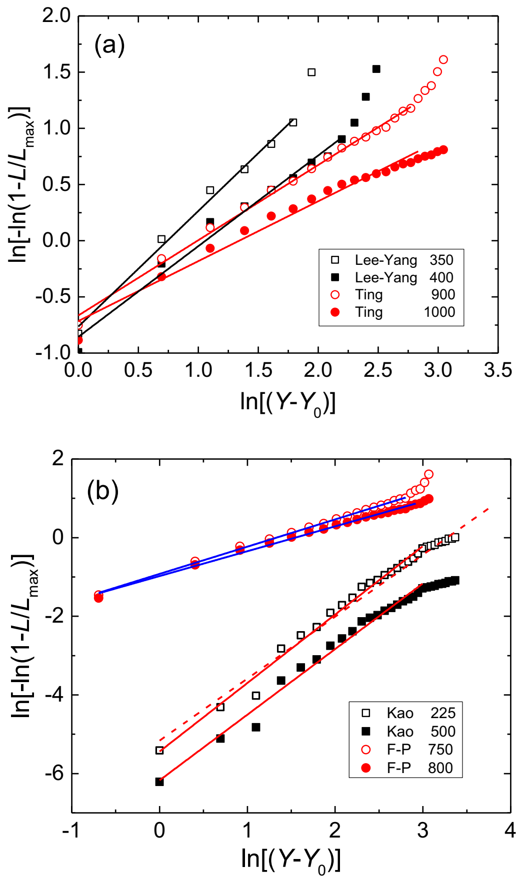

Fig. 3. Examples of dependence of $\ln[-\ln(1-L/L_{max})]$ on $\ln(Y-Y_0)$ for different authors: (a) Lee–Yang and Ting, and (b) Kao and F–P. For each article two values of L_{max}, given in the inset, were selected to establish linear dependence predicted by Eq. (11) for data of cumulative citations $L(t)$.

Fig. 3a and b shows examples of the dependence of $\ln[-\ln(1-L/L_{max})]$ on $\ln(Y-Y_0)$ according to linear relation (13) for Lee–Yang and Ting, and Kao and F-P, respectively. Plots are drawn for the $L(t)$ data of each article with two values of L_{max} selected to establish linear dependence predicted by Eq. (13). The selected values of L_{max} for the articles are given in the insets of the figures, whereas the values of the slope and the intercept of the linear parts of the plots, together with the selected values of L_{max}, are listed in Table 3.

The plots of Fig. 3 represent $L(t)$ data for not-too-large citation durations. Therefore, the procedure of obtaining the values of the exponent q and the time constant Θ from such data using Eq. (13) is reliable only when the value of L_{max} is chosen suitably such that the goodness-of-the-fit parameter R^2 attains a relatively high value. In fact, comparison of the values of q and Θ of Table 3 with those listed in Table 1 shows that they are in good agreement.

From the above discussion, it may be concluded that the initial concave and convex curvatures of the plots of cumulative citations $L(t)$ of individual articles against citation time t can be explained satisfactorily using unified Eq. (3), with the

Table 3
Best-fit values of intercept $-q\ln\Theta$ and slope q of Eq. (13) and calculated Θ for selected authors.

Authors	L_{max} (cites)	Fig., curve	$-q\ln\Theta$	q	$\ln\Theta$	Θ	R^2
Lee–Yang	350	3a, solid	0.764 ± 0.056	1.029 ± 0.005	0.742 ± 0.054	2.1	0.9906
	400	3a, solid	0.856 ± 0.068	0.807 ± 0.043	1.061 ± 0.084	2.9	0.9777
Ting	900	3a, solid	0.665 ± 0.024	0.670 ± 0.012	0.215 ± 0.008	2.7	0.9955
	1000	3a, solid	0.715 ± 0.045	0.533 ± 0.021	1.341 ± 0.084	3.8	0.9755
Kao	225[a]	3b, dash	5.161 ± 0.087	1.575 ± 0.029	1.263 ± 0.055	26.5	0.9853
	225	3b, solid	5.439 ± 0.101	1.742 ± 0.045	3.123 ± 0.058	22.7	0.9877
	500	3b, solid	6.171 ± 0.111	1.669 ± 0.049	3.698 ± 0.066	40.3	0.9839
F-P	750	3b, solid	0.922 ± 0.015	0.693 ± 0.007	1.263 ± 0.022	3.8	0.9983
	800	3b, solid	0.977 ± 0.024	0.632 ± 0.011	1.546 ± 0.038	4.7	0.9944

[a] Fit for entire data.

exponent q and the time constant Θ. This equation is based on the concept of generation of citations formed by instantaneous and progressive citation nucleation. The initial concave curvature of the $L(t)$ plots results when citations of an article are generated by the formation of citation nuclei instantaneously (instantaneous nucleation mechanism) and their subsequent growth. In this case, the values of the exponent $q<1$ and the values of the time constant Θ are relatively low (cf. Eq. (3)). However, the convex curvatures for the $L(t)$ plot results when citations of an article are generated by the formation of citation nuclei progressively (progressive nucleation mechanism) and their subsequent growth. When the initial curvature of the $L(t)$ plots is convex, the values of the exponent $q>1$ and relatively high values of the time constant Θ of Eq. (3).

It is interesting to know the value of dimensionality d of citation nuclei from the values of q obtained above. We recall that $q<1$ and $q>1$ for the growth of citations by instantaneous and progressive nucleation, respectively, and assume that the constant $\nu=1/2$ for the dissemination of scientific information contained in the articles by diffusion in the citation volume. Then one finds that, with the exception of $L(t)$ data for Chu, the dimensionality $d<2$ and $0<d<1.5$ for citations of papers occurring by instantaneous and progressive nucleations, respectively. In the case of Chu however, one finds that $d=4.4$, which is unrealistic. This suggests that a realistic value of d between 1 and 3 is possible when the dissemination process of information contained in articles involves the constant $\nu\approx 2/3$ for the generation of citations both instantaneous and progressive nucleation mechanisms.

Analysis of the $L(t)$ data for different individual papers shows that in some cases the data can equally be described by power-law relation (4). The power-law relation is a special case of unified Avrami–Weibull relation. As seen from Tables 1 and 2, the $L(t)$ data for a paper can be described by power-law relation when the time constant Θ of Avrami–Weibull relation (3) is relatively high. Examples are the data for papers by Lee–Yang, Ting, Tsui and Chu. Moreover, it should be noted that the $L(t)$ data for all of the papers cannot be described by unified Avrami–Weibull or power-law relation in the entire citation duration. In these cases two or three citation regions of an $L(t)$ plot follow different mechanisms or the same mechanism with different values of the parameters of its equation.

It should be noted that the plots of cumulative citations $L(t)$ for both Ting and F-P as a function of time t exhibit concave curvature. We recall that the ideas contained in the paper by F-P have remained unaccepted whereas those in the paper by Ting have been well-received. Therefore, it may be concluded that one cannot distinguish between initially much-praised articles and well received but later erroneously qualified work from the concave curvature of the plots of cumulative citations $L(t)$ of individual papers of different authors.

5. Summary and conclusions

The initial concave and convex curvatures of plots of cumulative citations $L(t)$ of individual articles against citation time t can be explained satisfactorily using unified Avrami–Weibull Eq. (3), with the exponent q and the time constant Θ. This equation is based on the concept of generation of citations formed by instantaneous and progressive citation nucleation. The initial concave curvature of the $L(t)$ plots results when citations of an article are generated by the formation of citation nuclei instantaneously (instantaneous nucleation mechanism). In this case, the values of the exponent $q<1$ and relatively low values of the time constant Θ of Eq. (3). However, the convex curvatures for the $L(t)$ plot results when citations of an article are generated by the formation of citation nuclei progressively (progressive nucleation mechanism). When the initial curvature of the $L(t)$ plots is convex, the values of the exponent $q>1$ and relatively high values of the time constant Θ of Eq. (3).

Analysis of the $L(t)$ data for different individual papers shows that in some cases the data can equally be described by power-law relation (4). The power-law relation is a special case of Avrami–Weibull relation.

Finally, it should be mentioned that some of the assumptions, especially analogy of citation process with overall crystallization and the dimensionality of citation nuclei, used in the derivation of unified Eq. (3) may be questioned. However, as shown in the present study, this equation is not only mathematically simple but explains reliably different aspects of growth of citation data, including trends of citation velocity and evaluation of maximum value of citation velocity at a particular time, in terms of its two parameters: the time constant Θ and the exponent q.

Acknowledgement

The author expresses his gratitude to the anonymous referee for his/her constructive comments on the manuscript.

Appendix A. Basic concepts of overall crystallization

The theory of overall crystallization is well developed in the field of crystal growth and involves simultaneous nucleation and growth of crystallites on a total of N_a active centers contained in fixed volume V (Kashchiev, 2000). Since several nuclei are formed on active centers in the volume V of the mother phase (melt or supersaturated solution), overall crystallization can occur on nuclei formed either instantaneously at the initial time $t=0$ or progressively during the entire crystallization process. These nucleation processes are known as instantaneous and progressive nucleations.

The fraction α of volume $V_c(t)$ of solid phase crystallized at time t in a crystallizing system of fixed volume V is given by the unified relation (Kashchiev, 2000)

$$\alpha(t) = \frac{V_c(t)}{V} = 1 - \exp\left\{-\left(\frac{t}{\Theta}\right)^q\right\}, \tag{A1}$$

where Θ is a time constant, and the exponent $q > 0$. Eq. (A1) is a general expression which accounts for both instantaneous and progressive nucleation, and $q > 1$ corresponding to progressive nucleation is a special case. The exponent q and the time constant Θ are given by

$$q = \nu d, \quad \Theta = \frac{1}{G}\left(\frac{V}{\kappa N_m}\right)^{1/\nu d}, \quad \text{for instantaneous nucleation,} \tag{A2}$$

$$q = 1 + \nu d, \quad \Theta = \left(\frac{1+\nu d}{\kappa G^{\nu d} J_s}\right)^{1/(1+\nu d)}, \quad \text{for progressive nucleation.} \tag{A3}$$

In Eqs. (A2) and (A3), N_m is the maximum number of nucleation centers (where $N_m = N_a$), J_s is the rate of stationary nucleation, κ is the shape factor for the nuclei (e.g., $\kappa = 4\pi/3$ for spherical nuclei) and the growth constant G is defined by

$$G = \frac{r^{1/\nu}}{t}, \tag{A4}$$

where r is the radius of the growing nucleus and the constant $\nu > 0$ is a number. Eq. (A4) describes the dependence of the radius r of the growth of *individual* nuclei on time t according to the traditional power-law relation

$$r(t) = At^{\nu}, \tag{A5}$$

where the values of ν are 1/2 and 1 for growth controlled by volume diffusion and interface transfer, respectively. In Eqs. (A2) and (A3) the parameter d denotes the dimensionality of growing nuclei. For nuclei growing in one-, two- and three-dimensions, $d = 1$, 2 and 3, respectively. However, when the nuclei do not grow, $d = 0$.

Depending on the values of the exponent ν and d, $0 < q < 1.5$ and $1 < q < 4$ for crystallization controlled by volume diffusion and interface transfer, respectively. Note that the lowest values of 0 and 1 for q, corresponding to crystallization by instantaneous and progressive nucleation mechanisms, respectively, are obtained when the nuclei do not grow with t. The time constant Θ is determined by the growth constant G and either by the maximum number N_m of crystallites in instantaneous nucleation mechanism (Eq. (A2)) or by stationary nucleation rate J_s in progressive nucleation mechanism (Eq. (A3)).

It should be mentioned that during 1939–1941 Avrami proposed and popularized Eq. (A1). It is also known as Weibull function in the literature on microbial growth (e.g. see: Aragao, Corradini, Nonmand, & Peleg, 2007; Fernandez, Salmeron, Fernandez, & Martinez, 1999) and provides a physical interpretation of its parameters. It has the form of Weibull distribution function (Weibull, 1951). In this paper Eq. (A1) is referred to as Avrami–Weibull equation in honor of its proponents.

References

Aragao, G. M. F., Corradini, M. G., Nonmand, M. D., & Peleg, M. (2007). Evaluation of the Weibull and log normal distribution functions as survival models of *Escherichia coli* under isothermal and non-isothermal conditions. *International Journal of Food Microbiology*, *119*(3), 243–257.

Ashkin, A., Dziedzic, J. M., Bjorkholm, J. E., & Chu, S. (1986). Observation of a single-beam gradient force optical trap for dielectric particles. *Optics Letters*, *11*(5), 288–290.

Aubert, J. J., Becker, U., Biggs, P. J., Burger, J., Chen, M., Everhart, G., et al. (1974). Experimental observation of a heavy particle-J. *Physical Review Letters*, *33*(23), 1404–1406.

Aversa, E. S. (1985). Citation patterns of highly cited papers and their relationship to literature aging: A study of the working literature. *Scientometrics*, *7*(3–6), 383–389.

Avramescu, A. (1979). Actuality and obsolescence of scientific literature. *Journal of American Society for Information Science*, *30*(4), 296–303.

Bharathi, G. D. (2011). Methodology for the evaluation of scientific journals: Aggregated citations of cited articles. *Scientometrics*, *86*(3), 563–574.

Chuang, K. Y., & Ho, Y. S. (2014). Bibliometric profile of top-cited single-author articles in the Science Citation Index Expanded. *Journal of Informetrics*, *8*(4), 951–962.

Egghe, L. (1993). On the influence of growth on obsolescence. *Scientometrics*, *27*(2), 195–214.

Egghe, L., & Ravichandra Rao, I. K. (1992). Classification of growth models based on growth rates and its applications. *Scientometrics*, *25*(1), 5–46.

Egghe, L., Ravichandra Rao, I. K., & Rousseau, R. (1995). On the influence of production on utilization functions: Obsolescence or increased use? *Scientometrics*, *34*(2), 285–315.

Fernandez, A., Salmeron, C., Fernandez, P. S., & Martinez, A. (1999). Application of a frequency distribution model to describe the thermal inactivation of two strains of *Becillus cereus*. *Trends in Food Science and Technology*, *10*(4–5), 158–162.

Fleischmann, M., & Pons, S. (1989). Electrochemically induced nuclear-fusion of deuterium. *Journal of Electroanalytical Chemistry*, *261*(2A), 301–308.

Gupta, B. M., Kumar, S., Sangam, S. L., & Karisiddappa, C. R. (2002). Modeling the growth of social science literature. *Scientometrics*, *53*(1), 161–164.

Gupta, U. (1990). Obsolescence of physics literature: Exponential decrease of the density of citations to Physical Review articles with age. *Journal of American Society for Information Science*, *41*(4), 282–287.

Ho, Y.-S., & Kahn, M. (2014). A bibliometric study of highly cited reviews in the Science Citation Index Expanded^{TM}. *Journal of American Society for Information Science and Technology*, *65*(2), 372–385.

Kao, K. C., & Hockham, G. A. (1966). Dielectric-fibre surface waveguides for optical frequencies. *Proceedings of the Institution of Electrical Engineers, London*, *113*(7), 1151–1158.

Kashchiev, D. (2000). *Nucleation: Basic theory with applications*. Oxford: Butterworth-Heinemann.

Lee, T. D., & Yang, C. N. (1956). Question of parity conservation in weak interactions. *Physical Review*, *104*(1), 254–258.

Liu, Y. X., & Rousseau, R. (2012). Towards a representation of diffusion and interaction of scientific ideas: The case of fiber optics communication. *Information Processing and Management*, *48*(4), 791–801.

Liu, Y. X., & Rousseau, R. (2013). Interestingness and the essence of citation. *Journal of Documentation*, *69*(4), 580–589.
Liu, Y. X., & Rousseau, R. (2014). Citation analysis and the development of science: A case study using articles by some Nobel Prize winners. *Journal of American Society for Information Science and Technology*, *65*(2), 281–289.
Nadarajan, S., & Kotz, S. (2007). Models for citation behavior. *Scientometrics*, *72*(2), 291–305.
Price, D. D. S. (1963). *Little science, big science*. New York & London: Columbia University Press.
Sangwal, K. (2011a). Progressive nucleation mechanism and its application to the growth of journals, articles and authors in scientific fields. *Journal of Informetrics*, *5*(4), 529–536.
Sangwal, K. (2011b). On the growth of citations of publication output of individual authors. *Journal of Informetrics*, *5*(4), 554–564.
Sangwal, K. (2012a). Progressive nucleation mechanism for the growth behavior of items and its application to cumulative papers and citations of individual authors. *Scientometrics*, *92*(2), 575–591.
Sangwal, K. (2012b). Application of progressive nucleation mechanism for the citation behavior of individual papers of different authors. *Scientometrics*, *92*(2), 643–655.
Sangwal, K. (2013a). Czochralski method of crystal growth in the scientific literature: an informetric study. *Acta Physica Polonica B*, *124*(2), 173–180.
Sangwal, K. (2013b). On the growth behavior of yearly citations of cumulative papers of individual authors. *Journal of Scientometric Research*, *2*(1), 30–36.
Sangwal, K. (2013c). Growth dynamics of citations of cumulative papers of individual authors according to progressive nucleation mechanism: Concept of citation acceleration. *Information Processing and Management*, *49*(4), 757–772.
Weibull, W. (1951). A statistical distribution function of wide applicability. *ASME Journal of Applied Mechanics*, *73*(9), 293–297.
Wong, C.-Y., & Goh, K.-L. (2010). Growth behavior of publications and patents: A comparative study on selected Asian economies. *Journal of Informetrics*, *4*(2), 460–474.
Tsui, D. C., Stormer, H. I., & Gossard, A. C. (1982). Two-dimensional magnetotransport in the extreme quantum limit. *Physical Review Letters*, *48*(22), 1559–1562.