Eur. Phys. J. B **27**, 445–447 (2002)
DOI: 10.1140/epjb/e2002-00176-y

THE EUROPEAN PHYSICAL JOURNAL B

The individual success of musicians, like that of physicists, follows a stretched exponential distribution

J.A. Davies[a]

University of Edinburgh, Teviot Place, Edinburgh EH8 9AG, UK

Received 23 October 2001

Abstract. Over the last five years or so, a number of studies have focussed on the distribution of ‘success’ in physics and other sciences; in these studies, ‘success’ is measured by the number of times a paper, or an author, is cited. The distribution of citations of individual papers approximates to a power-law [S. Redner, Eur. Phys. J. B **4**, 131 (1998)], while lifetime total citations of the 1120 most-cited physicists follows a stretched exponential [J. Laherrère, D. Sornette, Eur. Phys. J. B **2**, 525 (1998)]. Here, I examine the distribution of success in popular music, a field of creativity that has social structures very different from those of physics, and which is generally held to be controlled primarily by fashion. For this study, the lifetime total success of bands was measured by the total number of weeks they were in the weekly ‘top 75’ list of best-selling recordings. Like the lifetime success of physicists reported by Laherrere and Sornette, the success of the 6107 bands that appeared in the UK ‘top 75’ from 1950 until 2000 follows a stretched exponential of the form $P(x)\mathrm{d}x = c(x^{c-1}/x_0^c)\exp[-(x/x_0)^c]\mathrm{d}x$; for the music data, $c = 0.5$ and $x_0 = 9.37$.

PACS. 43.75.+a Music and musical instruments – 01.30.-y Physics literature and publications – 87.23.-n Ecology and evolution

Introduction

Over the last five years or so, several pioneering papers in this journal have examined the statistical properties of scientific activity. Science is unusual among creative fields, in that the impact of individual scientists and of their works can be measured by the number of times they are cited in subsequent research. Fortunately, a large and continuously-updated database of citations is available from the Institute of Scientific Information (ISI). Using this database, Laherrère and Sornette [1] studied the total citations to each of the 1120 most-cited physicists in the world, for the period 1981–1997 (the total period covered by the ISI database at the time). They found that the citations followed a stretched exponential distribution, so that where x is the number of citations, $P(x)\mathrm{d}x = c(x^{c-1}/x_0^c)\exp[-(x/x_0)^c]\mathrm{d}x$, with $c = 0.3$ and $x_0 = 2.7$. A few months later, Redner [2] performed a similar analysis but examined the citations per paper rather than total citations in the life of an author, and claimed that the distribution of citations for leading rank papers followed a simple power law, $N(x) \approx x^{-3}$, though this relationship did not hold as well in the region of relatively small x, where a stretched exponential provided a better description [3]. The $N(x) \approx x^{-3}$ distribution described in Redner's per-paper analysis and the stretched exponential in Laherre and Sornette's per-author analysis are not as incompatible as they appear at first sight, for the per author analysis is effectively a sum over all papers of nth convolution of the per-paper distribution. $N(x) \approx x^{-3}$ is unstable on addition and belongs to the domain of attraction of the Gaussian law; upon convolution it will slowly converge through a stretched exponential to a pure Gaussian (but the finite number of papers analysed here prevents this convergence to Gaussian proceeding all the way so, for the data set available, Laharerre and Sornette observed a stretched exponential).

As well as providing some disquieting information, such as the fact that that almost 47% of papers are simply not cited and are therefore arguably pointless, statistical studies of citation distributions raise interesting questions about the nature of scientific endeavour. In particular, they raise the question of why citation distributions should follow power laws or stretched exponentials at all. Commentators have highlighted various possible reasons for this behaviour. Sornette and Zajdenweber [4], and later Buchanan [5], for example, connect them to the paradigm shifts first described by Kuhn [6]. Kuhn proposed that the progress of science alternates between ‘normal science’, within an established set of ideas or paradigm, and occasional cataclysmic paradigm shifts. Paradigm shifts happen when an accumulation of data that do not fit an existing paradigm stimulates the formation of a revolutionary idea, which completely changes our view of the

---

[a] e-mail: `jamie.davies@ed.ac.uk`

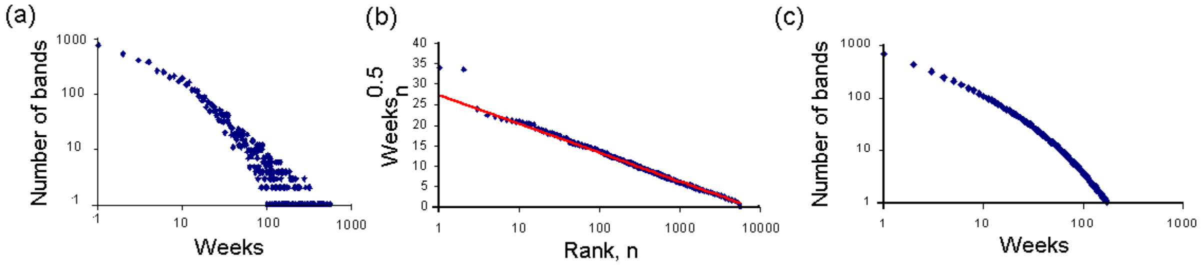

Fig. 1. (a) Log/Log histogram of the number of bands that achieved 1,2...n weeks in th UK top 40; (b) Plot of ranked musician's data, $(\text{weeks})^c$ ($c = 0.5$) for each band plotted against the log of that band's rank; the data fall on a straight line ($R^2 = 0.996$). (c) Reconstruction of the curve shown in part (a) using $c = 0.5$, $x_0 = 9.37$ in equation (1).

world; a much-quoted example is the rise of quantum mechanics. Buchanan suggests that paradigm shifts happen at all scales, from small ones that affect the view of only one tiny, specialised field, to huge ones that alter the whole of physics. In this view, the many papers that cause only small shifts attract few citations, while the few that cause great shifts attract many. In his book, Buchanan draws an explicit parallel between the distribution of the impact papers have when added to the sum of scientific knowledge, and the power-law distribution of the impacts (avalanche sizes) caused when a grain of rice is added to a pile of rice in a critical state [7]. Similar analogies could be made between the distribution of fault displacements that relieve accumulating strain in the Earth's crust, which follow a stretched exponential distribution [1], and paradigm shifts that relieve the "strain" of data that don't fit existing theories.

The distribution of citations might also arise from the internal organisation of science. Citations are made only by scientists who are themselves writing papers, so the whole enterprise forms a closed network. The properties of scientific collaboration networks can be studied using bibliographical databases. Newman [8,9] has studied such networks, and has reported a power-law relationship in the numbers of authors per paper, the exponent of the power law varying between different fields of science. The numbers of collaborators per author during a five year period also fitted a power law in some fields, though in other areas of science there was some curvature in a log-log plot. Furthermore, the distances between random pairs of scientists were small and scaled logarithmically with the total number of scientists in the network; they are therefore "small world" networks in the sense of Milgram [10].

In trying to understand more about the origin of the power law and stretched exponential distributions of scientists' impacts in their fields, it will be helpful to make a comparison with the other fields of creativity. The choice is limited by the requirement to have good quantitative data, but one field for which such data are available is that of popular music. This choice also has the advantage that music would seem to have little in common with physics.

The distribution of success in music

To analyse the distribution of the impacts of bands in the world of popular music, I used as a database the weekly lists of the 'Top 75' best-selling recordings in the UK popular music charts between the years 1950 and 2000 [11]. Impact was measured as the total number of weeks a band was in the 'Top 75' (the period 1950–2000 is much longer than the longevity of an average band). It can be noted in passing that measuring longevity of a band in the top 75 is reminiscent of measuring persistence phenomena in physical systems (see [12] for review), albeit with the constraint, for music, that a band's leaving the top 75 is always accompanied by another entering. This focus on impact per band, rather than per song, makes this study more comparable to the science citation study of Laherrère and Sornette [1] than to that of Redner [2]. The implied selection for only those bands good enough to have ever managed to be in the 'Top 75' also makes this study more like that of Laherrère and Sornette, who studied only the 1120 most cited physicists, than Redner's analysis of complete bibliographic databases; this selection is forced on me by the absence of data for bands of very low impact who never entered the Top 75. The total number of bands in this study is 6107.

A log-log plot of a histogram of the number of bands in the 'Top 75' for x weeks, against x, is shown in Figure 1a. The distribution falls along a fairly tightly-bounded line, which is reminiscent of a power-law though it is obviously too curved in to actually be a power-law throughout the range of the data. Since Laherrère and Sornette [1] found the life-long impact of individual physicists to follow a stretched exponential distribution, the musicians' data were analysed according to their method. The general form of a stretched exponential is;

$$P(x)\mathrm{d}x = c(x^{c-1}/x_0^c)\exp[-(x/x_0)^c]\mathrm{d}x, \text{ where } c < 1 \quad (1)$$

with a cumulative distribution;

$$P_c(x) = \exp[-(x/x_0)^c]. \quad (2)$$

In the method of Laherrère and Sornette [1], bands are rank ordered, the band with the highest impact having rank 1. From equation (2), it follows (see Ref. [1]) that within a rank-ordering plot, a stretched exponential gives rise to a straight line if the impact of each band, W_n, is raised to the power c and plotted against $\log(n)$;

$$W_n^c = -x_0^c \log(n) + b. \tag{3}$$

Figure 1b shows such a plot for ranked musicians' data; with the exception of the two highest ranking data points (Elvis Presley and Cliff Richard), the data all fall on or very close to a straight line with $c = 0.5$ (goodness of fit, $R^2 = 0.996$, even including those two outlying points). The gradient of the slope yields $x_0 = 9.37$. Figure 1c shows a 'reconstruction' of the original curve (Fig. 1) using these values of c and x_0 in equation (1). Because power-laws are so frequently believed to be the best description of all impact distributions in social and natural sciences [5], it is interesting to assess goodness of fit of these data to a power law (linearity on a Zipf plot) instead of the stretched exponential described above. Attempting to fit the data to a power-law yields an R^2 of only 0.842, much less than the $R^2 = 0.996$ for the stretched exponential.

Discussion

This study shows that, like citation impacts of physicists, the commercial impacts of musicians follow a stretched exponential distribution. There are, however, exceptions to this rule; the two highest-ranking data points are outliers with impacts higher than would be predicted by the general distribution. The tendency of the very highest ranking data points to be high-flying mavericks in otherwise well-behaved stretched exponential distributions has been described before in the context of stock market crashes, rupture of solids and even of human parturition [13]; the presence of high-ranking outliers is often called the "King effect" (a name strangely appropriate for Elvis' data). It has been suggested [14] that high-ranking outliers are caused by amplifying processes that can create orderly behaviour in normally chaotic systems. In the context of stock market crashes, Sornette suggested that a system of traders who are influenced by their neighbours can, under the right circumstances, propagate islands of local imitation into global cooperation and this precipitate a large scale "crash". Such behaviour is atypical of the general distribution of stock price adjustments which are usually the result of a system in which global order is absent. Music buyers too are influenced by their fellows, and the abnormally high impact of the two most popular artists suggests that the system of music buyers may also be constructed so that local imitation may occasionally generate global order.

It is interesting that the impacts of both physicists and musicians have similar distributions, for the activities of physics and popular music have few obvious similarities. The citation network of physics, for example, is closed; scientists are cited only by each other. This stands in contrast to music, where the output of professional musicians is bought mainly by people who do not themselves produce recordings. Similarly, while music may have its own version of paradigm shifts (for example the yielding of Swing to Rock'n'Roll, of Disco to Hip-hop), the pioneers of new forms tend not to be as commercially successful as those who capitalize on an already-obvious trend. When the record-buying masses do become interested in a new form of music, they are not compelled to buy a copy of the recording that originated it, but physicists are compelled to cite seminal papers; the link between pioneering work and measured impact is therefore much weaker in music than in science.

What (apart from dress sense!) might physicists and rock musicians have in common? One possibility is the presence of key 'gatekeepers'. In order to be cited, physicists need not only to produce interesting work, they also need to have it published by a journal – in order to be cited extensively, it helps to be published in a 'big-name' journal that is read widely. As journal editors are key gatekeepers in science, so talent scouts and radio stations are key gate-keepers in music, for no new band will sell large numbers of records unless their music is broadcast, preferably widely. In music, these gate-keepers are widely held to control fashion, so that success of a band depends as much on whether they are in tune with current fashions as on the intrinsic musical merit of a song. Controversially, this may also be true of physics; editors, conference organizers and funding councils are well-placed to dictate scientific fashion (either cryptically, or explicitly in strategic initiatives).

It would be interesting now to find fields of creativity that do *not* show a stretched exponential distribution of success, in order to better identify the characteristics of ones that do. Meanwhile, we can all tell our students that a career in physics is just like a life in rock'n'roll...

References

1. J. Laherrère, D. Sornette, Eur. Phys. J. B **2**, 525 (1998)
2. S. Redner, Eur. Phys. J. B **4**, 131 (1998)
3. C. Tsallis, M.P. de Albuquerque, Eur. Phys. J. B **13**, 777 (2000)
4. D. Sornette, D. Zajdenweber, Eur. Phys. J. B **8**, 653 (1999)
5. M. Buchanan, *Ubiquity* (Weidenfield, Nicholson, London, 2000)
6. T. Kuhn, *The structure of scientific revolutions* (Chicago University Press, 1962)
7. V. Frette, K. Christensen, A. Malthe-Sorenssen, J. Feder, T. Jossang, P. Meakin, Nature **379**, 49 (1996)
8. M.E.J. Newman, Phys. Rev. E **64**, 016131 (2001)
9. M.E.J. Newman, Phys. Rev. E **64**, 016132 (2001)
10. S. Milgram, Psychol. Today **2**, 60 (1967)
11. *Guinness British Hit Singles*, 13th edn. (Guinness Publications, 2000)
12. S.N. Majumdar, `cond-mat 9907407` (2000)
13. A. Johansen, D. Sornette, Eur. Phys. J. B **1**, 141 (1998)
14. D. Sornette, Proc. Natl. Acad. Sci. USA **99**, 2522 (2002)