Information Processing and Management 48 (2012) 791–801

Contents lists available at SciVerse ScienceDirect

Information Processing and Management

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

Towards a representation of diffusion and interaction of scientific ideas: The case of fiber optics communication

Yuxian Liu[a,b], Ronald Rousseau[b,c,d,*]

[a] Tongji University, Library of Tongji University, Siping Street 1239, 200092 Shanghai, PR China
[b] University of Antwerp, IBW, Venusstraat 35, B-2000 Antwerpen, Belgium
[c] KHBO (Association K.U. Leuven), Faculty of Engineering Technology, Zeedijk 101, B-8400 Oostende, Belgium
[d] K.U. Leuven, Dept. Mathematics, Celestijnenlaan 200B, B-3000 Leuven (Heverlee), Belgium

ARTICLE INFO

Article history:
Received 19 March 2011
Received in revised form 2 December 2011
Accepted 16 December 2011
Available online 9 January 2012

Keywords:
Informetrics
Representations
Fiber optics communication
Diffusion of scientific ideas
Evolution of science
Innovation
Academic space

ABSTRACT

The research question studied in this contribution is how to find an adequate representation to describe the diffusion of scientific ideas over time. We claim that citation data, at least of articles that act as concept symbols, can be considered to contain this information. As a case study we show how the founding article by Nobel Prize winner Kao illustrates the evolution of the field of fiber optics communication. We use a continuous description of discrete citation data in order to accentuate turning points and breakthroughs in the history of this field. Applying the principles explained in this contribution informetrics may reveal the trajectories along which science is developing.

1. Introduction

Informetrics can be defined as the study of the quantitative aspects of information in any form, not just records or bibliographies, and in any social group, not just scientists (Tague-Sutcliffe, 1992). It is a multifaceted approach that can be used to depict the structure and development of science (van Raan, 2008). To achieve such an ambitious aim one not only needs the graphical skills to display the structure of science (Boyack, Klavans, & Borner, 2005), but one also has to record the development of science during and between its intellectual development. This development occurs as the result of many influences and interactions, leading to diffusion and transformation of the original ideas. We imagine that these interactions during the diffusion process happen in *academic space*. Academic space is the "abstract intellectual landscape" of the scientific system, it is what happens in scientists' heads when they think, read and interact with colleagues and with the outside world. What is the role that citation analysis can play in the description of this process? We are convinced that its role can be essential. Hence our contribution should be seen as a description of a representation between academic space and citation space.

Studies of diffusion of ideas through citations and patterns of diffusion are standard topics in the field of the information sciences. Avramescu was one of the first to use the ideas of the physical diffusion process (conduction of heat in solids) to study and explain Bradford's law of scattering (Avramescu, 1975, 1980) while (Barnett, Fink, & Debus, 1989) study the dissemination of information and explicitly refer to Rogers' innovation theory (Rogers, 1962). Among other things they stress

* Corresponding author at: KHBO (Association K.U. Leuven), Faculty of Engineering Technology, Zeedijk 101, B-8400 Oostende, Belgium.
E-mail addresses: yxliu@tongji.edu.cn (Y. Liu), ronald.rousseau@khbo.be (R. Rousseau).

0306-4573/$ - see front matter
doi:10.1016/j.ipm.2011.12.001

the need of longitudinal data as provided by citation indices. In the introductory article of a special issue on bibliometric methods for the study of scholarly communication, Christine Borgman (1989) states that bibliometrics may be used to trace the evolution of an idea within and across disciplines, hence referring to the role of citation diffusion studies in scholarly communication. Le Coadic (1987) rejects the use of purely physical diffusion models in the information sciences as these physical models are conservative, in the sense that the quantity of diffused energy is conserved during physical processes. This assumption, however, does not hold for social diffusion processes because if one scientist 'gives' an idea to a colleague, she has not lost her idea, as it is assumed in physical theories.

One of the techniques to model the diffusion process is the use of theoretical epidemic models. In this approach the diffusion of ideas in populations of scientists is modeled via models derived from or inspired by epidemics and their description (Kiss, Broom, Craze, & Rafols, 2010). Though such models often fit data very well, in our opinion they do not describe the real diffusion of ideas among scientists over time, let alone delve deeper into how interactions of scientific ideas happen during the diffusion process. This leads to the research question of finding an adequate framework to describe actual knowledge diffusion over time. This question is of particular interest. Already in the eighties Schubert and Glänzel (1986) raised this problem in a citation context. They mentioned that the ideal temporal characterization of the citation process would require tracking the full citation "life histories" of single papers. Without a proper methodology this is, however, difficult to do. Although we will use elementary notions taken from physics, such as the notion of a velocity, we agree with Le Coadic that purely physical theories cannot be applied to the diffusion of scientific ideas.

Citations can be seen as historic remains left by interactions of different ideas related to the object of study (Smith, 1981). Hence, our aim is to recover and analyze the interaction of scientific ideas using citation analysis. Yet, the notion of interaction between different ideas is so abstract that we cannot grasp it directly. Interactions start in scientists' mind and as we are no neuroscientists we have to take another road instead. Although it stays difficult to measure such interactions, we think it is possible, and this for two reasons. The first is that these interactions do exist even if we cannot see them directly (Small, 1978, 1987); the second reason is that these interactions change the diffusion process of ideas in intensity and direction (different fields), and these two aspects can be detected (Liu & Rousseau, 2010). Moreover, knowledge originating in different fields may converge to a coherent new knowledge body, which is another detectable feature (Liu et al., in press). We would like to compare our approach to the bubble or cloud chamber method used in nuclear physics. Interactions of charged particles and resulting new particles cannot be seen directly, yet through such devices, placed in a magnetic field, their characteristics can be detected via the curvature of the resulting tracks.

Small (1978) argued that over time highly-cited documents become symbols for the ideas they contain. In this view referencing becomes a labeling process. In our contribution we follow Small's example and focus on two articles illustrating the evolution of the field which they initiated. Concretely, we study Kao and Hockham's Nobel Prize winning article (Kao & Hockham, 1966), as a symbol for the whole field of fiber optics. We like to stress that we do not study just one article, but the article that started a field and which led its main author to a Nobel Prize. In whatever way one would like to see it, the Kao–Hockham article is without any doubt a concept symbol. As a second case we briefly consider the Fleischmann–Pons article on cold fusion.

Our article is further organized as follows. First we describe the citation data we use. Next we recall how splines can be used to give a continuous, even differentiable, description of discrete data, followed by a short history of the field of fiber optics communication. We describe what we mean by movement in academic space and mention some factors influencing the academic movement process, leading to the notions of instantaneous velocity and acceleration. All this is brought together in an analysis of different phases in the history of fiber optic communication. It ends with proposing different uses of our theory and a conclusion. The contents of this article are based on Liu's doctoral dissertation (Liu, 2011).

2. Charles Kuen Kao, his Nobel Prize winning article and citation data

Charles Kuen Kao (Gāo Kūn in pinyin) received the Nobel Price (actually one half) in Physics 2009 for his contribution to the development of low-loss optic fiber used in optical fiber communication systems (Class for Physics of the Royal Swedish Academy of Sciences, 2009). We study citations to his most important, Nobel Prize winning article:

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

Together with a reprint (published in 1986, but which received very few citations) it received 201 citations during the period 1966–2009. As Kao received the Nobel Prize in the year 2009, this period is not 'tainted' by citations resulting from this honor. As we intend to find out how citations to Kao's article reflect the evolution of the field of fiber optics, we only kept citations given in 'normal' articles, proceedings and letters, and removed all other citations such as those in review articles and editorials. This leads to a total of 131 citations. Removing review articles and editorials was done because we want to study the actual development of a field and do not want to include mere recapitulations and introductions to the subject. A reviewer pointed out that the ideas contained in the cited article also are diffused by review articles and editorials. This is correct: these types of articles can even act as accelerants or catalysts. It so happened that we started our investigations including reviews and editorials but decided to remove them in the final analysis as they did not change the diffusion curve in a tangible way and because we wanted to study a 'pure' situation. We admit, though, that including these types of articles is a valid option. Citation data are shown in Table A1 (see Appendix A). One column shows the cumulative number of citations received by this article as collected from the Web of Science, while another column pertains to the same data but after removal of non-original publications, i.e. reviews, etc.

3. Discrete data representation with spline curves

As explained in an earlier publication (Liu & Rousseau, in press) a spline-fitted curve cannot only represent original discrete data in a continuous, even differentiable way, but can also display the change of information included in the original data. As all information about these changes resides in the original data the use of techniques from calculus can revive these and reproduce the hidden original information. Indeed, one may say that citation data have a kind of "holographic" property (Liu & Rousseau, 2011b). The use of spline functions offers a middle way between connecting discrete data by line segments and providing an overall best-fitting curve. As such this approach may lead to a better way to study informetric data, such as those used in the study of the development of science and technology. Adapting the spline-fitted function for the representation of discrete cumulative citation data we are able to calculate first and second order derivatives of the resulting curve. In this way we study how the citation diffusion process changes over time. The resulting curve of cubic spline interpolations is shown in Fig. 1. A small program to find these interpolating splines was written by us, based on (Rousseau, 1997).

Fig. 2 shows the spline-fitted cumulative citation curve, together with a best-fitting linear line (the trend line). This shows that Kao's cumulative citation data consist of consecutive S-shapes (from 1966 to 1986 one S shape, from 1986 to 2009 a converse S shape). There are some fluctuations, but their extent is rather small and can be ignored in a first approximation. Neglecting these small fluctuations, the curve can be divided into three phases. From 1966 to 1975, it is largely convex, from 1975 to 1994, it is slightly concave, and from 1994, after some oscillations, it continues parallel to the trend line. In order to have a proper view on these fluctuations, we consider a detrended curve (Shiavi, 1991). We recall that detrending consists of finding first a trend line and then calculating the difference between the observed data and the trend line. This results in a detrended description of the data. The detrended curve for Kao's data is shown in Fig. 3. Comparing with the trend line the year 1969 is the minimum, 1986 is the maximum; the years 1975 and 1995 are inflection points.

4. History of fiber optics communications and key points in the spline curve

If you really look at it, I was trying to sell a dream...
Charles K. Kao (from Hecht, 1999)

The complete history of fiber optics is described in (Hecht 1999; 2004). It is roughly divided into five periods, corresponding to early developments and four generations of fiber-optic communication systems. The first period begins with the Kao–Hockham article (Kao & Hockham, 1966) leading to the first development of optical fiber by Corning Glass Works in 1970 and ends in 1975. This is the experimental and academic research phase. From 1975 on, the field of fiber optic

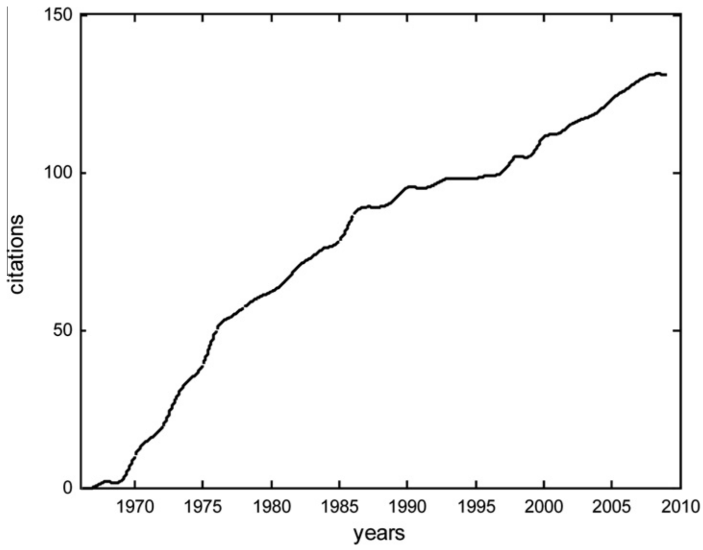

Fig. 1. Kao's cumulative citation data with spline interpolations.

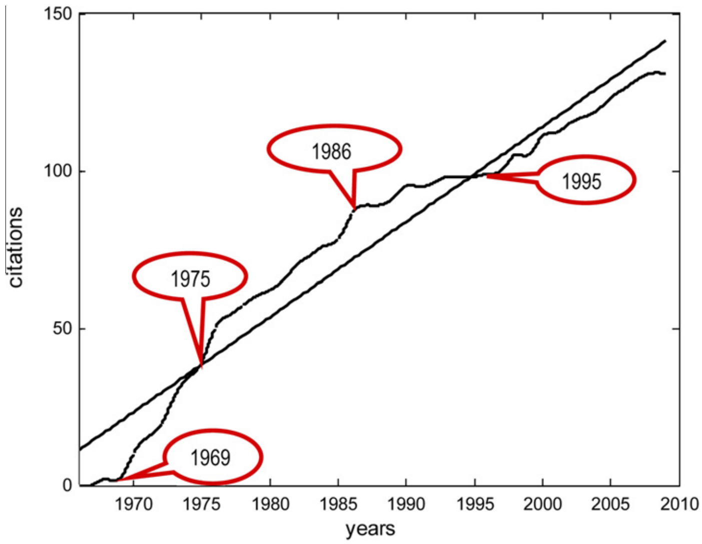

Fig. 2. Spline-fitted cumulative citation curve, best-fitting linear curve and indications of important years.

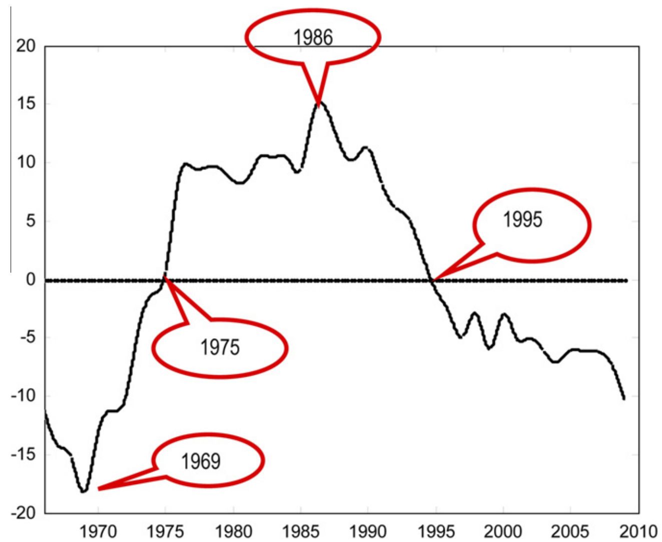

Fig. 3. Kao's cumulative citations: detrended form.

communication entered into its application period. Yet there were still two major obstacles restricting the capacity of fiber optic communication systems: one was related to the bandwidth that could be used to transmit data, and the other was the limited transmission distance of a fiber-optic communication system because of fiber attenuation and distortion. The former obstacle was overcome by enlarging bandwidth step by step. Once the technology that can transfer large bandwidths was developed, the production of a new generation of fiber optic communication system began. Commercial wavelength-division multiplexing systems, a transmission technology used in fiber optics that dramatically increases the data transmission capacity of fiber optics by dividing bandwidth based on wavelength, was introduced in 1996.

The other obstacle was eliminated by using opto-electronic repeaters. These repeaters convert the signal into an electrical signal, and then use a transmitter to send the signal again at a higher intensity than it was before. This procedure is, however, complex and costly. In 1987, David Payne and his group at Southampton (Mears, Reekie, Jauncey, & Payne, 1987) invented the erbium fiber amplifier (EFA), which amplifies the optical signal directly without having to convert the signal to the electrical domain. This technology made it possible, among other things, to install trans-atlantic and trans-pacific underwater cables with a capacity equivalent of 600,000 telephone circuits and this is at very low prices (Gambling, 2000). Before 1994, all efforts of fiber optic communication were based on improving the commercial telephone and television systems. In the 1980s, however, the Internet began to connect computers all over the world. In 1994, the World Wide Web had grown from 500 to 10,000 servers; in 1995–1996, internet traffic hit a peak growth, doubling in 3–4 months. This heavy demand stimulated the commercial application of high capacity optic fiber. From 1992 to 2001 system capacity doubled every 6 months, leading around 2000–2001 to an oversupply in capacity, referred to as the so-called "fiber glut". However, after this period of overcapacity, it turned out that people need more and more capacity and fiber optic communication systems are nowadays developing steadily.

Returning to the cumulative citation curve we see that in the spline-fitted curve shown in Figs. 2 and 3 the minimum occurs just before the biggest breakthrough was achieved in the academic phase. The maximum occurs just before the biggest breakthrough was achieved in the application phase. The inflection point from convex to concave corresponds to the time that the field went from the academic phase to the application phase, while the inflection point from concave to convex correspond to the time when the technologies in the fiber optic communication improved and became mature. Moreover economic rules of supply and demand also play a role shortly after technology matures. Is all this just a coincidence or is there something deeper going on?

5. "Movement" of a scientific idea in academic space

Normally, a scientific investigation is based on current knowledge, and publication logically infers the idea of new insight into some phenomenon of interest. When the investigation is carried out and made public by publication, the scientific ideas contained in these publication(s) begin to be diffused in academic space. We assume here that the publication contains some innovative ideas and not just facts or trivial ideas. By their nature such ideas and their consequences are not fully understood when they are first made public. The proposed ideas invite reactions and begin their journey to realize their scientific value. However, they may meet with some obstacles. They engage in interactions with these obstacles and try to break through them. If successful, scientific value is realized; if unsuccessful, scientists realize that their ideas cannot be fully realized (or maybe that they were totally wrong).

When reactions between the original ideas and new insights into the phenomenon happen, the resulting new insights may lead to a new publication, resulting in citations to the original article. What obstacles the original idea encounters and how it interacts with these obstacles, all this is reflected in the citation process. So we may use the citation process and its changes as a representation of evolution of scientific ideas in academic space. We assume that the number of citations received in a specific time can represent the position that the original ideas has reached in academic space. This position will change as the number of citations increases. The difference between positions at different points in time represents the distance that the scientific ideas have traveled in this abstract landscape during a certain period. In such a way, the citation diffusion process can be regarded as a movement process. We say that the citation diffusion process represents the academic movement process and consider the question: "Which indicators can be used to describe this motion?" When splines are used to describe discrete cumulative data the resulting curve describes how the number of citations grows over time. Yet, we like to stress that the use of splines is not at all essential here. Another twice differentiable function that contains all necessary information would do equally well. Yet, we just use splines because we think they are an excellent tool to describe discrete data in a differentiable way.

6. Factors influencing the academic movement process as reflected by citations – turning points

6.1. Development of a field

In the linear model, science and technology are developed from basic research and then their products find applications in society. Of course this model is too simplistic but in a first step we will use it to guide our ideas. During this process, products are introduced in the market, and diffused into social life. From the review of the process of fiber optics communications, we know that from 1966 to 1975, the field is in the academic phase, and from 1975 to 1996, the field is in the technology dominant phase, from 1996 on the field penetrated our daily life. We saw that the spline-fitted curve representing the academic movement is convex in the academic phase, and concave in the technology dominated phase. After the technology dominated period the curve oscillates for a while and then goes on parallel to trend line. Inflection points in the spline-fitted curve mark turning points between different phases.

6.2. Breakthroughs as shown in the diffusion process

Breakthroughs are characterized by major progress in overcoming obstacles. Although these breakthroughs happen in academic space we claim that, because of the representation that we have in mind, they must leave a trace in citation history. We clearly see the slow citation start of Kao's article reflecting the first difficult period, when fiber optics communication was mainly an idea, a dream as Kao recalls it in an interview mentioned in (Hecht, 1999, reference 1, Chap. 10). Received citations in this period are mainly from the fields of glass and of laser technology. In 1969, the spline-fitted curve reaches its minimum. At that time very few people believed that the dream could be realized. We may say with great probability that in this year the diffusion would have stopped if there were no breakthrough. Yet, the first commercial optical fiber was realized in 1970, overcoming the main obstacle to a practical realization of Kao's conclusion. This breakthrough led to a big enthusiasm resulting in an intense academic publication activity. These circumstances promoted the diffusion of Kao's idea in the abstract intellectual landscape.

It is widely believed that publishing activity ceases when a field moves into an application phase and instead patenting and industrial secrecy would hold sway, as Small and Upham (2009) so elegantly write. But in the field of fiber optic communication this was not yet the case since there were still two major obstacles restricting the capacity of fiber optic communication systems: one is the limited low-loss bandwidth that can be used to transmit the data, and the other is the limited transmission distance of a fiber-optic communication system because of fiber attenuation and fiber distortion. Though the Dorset police installed the first non-experimental fiber-optic link after lightning knocked out their communication system (bringing fiber optics communication into the application phase), Bell and the British Post office both abandoned the technology because of technological obstacles and costs. The former obstacle was overcome by enlarging bandwidth step by step from 1975 to 1996. The latter one was overcome by the invention of the erbium fiber amplifier (EFA) in 1987, 1 year after the maximum in the spline curve.

Consequently, publications continued in this phase and scientists still explored all possible ways to overcome these obstacles. Indeed, citations in this period mainly came from articles on technologies used in communication and related to repeaters, receivers, amplifiers, and so on.

Since breakthroughs constitute major progress achieved by overcoming obstacles, the bigger the obstacle, the more questions need to be solved, the more articles will be published, and the more citations the article that introduced the symbol idea will receive. In other words, interestingness of the phenomenon related to the symbol idea becomes more intense. This is consistent with the essence of citation (namely the notion of interestingness, (Liu, 2011)). After breakthroughs are achieved in the academic phase, the interestingness of the phenomenon will increase as more questions need to be answered. A breakthrough in the application phase, on the contrary, means that obstacles are solved, and debates are settled.

In Kao's curves, the breakthrough in academic phases appears near the minimum. In the application phase the breakthrough appears near the maximum. We may say that breakthroughs influence the diffusion rate directly. In the next section we provide a more detailed analysis using the fact that we have described the cumulative citation process by a differentiable curve.

7. Instantaneous velocity and instantaneous acceleration of the diffusion of Kao's idea

As in physics, we refer to the first derivative as instantaneous velocity and to the second derivative as instantaneous acceleration. The first derivative of a convex function is negative before and positive after a minimum. The second derivative of a convex function is always positive. The first derivative of a concave function is positive before and negative after a maximum. The second derivative of a concave function is always negative. Hence we may say that in the period corresponding to a convex curve diffusion is accelerated. In the period corresponding to a concave curve, the diffusion is decelerated, though at the beginning of this period, velocity still shows an increasing trend.

Since splines have the capability of storing how information is changed, we calculate the first order and the second order derivative to see what we can find in this view.

Clearly the first order and second derivative curves oscillate, see Fig. 4. The reason for this behavior is that spline functions are piecewise defined functions, and the first order derivative and second derivative of the function are also piecewise

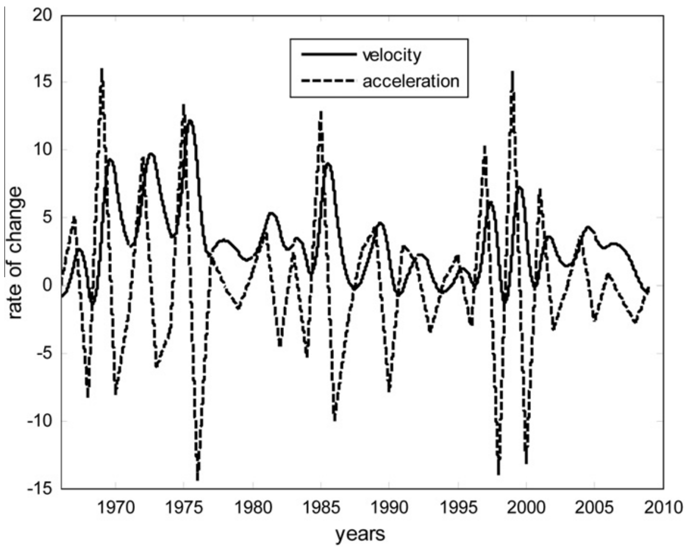

Fig. 4. Velocity and acceleration calculated from a continuous cumulative citation function.

functions, so velocity and acceleration reflect citation changes in the local sense. The frequency of velocity and acceleration are the same. Periods of acceleration start earlier than that of velocity. We notice that amplitudes around the breakthroughs (in 1970 and 1986), the phase change (in 1975) and the period of overcapacity (in 2000) are obviously bigger than during the normal periods. Also the periods near these years are smaller than during other times. The velocity curve is always situated above the *x*-axis except at the very beginning.

8. Analyzing different phases in the history of fiber optic communication

8.1. Diffusion of innovative ideas

S-shapes are traditionally used to describe innovation diffusion processes. Rogers (1962) defines innovation diffusion as the process by which an innovation is communicated over time, and this through certain channels among the members of a social system. The typical diffusion of innovations model is an "S-shaped curve". An innovation is adopted in a social system when the innovation is used by people. Similarly a scientific idea is diffused in the scientific communication system, when the article which proposes the scientific idea is cited. These two diffusion processes share the same mechanism. The scientific idea in Nobelist Kao's most important article is innovative, and hence it is no surprise that the diffusion of the article in the academic communication system through citations also follows an S-shaped curve.

Now we look in detail at the different phases of the evolution in the field of fiber optic communication. Instead of the simple linear model we distinguish three main phases in the history of a well-developed technology: the basic research phase, the technological phase and post-technology period. These phases, however, are often intertwined, as products tend to go through different steps of improvements.

We like to point out here why a curve with non-trivial second order derivatives is of the utmost importance in our work. Indeed: although a curve may still be increasing, the field may already be past a point of inflection (second derivative equal to zero) and in a concave phase. In this way future decrease is already announced (unless we first encounter another point of inflection). Of course, a similar phenomenon occurs when the field is past an inflection point and in a convex phase: then future increase is announced.

8.2. Basic research phase: from 1966 to 1975

The period 1966–1975 is the academic phase in the field of optic fiber communication. We know that the cumulative citation curve in this phase is largely convex. Hence citation diffusion accelerates in this phase. In the detrended curve we notice a minimum around 1969. In the history of optic fiber communication, a crucial material problem was effectively solved in 1970, the year next to the minimum point in the detrended spline curve. This invention promoted further research in the field of optic fiber communication systems, which – through citations – also promoted the diffusion of the scientific ideas in Kao's original article. In this phase, the largest amplitude (for the acceleration curve) occurs in the years 1968–1970 which was the period that the important breakthrough happened. The fact that the cumulative citation curve is in this phase mainly convex combined with the fact that the velocity and acceleration curves oscillate indicate changes in yearly diffusion although these changes do not influence the intrinsic diffusion trend.

8.3. The technology dominated phase: from 1975 to 1996

The period from 1975 to 1996 is the application phase in the history of optic fiber communication systems. Citations in this period are all from technology-oriented fields: we may say it is a diffusion phase dominated by technology. In this period, the spline-fitted cumulative citation curve first has (roughly) a linear growth (faster than that of the regression line), followed by a (roughly) linear decrease. The detrended curve in this period is largely concave with a maximum in 1986. A year later, the Erbium doped fiber amplifier was invented. This device is a crucial invention which made it possible to establish long distance optic fiber communication systems, such as the above mentioned underwater cables.

It is common sense that in the application phase of a field the academic publication activity declines, and consequently one expects the yearly citation curve to decrease even faster. But in our case, in this phase, the citation activity did not stop, and neither did publication activity. Kao's article was still cited but diffusion slowed down. Scientists still intensely studied the new technology to improve the system. This induced new citations to Kao's article, which means that its ideas still inspired new insights on fiber optic communication. After the technological breakthrough in the year 1986, the field of optic fiber communication matured and the diffusion process has a converse S-shape.

8.4. Post-technology phase: from 1996 on

From 1996 on, fiber optic communication became a part of public life. Optic fibers not only connected telephones and televisions, but also computers. Optic fibers led to the construction of the Internet infrastructure. More and more people are involved in this new form of communication and the speed of communication becomes faster and faster. The extent to which communication influences society increases in all senses of the word. The style of life and work, even the style of love-making (dating sites) is changing. So the field entered a period which we call the post-technology phase.

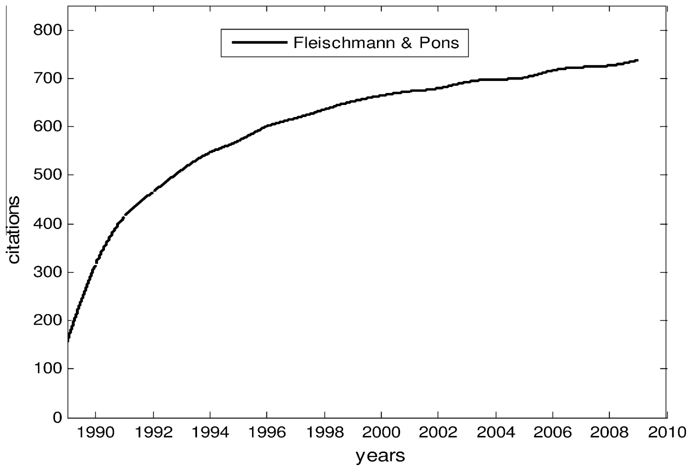

Fig. 5. Spline curve describing the academic movement process of Fleischmann and Pons.

9. Another case: the Fleischmann–Pons article on cold fusion

Without any doubt Kao's article is a concept symbol of a field that has gone through several stages. Similar cases are hard to find, yet we offer one other example of a concept symbol, namely the Fleischmann–Pons article (Fleischmann & Pons, 1989) on cold fusion. In this case the field has not (yet?) gone through several stages. Indeed, it has gone through an explosion of attention and soon thereafter came (almost) to a standstill. We write "almost" as there are still scientists, the "believers", who work on this topic (nowadays called LENR, for Low Energy Nuclear Reaction) and try to understand the phenomena they measure. Storms (2010) contains a recent review written from the point of view of the "believers".

Bibliographic data of the cold fusion article are:

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

This article has been cited more than 750 times in 21 years. Yet, the first 350 citations were received during the first 2 years after publication. Its citation history led to a fast increasing concave citation curve, as illustrated in Fig. 5 (based on WoS data). Indeed cold fusion was a new concept challenging the status quo of that time. The resulting discussions led to a large number of citations in a short time. In (Liu, 2011) it was observed that the cumulative citation curve of the most important article written by Nobel Prize winner S.C.C. Ting (Aubert et al., 1974) starts similarly. Yet, increases in citations to this article endured much longer and, after a period of slow grow, continued in a convex fashion. On Fig. 5 we see that the spline curve of the citations to Fleischmann and Pons' article does not show such a convex trend in the later phase.

This case shows that when considering the cumulative citation curve it is not immediately clear which represents a big breakthrough and which will turn out to be wrong (or at least not accepted by the scientific establishment). This leads to the interesting question of which factors determine the shape of an academic movement process? The number of citations is one factor, but how the academic position changes over time is another determining factor.

Fig. 6 shows the instantaneous velocity and instantaneous acceleration of the diffusion process of the ideas of Fleischmann and Pons. It shows the fast decline in velocity and the large negative acceleration almost immediately after the publication of the article. These are clear indications that the theory of cold fusion is not accepted.

10. Potential use of our theory of diffusion processes

In this contribution we proposed a method to study scientific ideas and their interactions by tracing corresponding citations. What are the potential uses of our suggestion?

10.1. Discerning breakthrough and turning points in an S&T area

With strict theoretical and experimental studies and intense research, difficult problems can be solved and a research field can totally change. Our study argues that the diffusion of really original scientific ideas illustrates the turning point from academic research to the application phase through points of inflection, while technical breakthroughs in the field show themselves in the extremes. An inverse question presents itself immediately. Can the shape of the cumulative citation curve be used to test if a scientific idea is innovative or not?

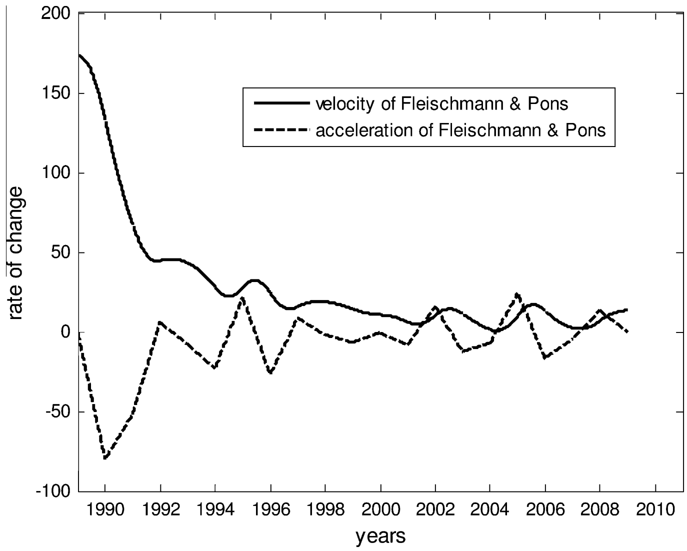

Fig. 6. The instantaneous velocity and instantaneous acceleration of the diffusion process of the Fleischmann–Pons article.

10.2. Detecting more detailed changes using the concepts of instantaneous velocity and instantaneous acceleration

In our empirical study the instantaneous velocity and instantaneous acceleration show special behavior near turning points and breakthrough points. This helped to discern special moments (or periods), even amid oscillations. Amplitude and frequency contain information about the characteristics of different phases. In our opinion, the instantaneous velocity and instantaneous acceleration reveal mostly information in the local sense. Even if the scientific ideas and technical breakthroughs are in their cradle, there are signs indicating future progress. These signs show themselves by the rate of change in the citation diffusion process. Because of the occurring oscillations we cannot directly obtain this information. The rate of change of cumulative citations, i.e. the yearly number of citations that the article receives is determined by two factors: one is a possible breakthrough or turning point while the other is pure randomness. We all know that publication and communication among scientists in the academic publication system are social activities. This social process is a stochastic process, influenced by random factors, i.e. there are certain factors as well as uncertain factors influencing the actual academic process. Certain factors come from scientific developments themselves such as breakthroughs in technical skill and other scientific ideas. Uncertain factors come, for instance, from the actual publication behavior of scientists and the time to review and publish. This applies not only to the original article but to those citing it as well. If we can remove, or at least reduce the influence of random factors in the academic movement, then the concepts of instantaneous velocity and acceleration may give us fine-grained information about the development of a scientific field. This leads to the problem of eliminating these random factors. We claim that discerning emerging trends is done by observing a change in the frequency of citations.

Actually obtaining this information is a serious, though important problem deserving further investigation. Eliminating the influence of random factors can contribute to the elimination of the oscillation in the velocity and acceleration curve. As an – unproven – suggestion we propose the use of Kalman filtering (Kalman, 1960). This technique is used, e.g. in astronomy where the signal is often buried under random noise. Taking it one level further we claim that, if we succeed in eliminating oscillation due to random factors, we may obtain the full story of how citations change in response to interactions between scientific ideas and how technical breakthroughs influence citation rates.

10.3. Describing the interaction of different subject areas and even of different ideas

We know that scientific ideas can be diffused into different subject areas, where subject areas are defined through the topics and ideas that are studied. If we can eliminate oscillations induced by random factors, then the rate of change is purely the result of the interactions of different scientific ideas, possibly corresponding to different subject areas. In this way we could establish a one–one correspondence between interactions among different ideas and the rate of change in the cumulative citation curve. If these ideas can be developed further we can really obtain a representation of the interaction of scientific ideas.

10.4. Finding social, technological, political and economic factors influencing the development of science

Social processes are to some extent stochastic processes, influenced by random factors. In scientific communication randomness is mostly visible during the submission, reviewing, revision and publication periods. However, also social factors including technical, political (science policy related) and economic factors directly influence the investigation process. In the short term these factors can be considered as deterministic. We claim that the study of diffusion as proposed by us can uncover the interaction between these factors and the development of science. As a point in case we saw that for Kao's article the citation rate was different during the academic, technological and post-technology phases. Moreover, we noticed that nowadays fiber-based communication, as initiated by Kao, influences society in all senses of the word: it changed the style of life and work of people all over the world.

11. Conclusion

In this contribution we studied articles acting as concept symbols and defined functions characterizing the actual temporal behavior of the citation diffusion process leading to "citation life histories". In doing this we assumed that a scientific idea that is put forward in an article is diffused into academic space and that this diffusion can be seen and measured through citations. During diffusion, the scientific idea follows a trajectory in academic space reflected by citations and changes in citation density. Clearly diffusion of scientific ideas can only be studied in the case of articles (or other publications) containing innovative ideas which are, indeed, diffused over many fields. "Normal" articles that are cited just a few times play no role in the diffusion of ideas and hence cannot be used for the type of studies we propose. Using spline interpolation we established a function that reflects the relation of the academic position (as measured by cumulative citations) with respect to time.

We presented two actual case studies in the field of science and technology. In this respect we add the caveat that actual scientific developments are probably different in the social sciences and humanities. Yet, we claim (be it without proof) that also in these fields citations can be viewed as traces of the development of ideas. Using citations of Nobelist K.C. Kao's highly original idea incorporated in his main article we established a concrete function reflecting the relation of cumulative citations with respect to time. With the help of the regression line and a detrended curve, we can see that the citation diffusion curve of an article containing a really original idea has an S-shape similar to the standard innovation diffusion curve. The convex part corresponds to the academic phase of the field that Kao's idea initiated, while the concave part corresponds to the technology dominated phase. The curve in the post-technology phase paralleled the regression line. The points of inflection correspond to the phase transition from academic to application research, while minima indicate a breakthrough in academic phase, and maxima indicate a breakthrough in the technology dominated phase. We further observe that breakthroughs may directly influence the rate of change of the diffusion process while phase transfers may influence the rate of change implicitly.

Publication and citation patterns in the academic and the technology dominated phase are different. It is believed that academic publication will cease when a field enters the application phase. Therefore citations to the original idea will also cease. Our case study, however, does not confirm this postulate. The citation diffusion process of the original article indeed slowed down when optic fiber communication enters the application phase, but citations did not decrease.

When Kao gave the world a potential new form of communication medium based on precise theoretical and experimental studies, this was just a potential device not yet a practical realization. Crucial material problems still had to be solved. Only when these investigations finally led to a commercial product people realized that Kao's ideas had revolutionized the whole communication technology and its industry. Kao's ideas are diffused in relation to many specialties. Indeed, checking the citations received by his article we find:

(1) investigations of the properties of materials related to the fiber optic communication system and improvement of technological details;
(2) investigations in paralleling fields such as sensors, measurement apparatus and astronomy;
(3) explorations of new applications in social life and education;
(4) studies in business management of corporations dealing in optic fiber communication systems;
(5) studies of hot topics in science such as network studies and purely journalistic articles, especially after Kao received the Nobel Prize (but these were not used in this article).

The Fleischmann–Pons article provides a case of an immediate explosion of attention, soon followed by the rejection of the ideas proposed in it.

The full realization of our ideas depends on the extent to which one is able to eliminate the randomness effect. We also need to investigate deeper how scientific ideas are produced, how they are diffused in the abstract landscape of the science system, and what happens to citations when interactions occur. Although, in our opinion, all this must still begin from scratch, we are convinced that it deserves our effort. Applying the principles explained in this contribution informetrics may reveal the regularities along which science is developing.

Acknowledgments

Work of the authors is supported by NSFC Grant No. 71173154. Moreover, R. Rousseau enjoyed support of an international grant from the National Natural Science Foundation of China (NSFC Grant No. 7101017006). The authors thank two anonymous reviewers for helpful suggestions to improve their article.

Appendix A

See Table A1.

Table A1
Citations to Kao–Hockham: distributions over time.

Year	Age	C1	C2	Year	Age	C1	C2
1966	0	1	0	1988	22	125	89
1967	1	1	0	1989	23	127	91
1968	2	3	2	1990	24	132	95
1969	3	4	2	1991	25	133	95
1970	4	13	10	1992	26	135	96
1971	5	18	15	1993	27	140	98
1972	6	22	19	1994	28	140	98
1973	7	31	28	1995	29	143	98
1974	8	37	34	1996	30	146	99
1975	9	44	39	1997	31	147	100
1976	10	56	50	1998	32	153	105
1977	11	61	54	1999	33	154	105
1978	12	65	57	2000	34	163	111
1979	13	72	60	2001	35	165	112
1980	14	77	62	2002	36	170	115
1981	15	83	65	2003	37	175	117
1982	16	90	70	2004	38	179	119
1983	17	94	73	2005	39	184	123
1984	18	100	76	2006	40	190	126
1985	19	109	78	2007	41	194	129
1986	20	120	86	2008	42	199	131
1987	21	123	89	2009	43	201	131

C1 refers to all received citations; C2 to citations received in articles, proceedings and letters.

References

Aubert, J. J., Becker, U., Biggs, P. J., Burger, J., Chen, M., Everhart, G., et al (1974). Experimental observation of a heavy particle. *Journal of Physical Review Letters, 33*(23), 1404–1406.
Avramescu, A. (1975). Modelling scientific information transfer. *International Forum on Information Documentation, 1*(1), 13–19.
Avramescu, A. (1980). Coherent informational energy and entropy. *Journal of Documentation, 36*(4), 293–312.
Barnett, G. A., Fink, E. L., & Debus, M. B. (1989). A mathematical model of academic citation age. *Communication Research, 16*(4), 510–531.
Borgman, C. L. (1989). Bibliometrics and scholarly communication: Editor's introduction. *Communication Research, 16*(5), 583–599.
Boyack, K. W., Klavans, R., & Borner, K. (2005). Mapping the backbone of science. *Scientometrics, 64*, 351–374.
Class for Physics of the Royal Swedish Academy of Sciences (2009). Two revolutionary optical technologies. Royal Swedish Academy of Sciences. <nobelprize.org/nobel_prizes/physics/laureates/2009/phyadv09.pdf>.
Fleischmann, M., & Pons, S. (1989). Electrochemically induced nuclear-fusion of deuterium. *Journal of Electroanalytical Chemistry*, *261*(2A), 301–308.
Gambling, W. A. (2000). The rise and rise of optical fibers. *IEEE Journal of Selected Topics in Quantum Electronics, 6*, 1084–1093.
Hecht, J. (1999). City of Light. In: *The story of Fiber Optics*. New York, Oxford: Oxford University Press.
Hecht, J. (2004). City of Light. In: *The story of Fiber Optics*. New York, Oxford: Oxford University Press.
Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering*, *82* (D), 35–45.
Kao, K. C., & Hockham, G. A. (1966). Dielectric-fibre surface waveguides for optical frequencies. *Proceedings of the Institution of Electrical Engineers, 113*, 1151–1158.
Kiss, I. Z., Broom, M., Craze, P. G., & Rafols, I. (2010). Can epidemic models describe the diffusion of topics across disciplines? *Journal of Informetrics, 4*, 74–82.
Le Coadic, Y. F. (1987). Modelling the communication, distribution, transmission or transfer of scientific communication. *Journal of Information Science, 13*, 143–148.
Liu, YX., & Rousseau, R. (2011a). A continuous description of discrete data points in informetrics: Using spline functions. *Aslib Proceedings*.
Liu, YX., & Rousseau, R. (2011b). Splines can recover dynamic information contained in discrete data. In: Noyons, Ngulube, & Leta, (Eds.), *Proceedings of the ISSI 2011 Conference* (pp. 1022–1024), ISSI, Leiden University & University of Zululand.
Liu, YX. (2011). *The diffusion of scientific ideas in time and indicators for the description of this process.* Doctoral dissertation, Antwerp University.
Liu, YX., Rafols, I., & Rousseau, R. (in press). A framework for knowledge integration and diffusion. *Journal of Documentation*.
Liu, Y. X., & Rousseau, R. (2010). Knowledge diffusion through publications and citations: A case study using ESI-fields as unit of diffusion. *Journal of the American Society for Information Science and Technology, 61*, 340–351.
Mears, R. J., Reekie, L., Jauncey, I. M., & Payne, D. N. (1987). Low-noise erbium-doped fiber amplifier operating at 1.54 μm. *Electronics Letters, 23*, 1026–1028.
Rogers, E. (1962). *Diffusion of Innovations*. New York: Free Press.
Rousseau, R. (1997). *Numerical mathematics (in Dutch)*. Course notes, Oostende: KHBO.
Schubert, A., & Glänzel, W. (1986). Mean response time – a new indicator of journal citation speed with application to physics journals. *Czechoslovak Journal of Physics, 36*, 121–125.
Shiavi, R. (1991). *Introduction to applied statistical signal analysis*. Homewood(IL), Boston (MA): Irwin, Aksen.
Small, H. (1978). Cited documents as concept symbols. *Social Studies of Science, 8*, 327–340.
Small, H. (1987). The significance of bibliographic references. *Scientometrics, 12*, 339–341.
Small, H., & Upham, P. (2009). Citation structure of an emerging research area on the verge of application. *Scientometrics, 79*(2), 365–375.
Smith, L. (1981). Citation analysis. *Library Trends, 30*, 83–106.
Storms, E. (2010). Status of cold fusion. *Naturwissenschaften, 97*(10), 861–881.
Tague-Sutcliffe, J. (1992). An introduction to informetrics. *Information Processing and Management, 28*(1), 1–3.
van Raan, A. F. J. (2008). Scaling rules in the science system: Influence of field-specific citation characteristics on the impact of research groups. *Journal of the American Society for Information Science and Technology, 59*, 565–576.