A chemical concept from the *Science Citation Index* database

Dominique Bissey and Vincente Viossat

Laboratoire des Systèmes Interfaciaux à l'Echelle Nanométrique, Université Pierre et Marie Curie, 4 Place Jussieu, 75252, Paris, France. E-mail: bissey@ccr.jussieu.fr, viv@ccr.jussieu.fr; Tel: +33 (0)1 44 27 55 46

Opinion

Received (in Montpellier, France) 12th December 2002, Accepted 18th February 2003
First published as an Advance Article on the web 16th June 2003

We study the concept "modeling in combustion" from 1980 to 2000 with data from the "Science Citation Index". A statistical study of this concept shows (*i*) the growth of the number of publications and their authors and (*ii*) the distribution laws concerning the cited references.

The chemical concept of "modeling in combustion" began in 1967 with dihydrogen combustion in a well-stirred reactor[1] and in a shock-tube.[2] The development of this concept can be attributed to the progress in "stiff" equation-solving techniques, the software package CHEMKIN being available since 1980.[3] Mathematical and chemical reduction techniques have been developed during the last decade in order to obtain tractable formulations for numerical combustion applications (method of Peters;[4] ILDM method;[5] the "*Perturbations Singulières*" method[6,7]).

We have tried to find a behaviour law for the "modeling in combustion" concept from 1980 to 2000. Although this idea can be dated back to 1967, its development has been really significant since 1980. Our "bibliometric" research is based on bibliographical data, which is taken from the multidisciplinary *Science Citation Index* database, each record containing full bibliographic details: title in English, abstract, journal name, author and his affiliation, and in particular a complete list of the references cited by each indexed publication. The searches were carried out with keywords in the title, using "combustion and model*ing" corresponding to "combustion and modeling" or "combustion and modelling"; we thus obtained 338 publications including 312 articles.

Temporal evolution of the concept

Growth or decline in the number of publications in a particular area reflects the popularity of the subject matter.[8] In our study, we supposed that all of the publications collected by the database reflect *exactly* the scientific activity relating to the subject. Consequently, the growth rate of the number of publications in a subfield can be directly correlated to the development of scientific progress in that field.

In a first step, we found that the number of publications follows an affine law according to the year with a rate equal to 1.3 publications per year; the number of authors interested in the subject grows in the same way with a rate equal to 4.2 authors per year. Then we tried to quantify this evolution by using the cumulative number of publications (or authors) since 1979, chosen as the origin. Two approaches were used to find, in an empirical way, a law that describes this evolution as accurately as possible.

The first approach, that was also developed for another concept,[9] uses the classic laws of chemical kinetics. Let P be the cumulative number of papers (or authors) that can be linked to a product concentration; the rate of growth ($\mathrm{d}P/\mathrm{d}t$) is assumed to be proportional to P as in eqn. (1):

$$\mathrm{d}P/\mathrm{d}t = kP, \qquad (1)$$

with k being a positive constant. By integration, we obtain:

$$P(t) = P_0 \mathrm{e}^{kt} = P_0 \mathrm{e}^{(\ln 2/T_\mathrm{d})t} \qquad (2)$$

with P_0 being the value of P at $t = 0$ and T_d the doubling time.

The curves represented in the Fig. 1 show two phases: the first between 1980 and 1985, and the other since 1986 to the present, during which the cumulative number of publications (or authors) *versus* time follows two different exponential laws. By comparing the curves representing literature and manpower, we observe the same growth behaviour: a break at $t = 6$ years, and, for each phase, the same doubling time T_d. That is due to the proportionality between papers and authors because the average yearly number of authors per article equals 3.5 ± 0.2.

The fast start of this evolution can be explained at first by the availability of the software CHEMKIN in 1980 but also in a more general context by the fashion effect that a new idea generates, like the "self-organised criticality"[10] concept: a break in the curve is noticed after approximatively five years.

The another approach we used is a statistical study of this evolution, by considering our data as finite and whole quantities. We sought a curve applicable to all the points representing the cumulative number P of publications (or authors) *versus* time; only the power law seems to be appropriate over the whole period studied.

We suggest a law of the type:

$$P = P_1 t^a \qquad (3)$$

with P_1 being the value of P at $t = 1$. This may be represented as in Fig. 2. The value of the exponent is appreciably constant ($1.56 \le a \le 1.59$).

We are thus confronted with two different laws that correctly describe the temporal evolution of the concept

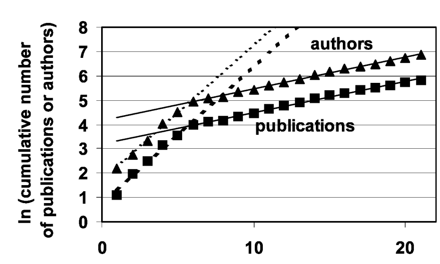

Fig. 1 Exponential laws for publications and authors.

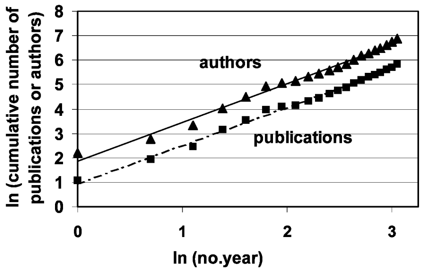

Fig. 2 Power law for the publications and authors.

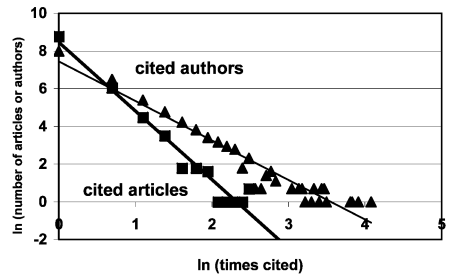

Fig. 3 Distribution laws for publications and authors.

"combustion modeling". We obtained the same conclusion in other similar studies with parameters such as the type of publication or the affiliation of authors. As we cannot give weight to one over the other approach, it is significant to explore the "why and how" of this behaviour.

Transmission of the concept

A new idea can be propagated if it is correctly transmitted. The transmission of ideas within a population is similar to the transmission of infectious diseases.[11] Some diseases may be transmitted if infective and susceptible subjects come into close proximity. Other diseases may be transmitted by means of biological carriers. The role of these carriers in the scientific community is played by publications. The development of an idea is thus closely associated with the cited references in publications. A thorough study of the references cited in publications has enabled us to gather much information such as the age of the references and the distribution function of the cited references or authors.

The cumulative number of cited references decreases exponentially with their age. If $R(t)$ is the number of references older than time t, this is expressed thus:

$$R(t) = N\mathrm{e}^{-(\ln 2/T_{1/2})t} \qquad (4)$$

with N being the total number of references and $T_{1/2}$ the half-life of references. An application to our data shows that about 50% of the references are less then 5 years old. This period being too short, we have decomposed all the references into two groups. The first, composed of references less than 20 years old, accounts for 91% of all the references. References older than 20 years form a slowly aging component of about 9% of the references.

All these results show the need for continuously renewing publications to preserve the topical character of an idea. Moreover, some publications do not find any "echo" in the scientific literature and they become "dead ends" in the development of a concept. The existence of these "dead ends" can explain the growth of the cumulative number of publications against time; following eqn. (3) we get a constant exponent value contained between 1 and 2.

That is why we propose another interpretation: the development of an idea through the scientific literature can be a "fractal" phenomenon.[12] The impact of one reference can be quantified by the number of times it has appeared in all the articles published. We classified 7862 references indexed in 312 articles published from 1980 to 2000, according to the number of times (n) that they appeared. The cited references $N(n)$ follow a distribution law of the type:

$$N(n) = K/n^b \qquad (5)$$

There is a similarity to the fractal distribution exposed by Turcotte[13] in the form:

$$N(r) = C/r^D \qquad (6)$$

where $N(r)$ is the number of objects with size r and D the fractal dimension introduced by Mandelbrot[14] in 1967.

The references cited in scientific articles are not only the reflection of the development of an idea but they also testify to the success of the authors. So we performed a similar study concerning the cited authors: the distribution functions, represented by similar graphs (see Fig. 3), behave according to the same laws for the cited references with $b = 3.46$ and for the cited authors where $b = 2.10$. If this distribution may be supposed "fractal",[14] then b would represent the "fractal dimension".

The irregularities in time concerning the evolution of our concept could explain the potentially "fractal" character of this phenomenon. These variations, compared to a continuous evolution managed by a traditional law, are indeed produced randomly. Nevertheless, they preserve a logic, an order in their apparent disorder.

The references cited in the articles are the links between authors and the idea that can be transmitted by them. The evolution of the concept can be schematised by a family tree. The tree trunk would represent a source event and the branches would be the references connecting two events. The number of authors or reference citations depends on the geometry of the branch-line system constituted by the references.

The increasing requirements for "modeling in combustion" by various applications (automobile and aeronautical engines, rocket motor engines, fuels, explosions and fires, pollution induced by combustion, environmental protection, *etc.*) are important enough to explain the development of this concept. However, we must note the contribution of physics, mathematics and computer science to this development, even if these disciplines play only an auxiliary role.

References

1 D. R. Jenkins, V. S. Yumlu and D. B. Spalding, in *11th Symposium (International) of Combustion*, The Combustion Institute, Pittsburgh, PA, USA, 1967, pp. 779–790.
2 C. W. Hamilton and G. L. Schott, in *11th Symposium (International) of Combustion*, The Combustion Institute, Pittsburgh, PA, USA, 1967, pp. 579–590.
3 (*a*) R. J. Kee, J. A. Miller and T. H. Jefferson, *CHEMKIN: A General Purpose, Problem-Independent Transportable Fortran Chemical Kinetics Code Package*, Sandia Report SAND80-8003, Sandia National Laboratories, Albuquerque, NM and Livermore, CA, USA, 1980; (*b*) R. J. Kee, F. M. Rupley and J. A. Miller, *CHEMKIN II: A Fortran Chemical Kinetics Package for the Analysis of Gas Phase Chemical Kinetics*, Sandia Report SAND89-8009B, UC-706, Sandia National Laboratories, Albuquerque, NM and Livermore, CA, USA, 1989.
4 N. Peters, in *Numerical Simulation of Combustion Phenomena*, (Lecture Notes in Physics 241), eds. R. Glowinski, B. Larrouturou and R. Temam, Springer-Verlag, Heidelberg, Germany, 1985, pp. 90–109.
5 U. Maas and S. B. Pope, *Combust. Flame*, 1992, **88**, 239–264.
6 S. H. Lam and D. A. Goussis, in *Twenty-Second Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, PA, USA, 1988, pp. 931–941.
7 P. Duchêne and P. Rouchon, *Chem. Eng. Sci.*, 1996, **1**, 4661–4672.

8 H. Boxenbaum, *J. Pharmacokin. Biopharm.*, 1982, **10**, 335.
9 M. Burger and E. Bujdoso, in *Oscillations and Traveling Waves in Chemical Systems*, ed. R. J. Field, Wiley, New York, 1985, pp. 563–604.
10 R. A. Bentley and H. D. G. Maschner, *Fractals*, 2000, **8**, 227–237.
11 W. Goffman and V. A. Newill, *Nature*, 1964, **204**, 225.
12 B. Sapoval, *Universalités et Fractales*, Flammarion, Paris, France, 1997.
13 D. L. Turcotte, *Fractals and Chaos in Geology and Geophysics*, Cambridge University Press, Cambridge, UK, 1997.
14 B. Mandelbrot, *Les Objets Fractals*, Flammarion, Paris, France, 1995.