

Articles

Trends in Condensed Matter Physics: Is Research Going Faster and Faster?

C. Attacalite^{a,1}, S. Barland^b

^a*Institut Néel, CNRS/UJF, 25 rue des Martyrs BP 166, Bâtiment D 38042 Grenoble cedex 9 France*

^b*Université de Nice - CNRS UMR 7335, Institut Non Linéaire de Nice, 1361 route des lucioles, 06560 Valbonne, France*

Received 30.08.2012, accepted 04.11.2012, published 06.11.2012

In this paper we study research trends in condensed matter physics. Trends are analyzed by means of the number of publications in the different sub-fields as function of the years. We found that many research topics have a similar behavior with an initial fast growth and a next slower exponential decay. We derived a simple model to describe this behavior and built up some predictions for future trends.

1 Introduction

The investigation of social phenomena has acquired large importance in the last years.[1] In particular statistical physics has been applied to study the interaction of individuals as elementary units of a social structure. In these models each individual interacts with a limited number of peers usually negligible compared with the total number of agents. In this ensemble of interacting agents, statistical physics has been used to investigate opinion and cultural dynamics, crowds behavior, social spreading phenomena and so on.[1]

On the other hand the structure of the human connections has been subject of an intense study by means of networks theory.[2] It has been shown that humans form particular kinds of networks, the so-called scale-free networks, that exhibit a very short path from one site to another thanks to the presence of large hubs connected to many nodes.[1] Also scientists and the scientific community have been subject of similar studies. In particular network theory has been employed to show that scientific collaborations form a scale-free network.[3] One important characteristic of these networks is that new ideas and trends can spread without any initial threshold due to the unbounded fluctuations in the connectivity.[4]

In this paper we will analyze the birth and death of trends in the scientific community, and in particular in condensed matter physics. In the last fifty years different topics have been subject of intense research in condensed matter physics, ranging from superconductivity to charge density waves or nanostructures. Researchers, during their career,

have usually worked on some of these subjects moving from one to another when they found the new one more stimulating or challenging. This collective behavior gave rise to the birth, growth and death of new trends. Here we want to investigate this dynamic without dealing directly with the diffusion of ideas in the scientists network, but devising a simple model to describe birth and death of new research topics in the community. In order to do so we will perform a bibliometric study based on lexical analysis of papers published in the last fifty years. Bibliometrics has been already employed to detect and investigate emerging research topics[5] by means of co-citation links or text-based techniques.[6] Here we will use a simple text-based approach to obtain the data that will then be analyzed by means of a model derived by population dynamics. The model resembles the dynamics of fashion in social science, that has been studied by means of Langevin and Fokker-Planck equations.[7,8] We will show that different sub-fields, irrespective of their importance and time location, exhibit a similar behavior that is well described by our simple model. The questions we would like to answer are: Are researches changing topics more often than in the past? Do research trends last less than in the past? And if this is true, is it due to the larger number of researchers? We do not have a definite answer to all these questions, but we will show that with a simple model we can get some insights for trends in condensed matter physics. The paper is organized as follows: in section 2 we discuss publication data; in section 3 we present the model used to describe trends in condensed matter physics; in section 4 we apply our model to the publication data and report the results; finally in section 5 we

¹E-mail: claudio.attacalite@gmail.com

draw conclusions and discuss open questions.

2 Publications and Data

The number of scientific publications in physics is exponentially growing (see Fig. 1). This trend has been found also in other research fields as chemistry, biology, and so on.[9] However the growth pattern of given sub-fields may be different from the total one, as already pointed out in previous studies.[9] This behavior is due to the fact that particular research areas are saturated, lack of interest or possible innovation.

In order to evaluate trends in condensed matter physics we consider the ratio between the number of publications in a given sub-field and the total one. The ratio is a good indicator for the importance of a given research sub-field because it is independent from the total number of publications and it allows to compare research topics at different times.

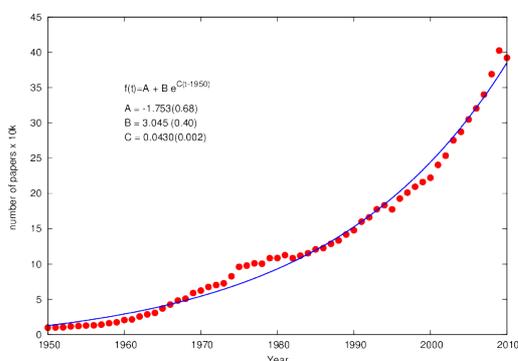


Figure 1: Normalization factor obtained from SAO/NASA Astrophysics Data System Abstract Service.[10] The total number of papers has been fitted with an exponential curve. Fit parameters are reported in the figure.

We started analyzing the number of abstracts published in physics from 1950 to 2011.[11] We found an exponential growth with a doubling time of about 16 years (data are shown in Figure 1) in agreement with previous studies.[9,12] Then we extracted the number of papers in the different sub-fields, and normalized this number with the total number of papers published each year. Some results are reported in Figure 3. The selection of papers belonging to a given sub-field has been done by means of a search in the abstract database using a combination of several keywords to best represent a research topic (see for instance the first column of Table 1).

3 The Model

In order to analyze trends in condensed matter physics we use a modified version of the logistic equation.[13] In his original formulation Pierre-François Verhulst applied the logistic equation to the population growth, assuming the reproduction rate being proportional to both the existing pop-

ulation and the amount of available resources:

$$\frac{dP}{dt} = rP \left(1 - \frac{P}{K} \right), \quad (1)$$

where r is the growth rate and K represents the environment carrying capacity.

In our model we consider P as the fraction of scientists that decides to work on a given research topic and we suppose P to be proportional to the paper ratio. As in the logistic model we suppose that the number of scientists working on a sub-field increases due to the available resources. In our case the ‘resources’ represent the possibility to make new discoveries in the sub-field, but also the possibility to get more grants and more citations in an emerging research area. This gain is equivalent to the so-called ‘bandwagon’ effect in fashion dynamics.[8]

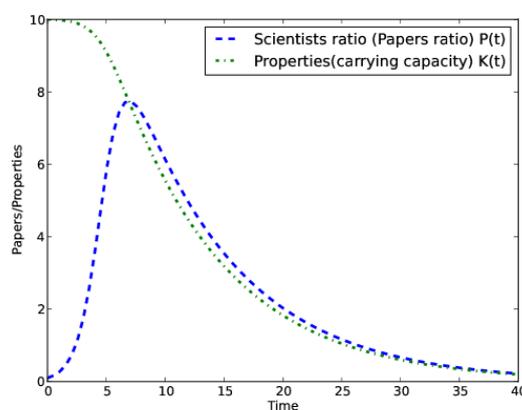


Figure 2: Simple model for fashion in physics, with parameters $K(t = 0) = 10$, $P(t = 0) = 0.1$ and parameters $r = 0.1$, $s = 0.01$.

Differently from the logistic model we suppose that the carrying capacity of a given research topic decreases with time. In fact one can expect that after some time a large part of the possible experiments and theory will be already over. Moreover there is also an economical component that contributes to the carrying capacity decrease: grants on a new and hot topic have clearly more successful probabilities than on an old and widely studied subject.[14] This fact automatically decreases the possibility to have many scientists working on old research topics due to the reduction of available positions. All these effects, that contribute to reduce the carrying capacity are equivalent to the so-called ‘snob effect’ in the fashion dynamics.

Therefore we suppose that K , the carrying capacity of our system, decreases proportionally to the number of scientists working on a given research topic multiplied by a *research rate coefficient* s

$$\frac{dK(t)}{dt} = -sP(t). \quad (2)$$

The *research rate coefficient* s represents the speed, at which the properties $K(t)$ are investigated by scientists $P(t)$. In principle the decay of $K(t)$ can have a more complex dependence from $P(t)$. In the literature, there are

several studies that include retardation effects or non-linear coupling.[15] however we did not find particular reasons for using more complicated models than a linear one. Combining equations 1 and 2 we get:

$$\frac{d^2 K(t)}{dt^2} = r \frac{dK(t)}{dt} \left(1 + \frac{1}{s} \frac{dK(t)}{dt} \frac{1}{K(t)} \right). \quad (3)$$

The solution of this equation has the form:

$$P(t) = \frac{\left(-\frac{rs}{(r-s)(C_2 r + C_1 e^{rt})} \right)^{\frac{s}{r-s}} r C_1 e^{rt}}{(r-s)(C_2 r + C_1 e^{rt})}, \quad (4)$$

$$K(t) = \left(\frac{rs}{(-r+s)(C_2 r + C_1 e^{rt})} \right)^{\frac{s}{r-s}}, \quad (5)$$

where C_1 and C_2 are the integration constants that depend from $P(0)$ and $K(0)$. In order to understand the behavior of this model we consider the limits at short and long times of the previous solution. For $t \rightarrow \infty$ we found that Eq. 4 decays exponentially as $P(t \rightarrow \infty) \simeq e^{-rs/(r-s)t}$. Since s is supposed to be much smaller than r , the exponent reduces to $P(t \rightarrow \infty) \simeq e^{-st}$. On the other hand for short time t , $P(t \rightarrow 0)$ goes as $P(t \rightarrow 0) \simeq c + e^{rt}$. A typical sketch of the solution is reported in Figure 2.

From the limits at short and long time we can see that the initial growth is dominated by the r coefficient while the exponential decay is dictated by the research rate coefficient s . In the language of population dynamics, the exponential decay is related to the decrease of carrying capacity, while the initial growth is due to the diffusion of a research topic in the scientific community. In order to analyze the publication data of different sub-fields, we simplify Eq. 4 in such a way to catch short and long time behaviors with a simple formula:

$$P_{model}(t) = P_0 \frac{1}{e^{-\alpha(t-t_0)} + e^{\beta(t-t_0)}} + c. \quad (6)$$

This model keeps only the main exponents for the growth and the decay part respectively, plus a constant P_0 that is the peak intensity while t_0 is related to the position of the maximum. We added a constant c not present in the initial equation that will be discussed in the next section.

Notice that this model does not apply to every research topic, in fact there are some active fields that continue to grow without any sign of decay, as for instance research on silicon, density functional theory, nanowires, and so on (see Figure 3). Moreover this model can be easily generalized to take into account more peaks as for instance in superconductivity, where the different discoveries gave rise to a sum of multiple curves, as the one in Figure 2, translated in time and with different intensities.

Table 1: Parameters obtained by fitting the normalized number of abstracts with the model in Eq. 2. The strength of each research topic has been normalized to the one of superconductivity $P_{sc} = 0.0262$.

Topic	t_0	(P_0/P_{sc})	β	α
crystal fields	1957.0(7)	0.095(5)	0.018(2)	0.27(4)
helium-3/-4	1959(2)	0.051(3)	0.06(1)	0.17(5)
alkali halides	1960(1)	0.186(6)	0.070(6)	0.12(1)
spin waves	1960.51(1)	0.125(12)	0.044(9)	0.30(4)
Haas-van-Alphen	1960.9(5)	0.023(3)	0.060(7)	0.5(2)
Brillouin scattering	1969(1)	0.048(3)	0.024(4)	0.24(3)
Compton scattering	1971.1(3)	0.0173(9)	0.060(4)	0.8(1)
spin glass	1977.9(3)	0.097(3)	0.018(2)	0.6(1)
charge density waves	1978.8(7)	0.054(3)	0.026(4)	0.33(4)
Josephson junctions	1983(1)	0.11(1)	0.038(8)	0.15(1)
fract. Quantum Hall	1983.8(3)	0.0162(1)	0.018(4)	1.6(5)
four wave mixing	1984.5(5)	0.091(5)	0.050(5)	0.30(2)
quasicrystals	1985.7(2)	0.031(2)	0.025(4)	1.6(5)
Penrose Tiles	1986.1(3)	0.022(1)	0.023(5)	1.0(2)
superconductivity	1986.7(2)	1.00(6)	0.045(5)	3(2)
slave bosons	1989.5(3)	0.0133(4)	0.062(6)	0.66(7)
t-J model	1990.1(1)	0.037(1)	0.090(5)	1.2(1)
fullerens	1991.2(1)	0.169(4)	0.030(2)	2.8(4)
Luttinger liquid	1993.1(3)	0.024(1)	0.032(5)	0.76(9)
giant-magnetores.	1993.0(1)	0.049(2)	0.071(7)	3(1)
Andreev reflection	1993.0(6)	0.013(1)	0.014(8)	0.54(9)
critical exponents	1993(2)	0.163(7)	0.078(1)	0.10(1)
Bethe ansatz	1994(1)	0.045(3)	0.08(1)	0.19(1)
composite fermions	1994.7(3)	0.022(2)	0.11(1)	1.4(3)
critical phenomena	1995(2)	0.096(4)	0.09(2)	0.10(1)
spin-charge separation	1999(1)	0.0125(1)	0.15(5)	1.0(4)
Coulomb blockade	1999(1)	0.069(6)	0.08(1)	0.26(2)
nanotubes	2003.1(4)	0.57(4)	0.08(1)	0.41(1)

4 Data Fit and Results

We suppose that the relevance of a research topic can be evaluated from the number of papers published on that sub-

ject. We extracted this number from the SAO/NASA Astrophysics Data System Abstract Service (SAS).[10] Since the total number of papers published every year is exponentially growing, we normalized the number of papers of a given research area with the total one in order to evaluate the importance of each research topic. Also the normalization factor is obtained by the SAO/NASA database (see Fig. 1). The data for the different topics were then fitted with the model described by Eq. 6, where the constant c has been always set to zero except for the case of ‘Superconductivity’. In fact, in this case, there was already an active research community working on ‘BCS’ superconductivity before the discovery of high- T_c superconductors boosted again the number of papers. In order not to miss papers and improve search results we included a list of synonymous for each research area as for instance ‘superconductors, superconductor, superconductivity etc .:’. The fit results are reported in Table 1, where we renormalized the peak intensity with the one of ‘superconductivity’, the largest peak in our sample, in such a way to underline the relevance of each sub-field. Selected data and relative fits are also plotted in

Fig. 3. Notice that recent research topics exhibit larger error bars because we could fit only the beginning of the decay tail. Moreover some active sub-fields as ‘‘graphene, topological insulators, quantum computing, photovoltaic, etc ...’’ are still in the growing part of the curve or have just started to decrease and this makes impossible to perform an accurate fit with our model.

In Table 1, we report the different parameters entering in the model that describes trends in condensed matter physics. Unfortunately the results for the α parameters are very noisy due to the fast diffusion of new research topics in the scientific community. This makes it more difficult to get an accurate estimate for this exponent. Nevertheless we found that α is slowly increasing with time (t_0), and this indicates a faster spread of ideas among scientists. This fast spread can be explained by the fact that the structure of the collaboration networks is scale-free[3] and in this kind of networks there is a vanishing threshold to spread ideas and trends.[4]

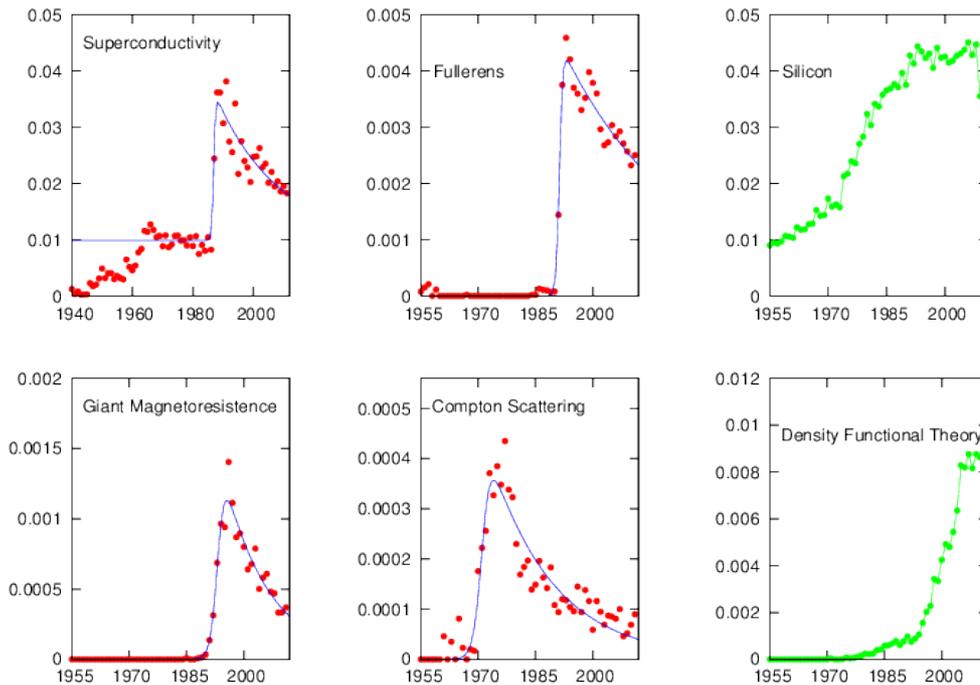


Figure 3: Number of papers for different research topics versus the years normalized to the total one in physics. The blue solid lines represent the fitting curves, fitting parameters are reported in Table 1. In green two topics that do not fit in our model.

Then we analyze the β parameter. In Figure 4, we report the parameter β as a function of the position of the maximum, which is given by $t_{max} = t_0 + \frac{1}{\alpha+\beta} \log(\frac{\alpha}{\beta})$, however, since α is usually one or two order of magnitudes larger than β , the position of the maximum can be in many cases approximated with t_0 . The parameter β is related to the life-time of a research

topic. We found that there is a variation of this parameter before and after 1995. In fact, we have an average life-time of about 25 years before 1995 while the average life-time drops to just 10 years in the last period.

5 Conclusions and Open Questions

In conclusion, we present a simple model to describe trends in science and applied it to condensed matter physics. We found that different research topics are well represented by our model. In particular topics that had limited applications in research or industry. These topics show a fast growth followed by an exponential decay in publication ratio. On the contrary, other topics that became relevant for industry (as silicon, transistors, lasers, etc.) and/or for basic research (as lasers, many-body perturbation theory, density functional theory ...) show a growth comparable or superior to the total number of publications in physics. We learn from these results that while new topics spread in the scientific community in a few years, their decay time is of the order of decades.

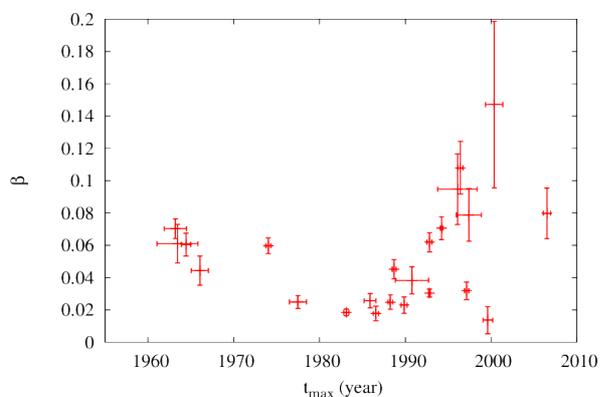


Figure 4: The β parameter as a function of t_{max} that represents the position of the maximum peak.

These results could be used for example to investigate “scientific bubbles”[16] or to discriminate between the importance of different sub-fields.[17] In fact one can expect that ‘bubbles’ will have a different decay in time than real research sub-fields, and try to detect them. Then we analyzed the α and β parameters that are related to the birth and death of a research sub-field. We found an increase of both β and α with time. Unfortunately, the data are too noisy to model this behavior and to draw reliable conclusions. Therefore the question we posed in the title remains open: Is research going faster and faster? Or in another way: Are trends in condensed matter physics changing faster than in the past? We have some indications but not a definite answer.

For sure, there are different possibilities to improve the present results that we have not considered yet, as for instance to extend this research to all physics branches, including more sources or combining lexical analysis with citation metrics. We hope this paper will stimulate future research on diffusion, birth, and depth of trends in the scientists network.

Acknowledgments

CA acknowledges Lorenzo Stella for useful discussions, and Elena Cannuccia for carefully reading this paper. This work has not been supported by any funding program.

- [1] C. Castellano, S. Fortunato, V. Loreto, *Rev. Mod. Phys.* **2009**, 81, 591–646.
- [2] R. Albert, A.-L. Barabási. *Rev. Mod. Phys.* **2002**, 74, 47–97.
- [3] M. E. J. Newman, *PNAS* **2001**, 98(2), 404.
- [4] R. Pastor-Satorras, A. Vespignani, *Phys. Rev. Lett.* **2001**, 86, 3200–3203.
- [5] W. Glänzel, *El profesional de la información* **2012**, 21(2), 194–201.
- [6] M. Zitt and E. Bassecoulard, *Scientometrics* **1994**, 30:333–351, 1994.
- [7] W. Weidlich, *Phys. Rep.* **1991**, 204, 1–163.
- [8] S. Nakayama, Y. Nakamura, *Physica A* **2004**, 337, 625.
- [9] J. Tague, J. Beheshti, L. Rees-Potter, *Library Trends* **1981**, 30, 125.
- [10] ADS, CfA, and NASA. SAO/NASA astrophysics data system abstract service.
- [11] All data are obtained from the SAO/NASA abstract service by means of a Python script that automatically performs the research and parses the results. We tried also to perform research on other databases as ISI Web of Science, APS journals or arXiv and we found a similar behaviour. However data in SAO/NASA are larger than the other mentioned databases and this allows to reduce fluctuation among the different years.
- [12] L. J. Anthony, H. East, M. J. Slater, *Reports on Progress in Physics* **1969**, 32(2), 709.
- [13] P.-F. Verhulst, *Correspondance mathématique et physique* **1838**, 10, 113.
- [14] S. Paula E, *Journal of Economic Literature* **1996**, 34(3), 1199.
- [15] V. I. Yukalov, E. P. Yukalova, D. Sornette. *Physica D* **2009**, 238(17), 1752–1767.
- [16] Abraham Loeb. <http://arxiv.org/abs/1108.5282>, **2011**.
- [17] Abraham Loeb. <http://arxiv.org/abs/1207.3812>, **2012**.