Scientometrics, Vol. 22. No. 1 (1991) 155-205

CO-WORD ANALYSIS AS A TOOL FOR DESCRIBING THE NETWORK OF INTERACTIONS BETWEEN BASIC AND TECHNOLOGICAL RESEARCH: THE CASE OF POLYMER CHEMISTRY

M. CALLON, J.P. COURTIAL, F. LAVILLE

Centre de Sociologie de l'Innovation, Ecole des Mines de Paris, 62 bd. St. Michel, 75006 Paris (France)

(Received August 20, 1990)

The goal of this paper is to show how co-word analysis techniques can be used to study interactions between academic and technological research. It is based upon a systematic content analysis of publications in the polymer science field over a period of 15 years. The results concern a.) the evolution of research in different subject areas and the patterns of their interaction; b.) a description of subject area "life cycles"; c.) an analysis of "research trajectories" given factors of stability and change in a research network; d.) the need to use both science push and technology pull theories to explain the interaction dynamics of a research field. The co-word techniques developed in this paper should help to build a bridge between research in scientometrics and work underway to better understand the economics of innovation.

1. Introduction

One of the major questions raised by the economy of technical change is the role played by scientific research in the innovation process. For many years, two diametrically opposed hypotheses were used in attempts to answer it: technology push and market pull. However, the majority specialists now recognize that the situation is much more complicated than these simple linear models imply. A network of interactions (techno-economic networks) between science (S), technology (T) and the market (M) explain the emergence of innovations and their diffusion. The nature of these interactions, their dynamic and their intensity depend upon the sectors studied and the period considered. (*Freeman,* 1982; *Nelson* and *Winter,* 1977; *Pavitt,* 1984; *Von Hippel,* 1988).

Once the importance of these interactions has been recognized, the need to develop tools to empirically describe the different forms they can take becomes evident. In this article, our goal is to show that scientometrics, and more precisely co-word analysis, can be useful in this task. This same assumption has been adopted by others. *Narin,* for example, has used patent citations to scientific articles as a means of studying the relationships which exist between science and technology. He has also looked at the correlation between highly cited patents and the economic performance of the firms that hold them. His work has produced a variety of useful indicators which can be used to describe the links between science, technology and market success (*Narin* and *Noma*, 1985; *Narin, Noma* and *Perry*, 1987).* Another study that merits attention is the one carried out by *Walsh* on the plastics industry (*Walsh*, 1984). Building upon work by *Freeman,* Walsh's research concerned the role played by science and technology in the long term dynamics of a science based industry. The indicators she employed were very simple: number of scientific publications; number of patents; and production volumes. Her analysis of S/T/M interactions was aimed at trying to understand the time lags she observed between these three variables. Despite the relative simplicity of the method' *Walsh* managed to demonstrate the rich variety of interactions which occur and, notably, the extent to which causal relationships can change over time.

These different studies show the utility of scientometry, or indeed technometry in describing the dynamics of techno-economic networks. However, the tools used up to now hove several serious limitations. The indicators employed show (or postulate) the existence of a link, but give no information concerning the subject or problem area concerned. In order to know if it is scientific research or, on the contrary, technology which has been the prime mover of an invention or an innovation, it is necessary to return to the documents themselves and to read the contents of the articles and patents identifies.

In this article, co-word analysis will be used both to describe the interactions which exist between different phases of the innovation processus and to show if basic research or applied research is the moving force. The results presented below come from a study concerning the impact of public funding on the development of a research field: in this case, polymer science.** The goal of this study was to understand the internal dynamics of this particular sector of science and technology and then, within this global framework, to compare the different research strategies

* The limits of this analysis are underlined: van Viannen et al, 1989.

** A variety of reasons led us to choose polymer science as a field of study. Among others, the following can be given: (a) it is the science base for the plastics industry, which makes it a very appropriate field for studying the relationships between science and technology; (b) it lends itself to the use of traditional bibliometric indicators and in particular to patents; (c) several excellent studies have already been carried out on the subject, which have constantly served as references to us in this study. For an exhaustive presentation of the results see (*M. Callon* et al, 1986).

of the teams which were awarded grants. In what follows, we will focus on the internal dynamics of the field and leave for another article our results concerning the impact of public funding on the development of science and technology.

Given that our goal is to demonstrate the interest of co-word analysis, two limitations must be kept in mind when reading this article. The first is that an exhaustive analysis of the domain will not be undertaken. Rather, we will present the technique, show how it can be used and provide examples which are chosen in order to illustrate how different types of results can be interpreted. The second is that we will concentrate our attention on a very limited part of the total techno-economic network picture in order to simplify the paper. A wide range of activities lie between basic research and technological invention. Generally, these activities are grouped together under the heading of applied research. Three poles can be thus identified: basic academic research, applied research and technological invention. The first two research activities use journal publications as a means of disseminating their results. Patents offer a more appropriate source of information for the third, although they are not the only source available. The results presented in this article concern specifically the interactions between basic and applied research. Over the twelve year period studied, the intensity of these interactions and the variety of their different configurations were greater than those we detected between either of these two types of research and technology. In a paper such as this, whose goal is essentially methodological, focusing on basic/applied relationships will allow us to present a wider range of results.

The paper is divided into three parts. In the first part we will present our data, that is, the various files that we constructed in order to carry out our study. In the second part, we will describe the co-word analysis techniques which enabled us to identify different research areas and to study their evolution over time. The results of this study will be presented in the third section. We will particularly focus our attention on the time lags observed in the interactions between basic and applied research. An interpretation will be given concerning the nature of the research which constitutes the driving force of activity in the polymer science field.

2. Presentation of the files and of the databases consulted

2.1. The files

In order to study the relationships between basic research and applied research, we draw a distinction between two main classes or files of publications.

– the first file (F1) is defined as made up of all the international literature in the field of polymer science, whether basic or applied – or even technical.

– the second file (F2), is the sub set of F1, containing all those documents which referred more particularly to what is generally called academic science.*

F1 and F2 should permit us to follow two types of production. F1 covers that of the totality of the international community of researchers or engineers who worked on or published in the area of polymer science. F2 highlights the part of this general production that corresponded to academic research. More briefly, we could say that F2 gives, by definition, an overview of basic or university research, whereas F1 gives a view of the overall organization of the field and in particular of the relationships between academic and applied research.

The two files F1 and F2 thus allow us to explore the relationships between basic research and what we could label general research (including basic and applied).

2.2. Choice and description of the base

By definition, a file contains all literature considered as relevant. For example, F1 should contain all articles dealing with polymers published by researchers anywhere in the world in any journal. In practice, such exhaustiveness is not possible. We must substitute for each file defined abstractly and generally a real, incomplete but representative file. This is far from being easy. The choices made are necessarily imperfect and difficult to justify rigorously. We will now present the database used and indicate how the actual files which we worked on were built.

* A third file (F3) corresponding to technological literature dealing with applications of knowledge and know-how about polymers (their industrialization and the production of potentially marketable goods) was also built.

We drew on the CNRS's PASCAL database to make up files F1 and F2.** This is multidisciplinary base in which each document is indexed by title, summary and a certain number of restricted key words (descriptors) drawn from a CNRS lexicon.

F1 was obtained by selecting from the PASCAL base all references having a link with the field of macromolecular compounds. There were too many references over the period from the setting-up of the PASCAL database in 1973 for us to be able to take the whole file into account over the period 1973-1986. In order to make the task manageable, we decided to limit ourselves to a few years, chosen so that we could follow the evolution in polymer research over the period under consideration. The years chosen were the following: 1985, 1980, 1975 and 1973. Table 1 gives the number of documents corresponding to these different years.

Table 1

Distribution of the publications in the different years studied in the set F 1

Year of publication	1973	1975	1980	1985
Number of publication	14170	15162	17297	9334

Distribution of the number of publications in the different periods in the set F 2

Year of publication	1973/75	1976/78	1979/80	1981/83	1984/86
Number of publications	3447	3106	2717	3456	2749

File F2 was selected from F1 by extracting from it articles appearing in following three journals:

- Journal of Polymer Science (Chemical and Physical editions),
- Journal of Applied Polymer Science,
- Die Makromoleculare Chemie.

** PASCAL contains more than 6 million abstracts from all scientific and technical fields. Each year some 500,000 references are added (492,000 in 1987) covering 8000 journals. French theses and unpublished reports are also included. The indexing by key word is done manually by ex-researchers or by specialist engineers from the fields covered. Since it relies in this way on human input, the indexing can be of variable quality. Further, the following information is given for each document: the name of the authors; their institutional affiliation (this information is not always given); the name of the source (title of the periodical, congress etc.); the country of publication; the date of publication. All the documents analyzed are classed by area using an exhaustive and detailed classification scheme.

As Zeldenrust's study has shown, these journals deal with what might be called academic research in the field of polymer science – in some sense they are its core. Their impact is not very widespread, however, since unlike other chemistry journals they are above all cited by other journals in the same field. This in turn confirms their representative character. In terms of number of publications – excepting Soviet journals – they nevertheless constitute the three main journals in all the field (Table 2).

Table 2
Main reviews to be found in the Polymer section of the PASCAL database

	Journal	Impact factor	Number of articles (1980)
1	J. Polym. Sci. Polym. Che. Ed.	1.03	389 } 623
2	J. Polym. Sci. Polym. Phys. Ed.	1.84	234 }
3	Die Makromolekular Chemie	1.18	349
4	J. Applied Polym. Science	0.89	288
5	Polymer	1.70	285
6	Macromolecules	2.60	280
7	Kunstoffe	0.44	235
8	Kauch i Rezina	-	217
9	European Pol. J.	0.85	201
10	Acta Polymerica	0.49	198
11	Polymer Engineering and Science	0.92	196
12	Plast. und Kautsch.	0.22	186

The number of references included in file F2 being fewer than in F1, we were able to cover the whole period (1973-1986). We drew up five sub-files, each one corresponding to two or three successive years. We give a detailed breakdown in Table 1. This division allows us to follow the development of the field in international academic science over the whole period of existence of the PASCAL base.

3. Co-word analysis: definition of indices and visualization of networks

In this section, we will give a brief presentation of the co-word analysis, emphasizing the new tools that have been developed during the last months for the analysis and visualization of the dynamics of problem networks (*Turner* et al, 1988). The methodological foundation of co-word analysis is the idea that the co-occurrence of key words describes the contents of the documents in a file. From a methodological point of view, it is therefore a question of using one (or more) index (or indices) to measure the relative intensity of these co-occurrences and to achieve simplified representation of the networks to which they give rise.

3.1. Measuring the links between keywords: the equivalence index

It is said that two key words, i and j, co-occur if they are used together in the description of a single document. It is clear that a simple counting of co-occurrences is not a good method for evaluating the links between co-words. Words which are used very frequently, and which are indeed used almost systematically in the indexation of the documents in a file being studied, will be advantaged with respect to words used less often. Thus we need to adopt statistical indices in order to avoid this bias. A systematic study of all such possible indices led us to choose one, which we have called the index of equivalence (*Michelet*, 1988). It is defined as follows.

Take file S consisting of N documents. Let C_i be the number of occurrences of the word i, that is say the number of times that it is used for indexing a document within S. Let C_{ij} be the number of co-occurrences of the words i and j, that is to say the number of documents that are described by both words in the set of key words used to index them. The coefficient whose value is given by the following equation will be called the equivalence index (E_{ij}):

$$E_{ij} = (Cij^2)/(C_i \times C_j)$$

The calculation of all coefficients between all possible word pairs (even if the value of these coefficients is often equal to zero) generates a considerable number of links – far too many to able to represent graphically. This is why we have developed algorithms for generating sub-groups that can be more easily visualized and interpreted.

3.2. The construction of clusters (or sub-networks)

The goal of co-word analysis is to bring the relationships between key words that might at a given moment be considered as the most significant. We thus want to locate sub-groups that are tightly linked between themselves and which correspond to centers of interest or to research problems that are the object of significant investments by researchers. In other words, if you imagine a graph linking key words to each other (the strength of these links being measured by the equivalence index defined above), we have to think of a procedure for isolating sub-graphs of words (which are represented by points) that are more strongly linked internally than with points not part of the sub-graph.

There are several possible ways of creating such a partition of the key word network associated to one file. The one that we have used for this study has the merit of avoiding problems arising from the choice of a – necessarily arbitrary – threshold. In the method chosen the clusters are formed with variable thresholds, characterized by the value of the first link refused, which is called the saturation threshold of the cluster. These threshold are automatically determined in such a way that no cluster contains more than ten words. It can ever have fewer than ten words if, by addition of a new word, two clusters are merged and the two of them together have more than the ten words. When a cluster saturates, a new cluster is started, the association value of the first link being this new cluster's "ceiling" threshold. Each cluster is consequently characterized by its "ceiling" and "saturation" threshold, which will vary from one cluster to the next.

This division into clusters of not more than ten words can lead to arbitrary divisions. We consequently keep track of the links which exist between words found in different clusters. The "external links" can be of two types. When the association value of an inter-cluster link is higher than the ceiling threshold of the examined cluster, this suggests that the latter is simply the continuation of another cluster that appeared before it and was already saturated. When the association value of the inter-cluster link is lower than the saturation threshold, this suggests that the examined cluster spills over into another cluster. In both cases, the division of clusters into groups of ten words is artificial and a new round of classification is needed. Three distinct categories in the initial list of clusters can be identified:

– *isolated clusters,* which are characterized by an absence (or low intensity) of links with other clusters. The only question regarding them is their internal homogeneity: they might consist of several sub-clusters that are worth identifying when examining the structure of their internal relationships;

– *secondary clusters,* whose external links with other clusters above the ceiling threshold are sufficiently strong that it is legitimate to consider that they are the natural extension of one of these;

– *principal clusters,* to which one or more other (secondary) clusters are associated by links whose value is lower than the saturation threshold.

In order to give these definitions an operational value, we need to lay down some general rules for determining the association of two clusters (of which one becomes the principal and the other the secondary). We have chosen to consider that two clusters are associated to each other if they are tied by at least three links.

Thus we have two lists for each file. The first give all the *isolated clusters*. The second consists of group each containing a *principal cluster* to which are linked (from "below") one or several *secondary clusters*. An isolated cluster, as well as a group consisting of a principal and several secondary clusters represents what can be considered as sub-networks (of the general network associated to one particular file) with their own coherence and autonomy.

Describing the dynamics of a sector comes down to identifying these sub-networks (we will use indifferently the two notions of cluster or sub-network for these associations of words in a file), to characterizing their content and to following their evolution. So as to simplify the analysis, we propose an additional distinction, whose aim is to select clusters with a strong ability to structure the general network. It is with this in mind that within the set of principal clusters in a file we highlight those that have at least two secondary clusters. We call them *crossroads clusters*. By their power to connect, crossroads clusters play an essential role in the transformation of a network.

One last remark can be made about this classification (isolated, crossroads, principal and secondary clusters). It could be considered that the distinction between principal and secondary clusters is decidedly artificial, since there is an underlying unity between groups that have been arbitrarily separated. In fact, from the point of view of analysis and interpretation, this distinction is a very useful one, and must be retained. The principal cluster in effect designates the heart of a given sub-network – it is in some sense its core. As a result, principal clusters – and in particular crossroads clusters – identify the focal and structuring problems of the field being studied. Any analysis must therefore start with them.

3.3. Characterization of clusters as a function of their participation in the organization of the network: centrality and density

The first stage of the description of a network (we mean the whole network of words for a given file) is the identification of clusters, the description of the links that unite them, and the representation of their internal organization. We then have to characterize the morphology of the network as a whole, and the contribution of each of these clusters to its structure.

A cluster can be defined in two different ways. Firstly, it can be seen as a point in a general network, one which is characterized by its position, that is to say by the bundle of links uniting it to other clusters/points in the general network. Secondly, it can be seen as a cluster, made up of words linked with each other – it itself defines a more or less dense network, one which is more or less coherent and robust. We need this double analytical perspective in order to appreciate the dynamics of the whole. In effect, the general network can develop in two ways. Firstly, by the reorganization of the relationships between clusters with a stable internal composition and secondly, by the reconstruction, redefinition of the clusters that make it up or by the appearance of new clusters (whether these emerge progressively or result from the fusion of existing clusters) or by the disappearance of clusters (which are either progressively wiped out, or split). These two mechanism are only rarely mutually independent. In general, one observes a modification of the content of clusters and of their list at the same time as a redefinition of the links that unite them. This complexity is the essence of research: it is only by making wild over-simplifications and looking at very large clusters that anyone can talk about specialities, fields or research themes that are stable over time. The problem that we are posing is that of the identification of what changes and what transforms. Thus we need to avoid a priori answers and furnish ourselves with the means and tools that allow an empirical response to the question. It is to this end that we bring in the following two notions: *centrality* and *density*. These are designed to bring out the contribution of different clusters to the structuring of the general network.

Centrality measures for a given cluster the intensity of its links with other clusters. The more numerous and stronger are these links, the more this cluster designates a set of research problems considered crucial by the scientific or technological community. To put it in the language of the sociology of translation, this proposition means that the cluster in question is an obligatory passage point (*Callon,* 1986). It is essential for anyone interested in the clusters associated with it to invest, directly or indirectly, in this sector. It occupies a strategical position.

The centrality of a given cluster could be measured in several ways. For example, we could sum the squares of all the links (measured by the index of equivalence) uniting it with other clusters.* More simply, and this is the method we have adopted,

* The sum of the values of the indices is divergent. This is why, when we want to take all the links into account, we sum the squares. This is the procedure adopted by *S. Bauin* and *B. Michelet*, 1988.

we can calculate the mean of the first six links with other clusters. Whichever measure is adopted, it permits us to order the different clusters in a file by decreasing order of centrality.

Density characterizes the strength of the links that tie the words making up the cluster together. The stronger these links are, the more the research problems corresponding to the cluster constitute a coherent and integrated whole. It could be said that density provides a good representation of the cluster's capacity to maintain itself and to develop over the course of time in the field under consideration.

The value of the density of a given cluster can be measured in several ways. We have chosen quite simply to calculate for each cluster the mean value of its internal links.* This evaluation allows us to rank clusters by decreasing order of density.

3.4. Strategical diagram as synthetic representation of one general network

It is difficult enough to give a static description of a network. It would seem to be impossible to give a satisfactory description of its dynamics. We need to simplify our descriptions, so as to render them usable while keeping them faithful.

Our point of departure is the list of clusters (whether isolated, principal or secondary). This list furnishes the simplest possible description of the network. More details can be added by listing the links between principal and secondary clusters, as well as by indicating which of the principal clusters are crossroads clusters. But such a list, which helps to highlight the sub-networks around which the field is organized, is not sufficient. We have to fill it out with a presentation of their respective positions and their relative stabilities.

The ideas of centrality and density allow us to give a synthetic and simplified presentation of the network's morphology, and provide a stepping-stone for a dynamic analysis. Since each cluster can be defined by its centrality and density, it is possible to trace what we will call – for reasons that will become obvious – a strategical graph. This graph is obtained by ordering clusters horizontally (along the x-axis) by increasing order of centrality, and vertically (along the y-axis) by increasing order of density. This operation allows us to classify all aggregates into four general categories, which correspond to the four quadrants of the graph (Fig. 1).

* In the document cited above, *Bauin* and *Michelet* have chosen a different measure, based on the evaluation of an index of representativity of the words in the cluster. For them density is the sum of the squares of the values of this index.

Clusters of *type 1* are both central to the general network (they are strongly connected to other clusters) and have intense internal links (they display a high degree of development). These clusters in some sense constitute the file's core. Their position is strategical, and they are probably dealt with systematically and over a long period by a well-defined group of researchers.

Clusters of *type 2* are central, that is to say that they are strongly connected to other clusters, but the density of their internal links is relatively low. Such aggregates, although strategical to the file under consideration, might in reality be the object of investments in other, connected, files. They correspond in this case to points of transfer between separate but linked networks. They can also signal the appearance within a given network of research problems that are becoming central, but which are not yet the object of significant investments: they are becoming mature, and their importance for the field is already indicated by their degree of centrality.

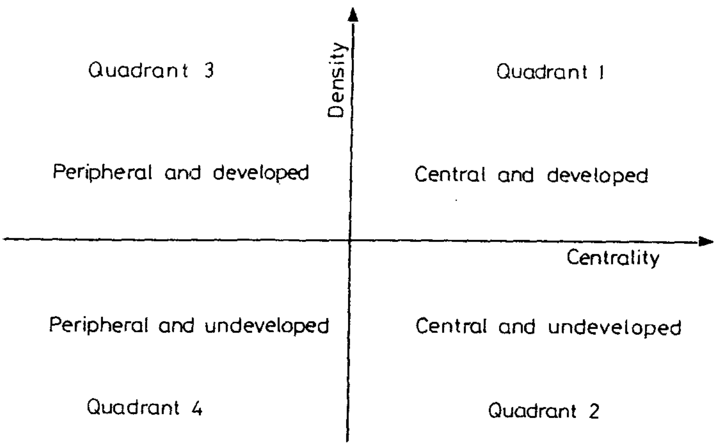


Fig. 1. Strategic diagram and characterization of the clusters in the file

Clusters of *type 3* are not central – we will call them peripheral – and the strength of their internal links leads us to suppose that they correspond to research problems whose study has already been well-developed. They might be clusters which at an earlier time were central, but which – while remaining the object of significant investments (it is not so difficult to explain such permanence) – have been progressively marginalized, generating less and less interest. They appear as specializations that interact weakly with the other sub-networks of the field under study.

Clusters of *type 4* are both peripheral and little developed. They represents the margins of the network. Only a dynamic analysis (the evolution of a network over several periods) or a comparative one (the relationship of the network with other networks) allows us to determine their contribution to the field.

The classification of clusters, depicted in the form of an easily readable strategical graph, gives a more exact description of the state of a given network, and of the position and degree of development of the clusters that constitute it. Of course, the fact that two clusters are close to one another in the strategical graph does not mean that they are closely linked to each other. The only conclusion that can be drawn from this observation is that their indices of centrality and density have neighboring values!

To sum up, we can describe a network as follows:

(a) a list of clusters, split into four classes (isolated, secondary, principal and crossroads) and identified by the most central key word of the cluster;

(b) each cluster is characterized by an index of centrality and an index of density;

(c) a strategical graph is drawn, which enables a classification to be made of the clusters into four major families, corresponding to the different *relative* values of their two indices – centrality and density.

This description provides the basis for a comparative analysis of the different networks under study.*

4. Analysis of the transformation of networks and of their interactions: presentation of the method

In the preceding section, we dealt with methods for the static description of problem networks. Now it remains for us to develop tools for studying the transformation of a network (or file) over time (dynamic analysis) as well as the relations, at a given moment or over several different periods, between two distinct networks (analysis of interactions). From a strictly methodological point of view, the analysis of the temporal dynamics of the various configurations taken by only one network (associated to one particular file) and the analysis of the interactions between several networks (or files) can be carried out in the same way. In both cases, the questions are the same. Given two (or several) networks of words, in what way

* A reordering of the documents concerned into their different clusters is made. This reordering opens the way for a set of descriptions in which factors such as the authors, their countries, journals published in and so on intervene.

are they comparable? How can we describe their similarities and differences? What do they have in common? How we get from one to the other? These are difficult questions. It would be vain to attempt to give an exhaustive answer to each of them, such is the complexity of the networks being compared. In order to achieve partial but satisfactory results, we will use as our point of departure the simplified description of the networks that we developed before. Then we will describe the procedure used for comparing two networks (belonging to one or several files).

4.1. Comparative analysis of networks: dynamic analysis and analysis of interactions

Take two networks, N1 and N2. These might be two networks at different times (for example the network of academic science in 1970 or 1980) or two distinct networks at the same time (for example academic science in 1970 and applied research in 1970), or finally two distinct networks at different times (for example academic science in 1970 and applied research in 1980). In each case, we will follow the same procedure, since the problems of comparison are identical. The different stages of the comparison can be broken down as follows:

(a) Comparison of clusters.

Let C_{1i} be the set of clusters of network N1 and C_{2j} be the set of clusters of network N2. The degree of similarity of any two clusters will be measured by the number of the words that they have in common (this is the simplest way of calculating what might be called their intersection index). We can thus draw up for each cluster C_{1i} in N1 a list of clusters in N2 having at least one word in common with C_{1i}. We can then rank these clusters by decreasing number of key words shared with C_{1i}. Once we have completed this operation for all the clusters in N1, we can make a table of the correspondences between the two lists, bringing out the clusters which have been found to have some similarity with each other. In order to simplify this table, we will then adopt a consistent rule of only retaining the most striking correspondences (for example, two clusters will not be deemed comparable if they do not have more than three words in common). We define a transformation index t in order to measure the degree of dissimilarity between two given clusters. This index is the division of the total number of words used to describe the two clusters (a word is counted twice if it is present in the two clusters) by the number of words in common. If the cluster C_i is defined by 7 words and the cluster C_j by 4 words, and if 4 words amongst these 11 words are common to the two clusters, the transformation index is:

$t = 11/4 = 2.8$

As Figure 2 shows, several clusters in N2 can correspond to one cluster in N1. These correspondences have to be examined on a case by case basis as a function of the questions being asked. For example, we will see that the cluster *epoxy resin* that comes from the file of basic academic research is comparable with several clusters of a file containing articles closer to industrial applications (there is nothing surprising in this – there are numerous industrial uses of resins). And we will decide to follow one cluster rather than the other according to whether we want to investigate epoxy resins per se, or more generally the ways that resins harden.

(b) Comparison of positions on the strategical graphs.

Once we have established the correspondences between the clusters of two networks, the next stage in the comparative study of these networks, is an examination of the position on the strategical graph of those clusters which proved to be similar.

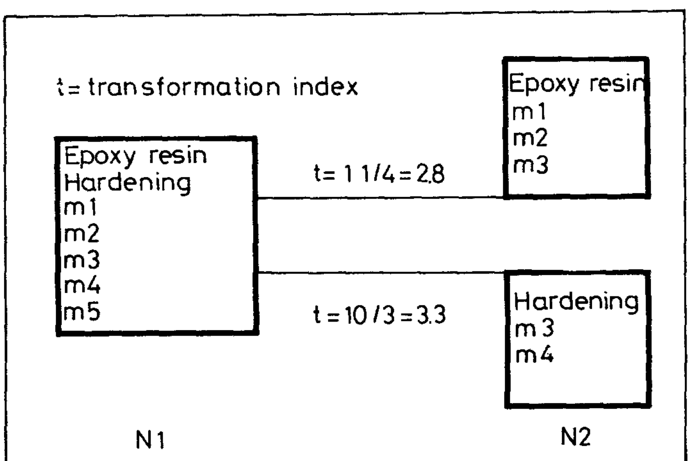


Fig. 2. Comparison between clusters of networks N1 and N2

This examination leads to finding such as, for example, that such and such a cluster that is central and dense in N1 occupies a peripheral position and is little developed in N2. Or perhaps another cluster that is central but not yet very developed in N1 is in quadrant 1 of N2. These analyses provide a wealth of information, whether we want to compare a single network at two different times (*dynamic analysis*) or to examine the interactions between two different networks at the same time or at different times (*interactional analysis*). They help us go beyond a simple enumeration of correspondences between clusters and bring out the relative position and degree of development of "similar" clusters within their respective networks.

(c) The cluster's life cycle curve.

In the case of the dynamic analysis of one file, it is possible to deepen the study by following the change in the indices of centrality and density and also the progressive transformation of content, of clusters that remain stable over time. Consider, for example, the network R at times T0, T1, T2, T3, T4 and T5. Suppose that a comparative analysis of the states of the network at these different times has identified a set of similar clusters. C10 from the first period corresponds to C11 from the second; C11 corresponds to C12 from the third period, which corresponds to C14 from the fourth and so on: we will call series such a set of similar clusters.* Other series can be found: C20 (T1), C9 (T2), C34 (T3), C25 (T4)...; C30, C35, C3, C76 etc. One series might come to an end, another might only start after a certain time (for example at T5). It is clear that the more stable a network is, the more series there are indicating the temporal propagation of its clusters. The existence of these series in itself provides interesting information. But it is easy to enrich this analysis by bringing out, for each series, the following:

– The progressive transformation of the clusters through time. In effect, it is very possible that by the principle of construction of these series cluster C25 id very different from C20, and that it does not even have a single word in common with it. In this case, progressive deformations have led, after a certain time, to a radical redefinition of the research problems. These latter retain, however, an affiliation with each other. Inversely, in other cases the series gets set up without the content of the cluster being much transformed. As it turns out, C76 is very like C30. We can display the degree of these transformation quite simply by drawing a graph in which the x-axis corresponds to time and the y-axis to the cumulative transformation of successive clusters (measured by the reverse of their intersection index). The steeper the slope, the faster the transformation of the clusters; the flatter it is, the greater their stability.

– The evolution of the indices of centrality and density over time for a given series. The variations of these two values, for a given series of clusters, are followed over the course of time and then mapped along two axes: the x-axis corresponds to

* The coefficient associated with each cluster is an order number of the cluster in the network at a given period.

time and the y-axis to the density and the to the centrality. On the same graph we also put the percentage and the cumulative percentages of documents corresponding to each cluster in the series (this allows us to follow the volume of activity over time and also their cumulative total). These graphs provide valuable information. For example, they might demonstrate that a given cluster becomes more and more (or less and less) central over time, and that at the same time its density increases (or diminishes). This information can be related to the significance of the effort (in terms of the number of documents) devoted to these research problems by the community of specialists. As we will see below, these graphs can be classified into several large groups, which recall product life-cycles studied by economists.

All of this information (the rate of transformation of the cluster in a series, the evolution of the indices of centrality and density, the number of documents in the clusters of the series) can be set out in a single diagram, which provides a picture of the network's dynamics (Fig. 3). For example, we might learn that a given cluster is becoming more and more central over time, and is receiving ever more attention, but that at the same time as its research problems are becoming strategically important, they are being profoundly reformulated. This latter is shown by the degree of transformation of the clusters over time. Of course there are other possible configurations – some of which we will encounter in the empirical study below.

4.2. Some practical rules for simplifying the comparative analysis

There are no particular difficulties attached to using the procedure that we have just described. However, just going ahead and applying it mechanically might produce a wealth of results that it would be difficult to synthesize. In order to simplify the analysis, we have adopted a course of action whose object is to firstly bring out the largest transformations, when studying the dynamics of a network, and the most prominent ones (both in terms of their permanence and the significance of the clusters they relate to) when looking at the relationships between several networks. This does not rule out a more fine-grained and exhaustive analysis at a later time.

In order to characterize the major changes and the most significant interactions, we decided to limit our comparisons to those clusters which, at a given moment and/or in a given network, were *crossroads clusters*. This means that we drew up a complete list of all the crossroads clusters of the relevant files for all the relevant

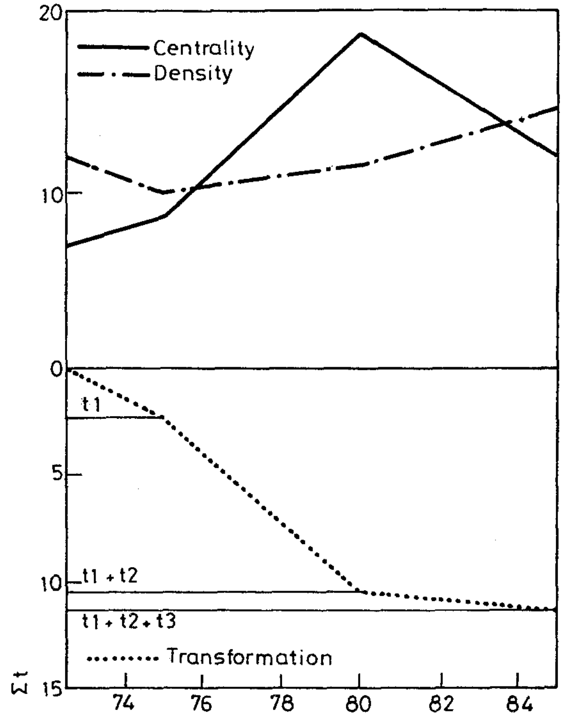


Fig. 3. Visualization of a clusters set evolution in the same file: the three curves of centrality, density, and transformation

periods. Each of these clusters was looked at, in all files at all times, and was used as the point of departure for the comparative analysis. The aim of the latter was to establish: (a) for a given file the series of clusters between which there were connections; (b) the correspondences between clusters in several different files. This analysis, which deals exhaustively with all the crossroads clusters (for which we draw up a list of the clusters to which they correspond in the networks under consideration) is not strictly limited to them. In fact, it may very well be that a crossroads cluster of N1 at T3 corresponds to an isolated, secondary, or even principal but not crossroads cluster of N2 at T2. In this case, this cluster is included in the comparative analysis, even though it is not a crossroads cluster. The more general rule is as follows: of all correspondences, we only consider those that (at some point in a series or in a comparison between files), include at least one crossroads cluster. The meaning of this rule is clear: we limit ourselves to evolutions and interactions dealing with research problems which at some given moment have a strong influence on the structure of the networks they are part of.

5. Academic and general networks: analysis of their internal dynamics and their interactions

In this part, we will look in turn at the evolution of the networks linked to F1 and F2, and then at their interactions. This procedure followed is that defined in the first part of this paper. We shall limit the presentation to some significant examples in order to illustrate the method and to show its interest.

5.1. The internal dynamic

One simple way of determining what themes played a role in structuring a file over time is to establish the list of the series of clusters containing at least one crossroads cluster. Table 3 gives this list for F1 (no cross next to a theme can have two meanings: either it is absent, or it is isolated).

For example, we see that there is a series of clusters covering 1973, 1975, 1980 and 1985 treating the study of *water-based solutions of polyelectrolytes.* This means, it will be remembered, that the successive clusters belonging to this series (for example that of 1973 and 1975, or that of 1975 and 1980) have a significant number of identical key words in common. We stress that this series may very well lead via successive transformations to a situation in which the cluster for 1975 and the cluster for 1985 only hold very few words in common. Also, this series has been retained because at least one of the clusters that make it up is a crossroads cluster (the others might be principal without being crossroads, or even secondary). Furthermore, as the table shows, certain series come to an end either definitively (there are no similar clusters for following years) or temporarily (such as, for example, the cluster *stereospecific polymerization*). Still others begin late in the day, demonstrating the birth and then the development of new lines of research. This turbulent history – which it will be recalled is limited to the most notable events and to the most significant regroupings or most stable factors – is clearly traced in Table 3.

For each of the series, we give the successive positions in the strategical graph of the clusters that make it up. For example, we learn that the cluster *Thermostable* has followed a tortuous path, starting in quadrant 3 and then oscillating between quadrants 1 and 2. Others remained in the same quadrant for the whole period.

This type of analysis can be carried out in the same way for File F2, which includes all articles appearing in journals constituting the core of academic research in the field of polymer science. Table 4 was drawn up following the same principles as Table 3. It presents all the series which contain at least one crossroads cluster. For each of these series we indicate the year of appearance or disappearance of the cluster (it will be recalled that a cluster is not considered present if it is isolated), as well as its movement in the strategical graph.

(a) The series of the clusters: appearances, disappearances, transformations and movements.

Three examples will serve to show how the dynamic of a research network can be analysed through cluster changes. The first case is in the paint binder series in F1 which is present throughout the period. The two other cases, taken from F2, concern the thermostable series (also present over the whole period) and the NMR spectrometry techniques which appear as theme within the periods studied. The following comments give the flavour of the types of interpretation authorized by our examination of this series of clusters:

Table 3

Transformation of all articles "Polymers" network (F1) from 1973 to 1985

	73	75	80	85
Thermostable	3	1	2	1
Radiochemical polymerization	1	1	3	3
Moulding/Packaging	1	2	1	1
Free radical polymerization	1	1	1	1
Reinforced plastics	1	1	1	1
Binder (paint)	3	1	1	3
Polymerization/Preparation/Condensation	2	2	2	1
Polyelectrolyte aqueous solution	4	1	1	1
Light scattering	1	2	1	1
Reaction mechanism	2	2		2
Stereospecific polymerization	2	1		1
Complex catalyst polymerization	1	1		2
Viscosity solution	1	2		
Ion exchange resin	1	3		
Grafting		1	2	2
Mechanical property		2	2	2
Natural rubber		1	1	1
Characterization crystallinity		2	1	2
PVC plasticizer		4	4	
Charge transfer compound		4	1	
Productivity		1	1	1
Paper pulp			1	1
Wood protection				1
Electrolytic polymerization				1
NMR spectometry characterization				1

For each non isolated cluster, location on the strategic diagram (number of quadrant) is indicated 1: upper right 2: lower right 3: upper left 4: lower left The absence of a number indicates either the cluster is absent or isolated: in both these cases, it is not included in the structure of the network.

– The series of clusters *paint binder* follows a very interesting trajectory. Table 5 brings out the relative stability of the contents of the cluster. This series of clusters clearly corresponds to a research theme that remains for the entire period studied the object of a full-fledged research programme that develops in time: this is confirmed by the evolution of the index of density, which remains high and even increases over time. What varies, however, is the degree of centrality of the clusters in the series, which move in the strategical graph from quadrant 3 to 1 and back again. At the middle of the period, the cluster becomes central (while remaining

Table 4
Main clusters in F2 and their evolution in the strategic diagram

	73/75	76/78	79/80	81/83	84/86
Condensation pol./thermostable	1	1	1	1	1
Viscosity molecular weight	1		2		2
Reverse osmose membran	1		3	4	3
Reaction mechanism	1	1	2	1	
Grafting	2		2	1	
Characterization styrene polymer	2	1			
Block copolymer	1	1			
Crystallization spherulite	2	1			
Reinforced plastic	3	3			
Crystal structure determination	2	2	2	2	2
Mechanical properties	1	2		1	
Reactivity/copolymerization ratio			1	1	
Conductivity			1		3
Stereospecific polymerization			4		1
Liquid crystal				1	3
Glass transition characterization				2	2
NMR spectometry characterization				1	1

For each non isolated cluster, location on the strategic diagram (number of quadrant) is indicated 1: upper right 2: lower right 3: upper left 4: lower left The absence of a number indicates either the cluster is absent or isolated: in both these cases, it is not included in the structure of the network.

coherent and developed). At this time it is closely linked to other clusters (see Tables 6). This evolution is confirmed by the fact that in 1973 the cluster *paint binder* is principal but not crossroads. In 1975, 1980 it becomes a crossroads cluster, and then is transformed into a principal cluster in 1985. This spectacular change would seem to prove that this cluster has progressively become autonomous in the course of an evolution which has led it to become linked to several other research sectors, and to profit from this contact. It has done this while maintaining its coherence and its high degree of development. If we wish to go further with the analysis, we can identify the clusters to which it was linked in 1973, 1975, 1980 and 1985. As a principal cluster in 1973, it had one satellite (and secondary cluster) called *aminoplast phenolics*. In 1975, as a crossroads cluster, it was linked to two secondary clusters, the first being *constraint deformation* and the second *glass transition temperature*. In 1980, still a crossroads cluster, it was associated with two secondary clusters: *anticorrosion paint* and *evaporation solvent*. In 1985 it is only linked to one cluster: *solvent*. This simple enumeration demonstrates the variability over time of the cluster's context. This variability and instability are made all the more apparent by the fact that the secondary clusters that are dependant on it never appear in Table 4. These clusters are at no point central to the structure of the general network. If we wanted to know more about the nature of the links with the various secondary clusters at different periods, all we would have to do is display the cluster's external links – that is to say the key word in the secondary clusters to which it is associated.

Table 5
Evolution of "Paint binder" cluster over time

	1973	1975	1980	1985
Binder	*	*	*	*
Alkyd resin	*	*	*	*
Paint	*	*	*	*
Varnish	*	*		
Acrylic copolymer	*	*		
Prothesis	*			
Dentist	*			
Paint film		*	*	*
Moulding compound		*		
Formulation		*		
Chlorinated rubber		*		
Anticorrosion material		*		
High solid paint			*	*
Water based paint			*	*
DIN standard			*	
Use			*	
Emulsion paint				*
Vegetable fat				*
Aqueous dispersion				*
Paint film formation				*

Table 6a
Associated clusters to "Paint binder" (F1) – 1973

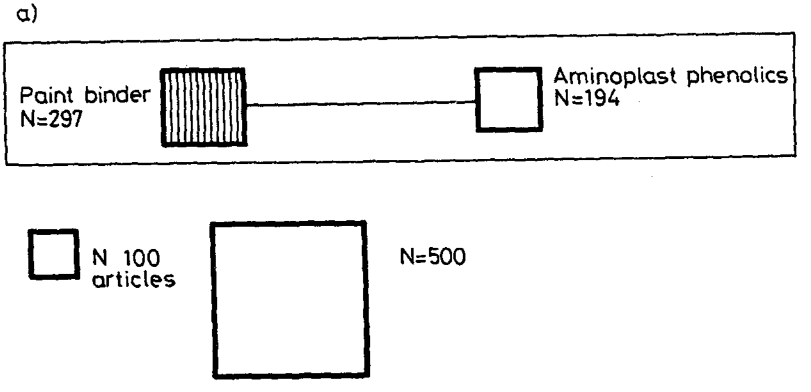


Table 6b
Associated clusters to "Paint binder" (F1) – 1975

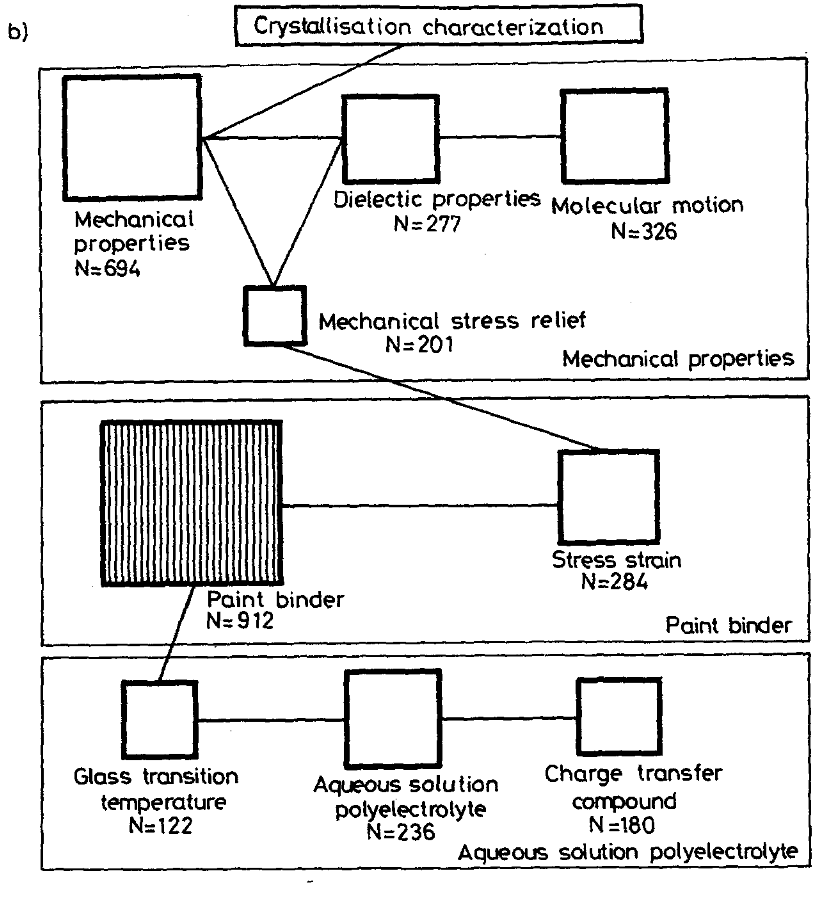


Table 6c
Associated clusters to "Paint binder" (F1) – 1980

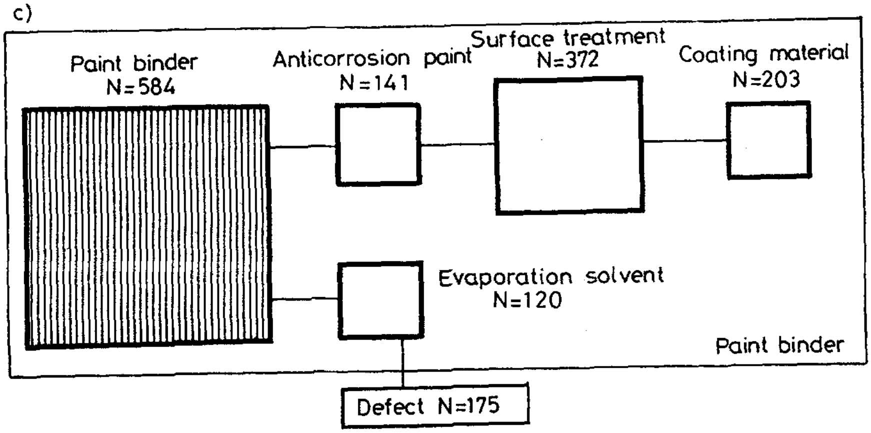
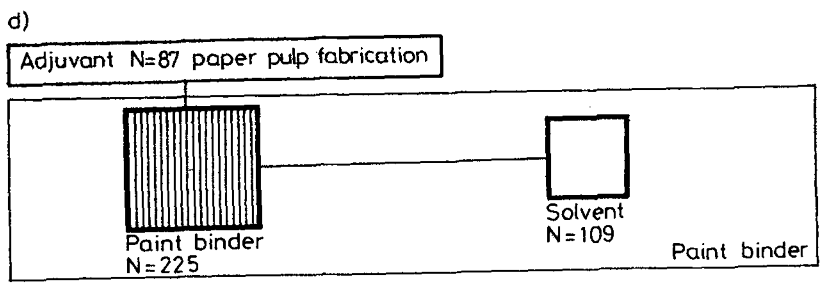


Table 6d
Associated clusters to "Paint binder" (F1) – 1985



The final stage in this examination is the drawing of a graph depicting the evolution of the indices of density and centrality, as well as the number of documents defining the cluster. The graphs we have drawn confirm the preceding interpretations, and go further. The index of centrality reaches a maximum in 1980, the form of the graph being an inverted U. This translates the evolution described above. The theme only holds a strategical position at the middle of the period. The index of density, on the other hand, increases continually until it reaches a plateau. This confirms that despite the variations in its centrality (and its eventual "collapse") the cluster maintains and even increases its internal coherence. But this evolution is accompanied by a progressive disaffection of the research community with this theme. In effect, the relative number of documents (as a percentage of the total number of documents in the file) goes up from the start of the period covered to reach a maximum in 1975, precisely at the moment when the cluster begins to become central (which it still is in 1980). It then begins to go continuously downwards. Everything seems to point to researchers anticipating the lessening centrality of the cluster, whose strategical position is weakening. In 1985 the theme is in much the same position as in 1973: there are still significant investments, but unless something unexpected happens in the future the theme has settled down to a tranquil rate of development. A small community of specialists deals with it. Figure 4 gathers all this information together in a single graph.

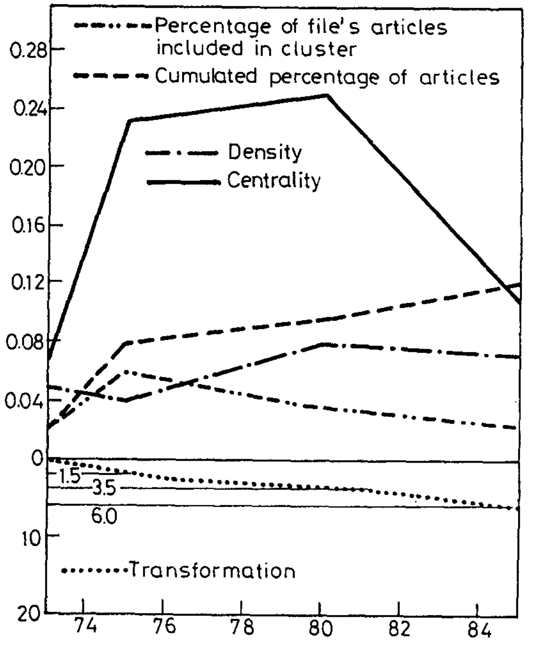


Fig. 4. "Paint binder" evolution in F1

– In F2, the cluster *thermostability* constitutes (as in F1) one of the themes that remains constantly strategical. Indeed, it remains in quadrant 1 from 1973 to 1975. Table 7 enables us to follow this cluster's transformation over time. Between 1973 and 1976, its content changes markedly.* It then remained stable from 1976 to 1983,

* For a similar description see: *Oehler* and al, (1989).

and then there was further change (1984-1986). At first, the cluster was largely centered on polycondensation; by the end of the period interest had clearly shifted to questions of behavior in heat and fire. Between these two extremes, the problematic was hybrid: polycondensation was studied as one of the methods of polymerization that might be of interest with respect to thermostability.

Table 7
Evolution of "Thermostability" cluster over time

	73/75	76/78	79/80	81/83	84/86
Preparation	*	*	*	*	*
Condensation polymerization	*	*	*	*	
Polymerization	*				
NMR spectrum	*				
IR spectrum	*				
Benzoic acid	*				
Aminoacid	*				
Thermostable	*	*	*	*	*
Thermogravimetry		*	*	*	*
Differential thermal analysis		*		*	*
Pyromellitimide polymer		*			
Polyimide					
Nylon			*	*	
Aromatic polymer			*	*	
Solubility			*	*	
Amine polymer			*		
Diepoxyde			*		
Diamine			*		
Thermostability				*	
Benzimidazole derivative pol.				*	
Decomposition temperature					*
Oxygen index					*
Fire resistance					*
Phosphorus polymer					*

Table 8, which gives the associations between clusters at different periods and thus enables us to see research contexts, confirms this interpretation. In 1973-1975, the crossroads cluster *polymerization by condensation* was to all evidence centered on concerns intrinsic to polycondensation. In 1976-1978 there was a significant transformation, since the cluster, which remained crossroads, could now be found at the center of a constellation merging polycondensation and problems of thermostability. This movement was accentuated in the next two periods, over the course of which thermostability became a crossroads cluster and polycondensation was distributed over secondary clusters. Finally, in 1984-1986 polycondensation disappeared from the list of principal or secondary clusters.

Table 8a
Associated clusters to "Thermostable" (F2) – 1973-1975

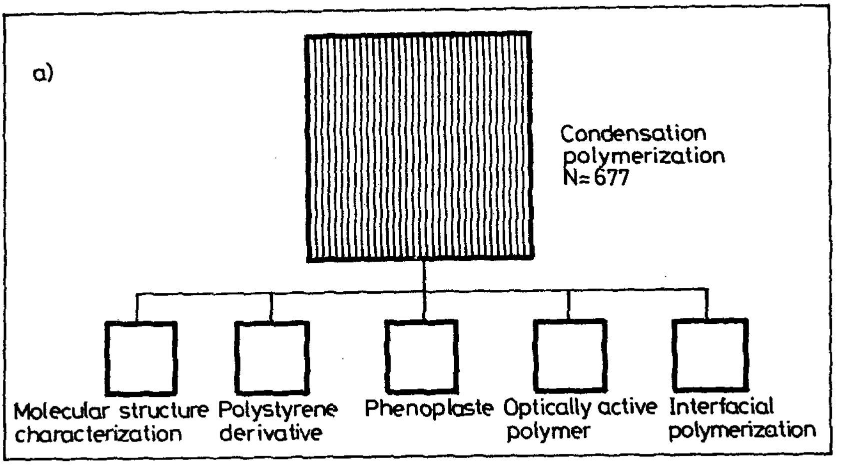


Table 8b
Associated clusters to "Thermostable" (F2) – 1976-1978

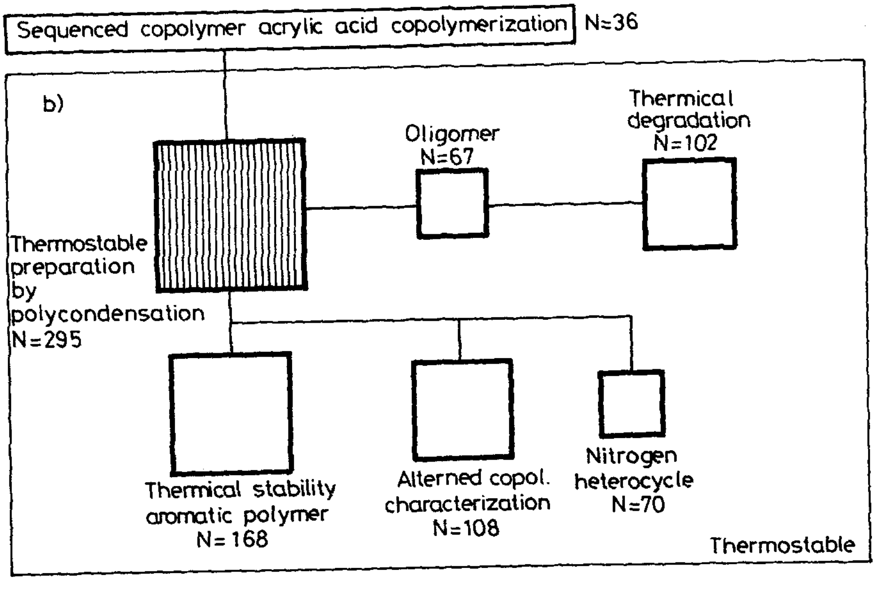


Table 8c
Associated clusters to "Thermostable" (F2) – 1979-1980

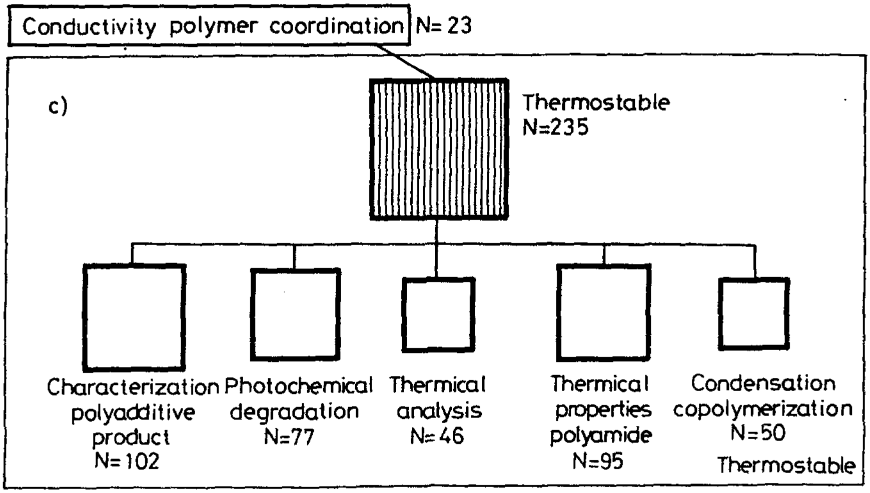


Table 8d
Associated clusters to "Thermostable" (F2) – 1981-1983

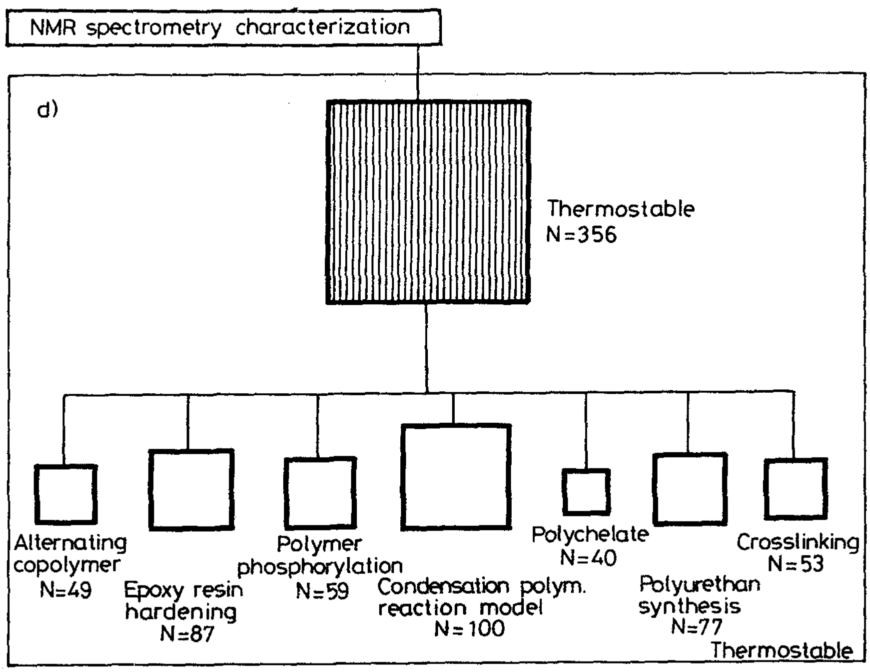
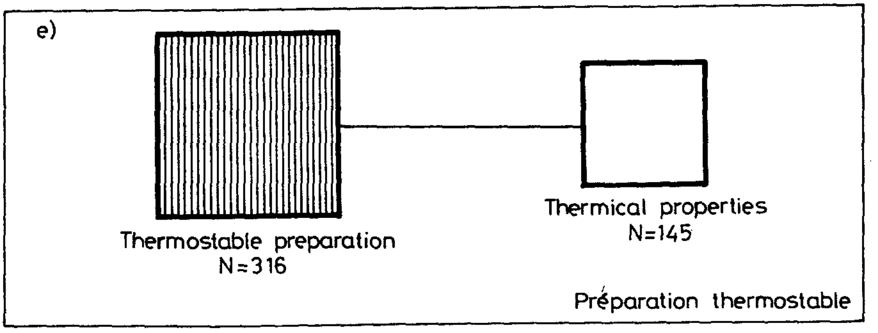


Table 8e
Associated clusters to "Thermostable" (F2) – 1984-1986



All these comments could be confirmed and synthetized by the evolution of the indices of density, centrality and transformation, as shown on Fig. 5. These graphs clearly show that the increase in the centrality of the theme results from a major change in its content. In order to maintain their research programme, the scientists are obliged to adapt it to a changing research context. By acting in this way, they contribute themselves to structuring this context.

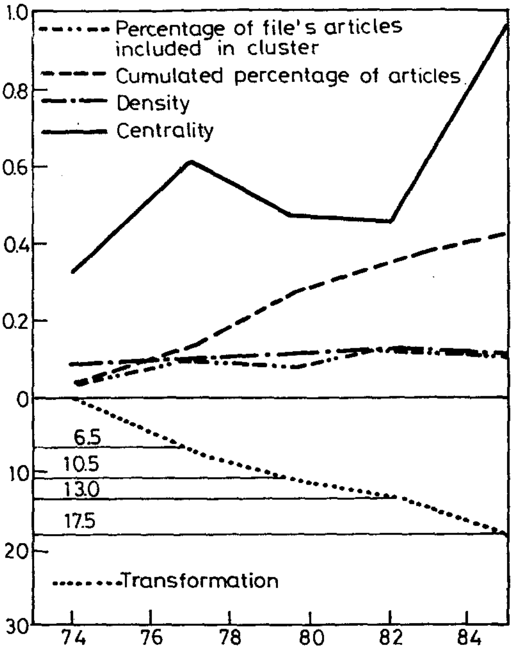


Fig. 5. "Thermostable" evolution in F2

– The cluster *Characterization by NMR spectrometry* made its appearance in the fourth period (1982-1983). It came into the first quadrant, where it remained in the fifth period. Table 9 shows that the contents of this cluster have a very stable, technological, kernel. An examination of its associated clusters (Table 10) brings out the relative diversity of the contexts of this technology's applications. Figure 6, which gives a representation of the evolution of the indices, brings out the fact that this theme was already an old one, but that it remained relatively unconnected to the rest of the network. This latter only integrated it during the fourth period. Its centrality and even its density went up constantly, at first spectacularly. The number of articles also went up. All this tends to show that this cluster has played a more and more important role in academic research on polymers. It has become an obligatory passage point. However, it remains to be explained why this theme, whose centrality increases regularly, does not experience a proportional increase in its density. One possible explanation could be that studies devoted to spectrometric techniques are a source of methodological inspiration, but not really a subject of interest in their own right for the polymer science network. Research specifically interested in developing the techniques themselves is no doubt undertaken elsewhere (in another network).

Table 9
"NMR spectometry characterization" cluster changes over time

	81/83	84/86
NMR spectrometry	*	*
NMR spectrum	*	*
IR spectrum	*	*
Hydrogen 1	*	*
Carbon 13	*	*
Proton	*	*
IR	*	
UV spectometry	*	
Cross polarisation	*	
Magic angle		
IR spectometry		*
Fourier transformation		*

Table 10a
Associated cluster to "NMR spectrometry characterization" – 1981-1983

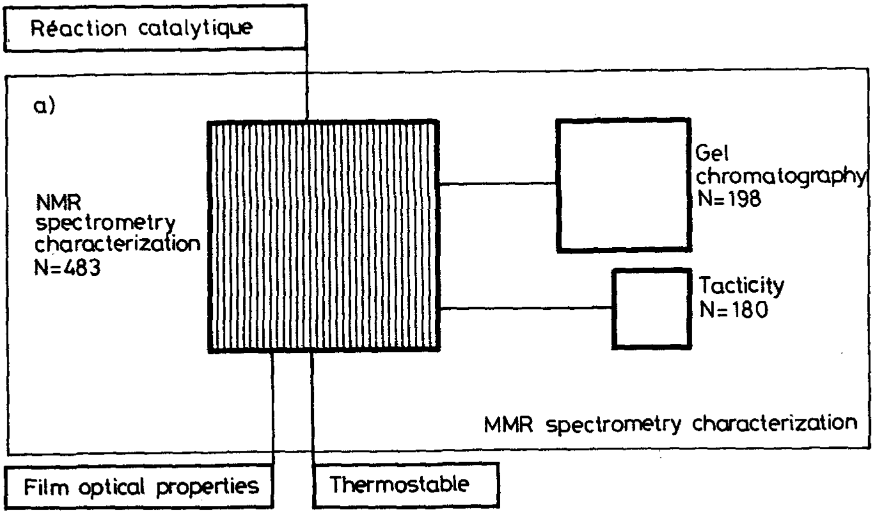
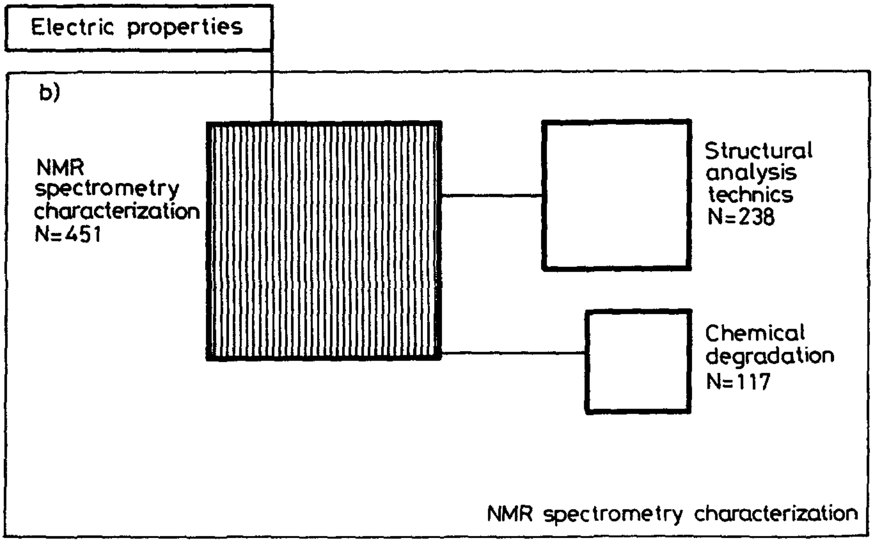


Table 10b
Associated cluster to "NMR spectrometry characterization" – (F2)-1985



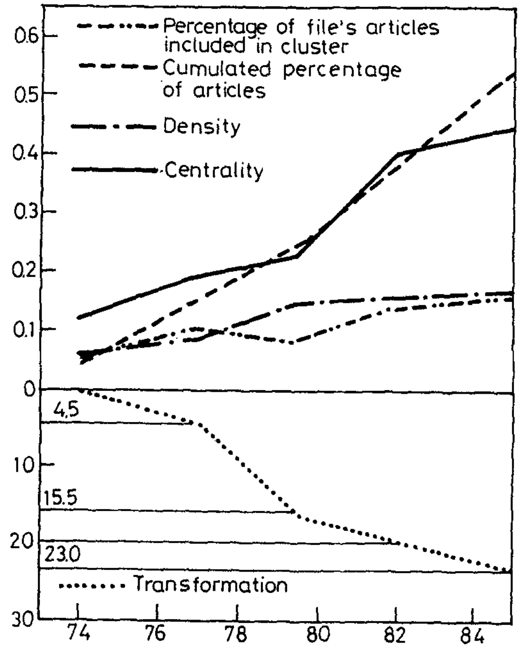


Fig. 6. "NMR spectrometry characterization" evolution in F2

(b) Life cycles of clusters.

The various cases we have looked at in our study bring out three main groups of graph:

– The first group contains graph of centrality and density which pass through a single maximum over the period being considered. The (relative) number of documents evolves identically (with one peak for the period), while the total number of documents corresponding to the cluster (still as a percentage of the total number of documents) increases constantly (Fig. 7a). There is one parameter that enables us to distinguish several types of evolution within this group. This is the order in which the two maxima (that of the indices of density and centrality) appear, as well as the interval of time that separates the two peaks. There is frequently, particularly in the case of academic science, a precession of centrality with respect to density. The cluster begins by linking itself closely with other clusters, then undergoes significant internal development. It is as if the search for outlets and strategical positions were a preliminary to committing the investments necessary to developing of a line of research.

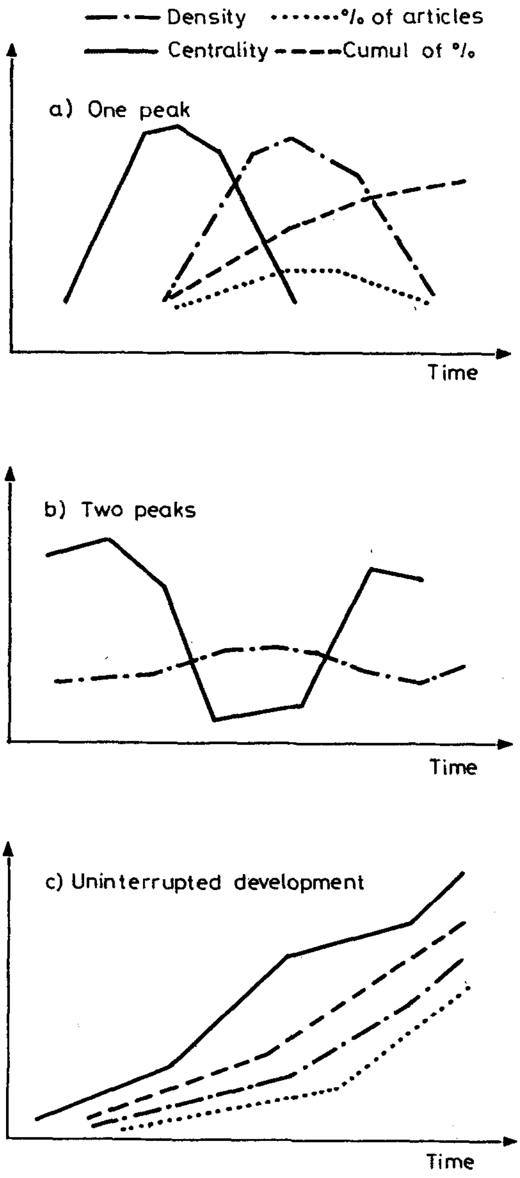


Fig. 7. Three kinds of cluster: centrality, density, % of articles, cumul of % of articles over time

– The second group corresponds to graphs of centrality and/or density with two maxima (Fig. 7b). In this case, one very often observes a profound change in the cluster's content. There is continuity, but it is as if a new series had appeared, linked to the preceding one but following its own line of development. Here again the index of density lags behind the index of centrality to varying degrees.

– The third family includes those graphs which increase or decrease constantly over time. Here centrality, density and the number of documents vary continuously in the same way (Fig. 7c).

This classification into three groups should not be considered as a definitive and general result. It is, for example, possible that the third group is a special case of the first. We would need to look at other files in order to validate the result.

One important point is that the analysis we are proposing differs from traditional studies, which only follow the change in the volume of publications. Unlike graphs such as the economists' life cycles of products, ours allow us to map the position of objects (here knowledge) in the networks they are defined in, and to show the significance of the transformations they undergo. We think that this tool should be useful for planning (*Courtial* 1989), since it enables a better grasp of the conjoint evolution of themes and their contexts.

(c) The overall integration and development of F1 and F2.

Up until now, we have characterized the evolution of clusters by following their indices of centrality and density. It is possible to apply this analysis to the whole of a network, thus measuring the overall degree of integration and development of the network. The simplest way of doing so is to calculate the average centrality and density of all the clusters composing the network at a given moment. The mean index of centrality gives an indication of the connections of the clusters between themselves. The mean index of density provides information about the degree of coherence of each of the themes constituting the field.

The Table 11 traces the variation in these mean indices over time for F1.

Table 11
The mean centrality and density indices over time for F1

Year	73	75	80	85
Mean centrality	143	175	205	182
Mean density	52	65	74	62

It can be seen that two indices increase until 1980, at which time they reach their maximum and then decline. We must be prudent in our interpretation of such an evolution, since we are only dealing with a short period of time. But it seems to be the case that from 1973 to 1985 the field's themes became unified and increased in coherence and integration, and that they were then torn apart by centrifugal forces. The interest of such an observation is apparent. It goes beyond the classical analysis of research fields, which too often looks at a single dimension – that of maturity, to the extent that it describes their dynamic. In reality, networks can go through phases of integration and redistribution, of deployment or cut-backs, without any simple evolutionary law being applicable to them.

As for F1, we can calculate the mean indices of centrality and density of F2, so as to chart their evolution over time (Table 12).

Table 12
The mean centrality and density indices over time for F2

Year	73-75	76-78	79-80	81-83	84-86
Mean centrality	128	138	160	162	198
Mean density	102	121	133	94	87

Two things can be read from Table 12:

– the constant increase in the mean centrality indicates an increased integration of the overall themes of this network;

– the passage through a maximal density shows that at the same time as the network becomes integrated, the themes that compose it display internal diversification.

With all due prudence, it seems that we can say that over the period 1973-1985 the academic network diversifies itself and becomes more polycentric. At the same time there is increasing integration within F1. It seems, therefore, that the field as a whole has become more bottom-up, and that as a result "fundamental" research has developed its ideas and new themes as a response to problems encountered by various technological applications.

5.2. Interactions between networks

We showed in the last section how the dynamics of a given network (in this case: F1 and F2) can be analyzed. Now we turn to an analysis of the interactions between two different networks with a view to examining the mechanisms by which relationships between academic and general research are established. The ultimate goal of this analysis is to be able to respond to questions like the following. Was such and such axis of research initiated by industrial or by academical laboratories? When was a given theme developed in the university taken up by technologists? Did academic scientists continue to study it? Were the contents modified, and how? In brief, we are attempting to find a way of comparing networks and their evolution which gives a representation of the dynamics of a field of research, in which heterogeneous actors, often with varying or contradictory goals, interact.

The method adopted for the comparative analysis of two different files is the same as that used for studying the dynamics of a single one. For the purposes of this presentation, we will limit ourselves to the analysis of a few typical interactions over the period studied between F1 and F2. This will lead us to distinguish between those cases in which the initiative (over this period) came from applied research, and those cases in which on the contrary academic research was the prime mover.

As will be seen however, these interactions are always very complex, involving a number of trials and errors and significant transformations of research content.

In order to better understand these transformations and enter more directly into the content of research being carried out in a given context, we will now consider all the links a cluster has with the other clusters in the network. Up to now we have only taken into consideration the clusters associated by more than three links. From now on we shall also consider clusters associated below this threshold. The links are represented by a continuous line when their number is equal to, or more than, three, and by a dotted line in other cases.

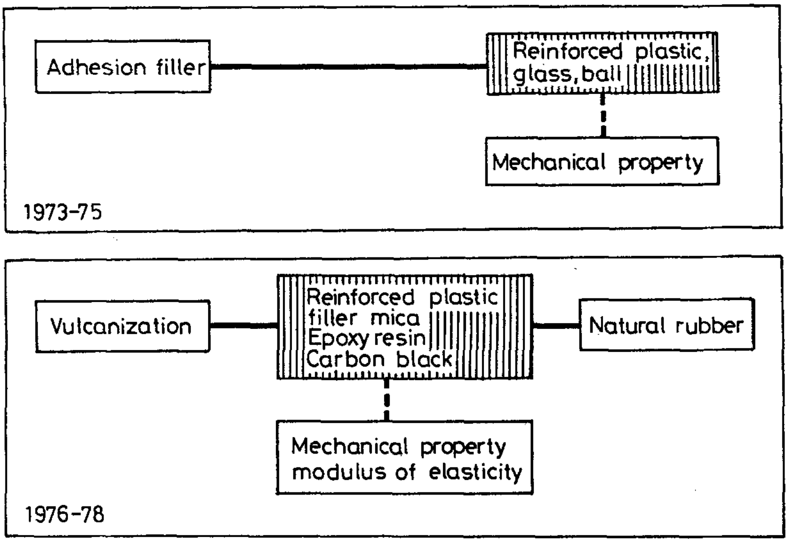
(a) From the academic network to the general network.

A systematic comparison of all the clusters (the point of departure for the comparison being a crossroads cluster: see Tables 4 and 5) identified in F1 and F2 at different periods can help us to establish the similarities between the two networks. That is to say, it helps us to identify correspondences between clusters (such and such a cluster from F1 at T1 is comparable to another in F2 and T2 and T3 etc.). an interesting case is that in which one of the networks initiates a movement which is taken up only later by the other network. In the case we are looking at, this "anticipation" might be made by the general network or by academic research. We will begin by giving an example in which the initiative comes from network F2.

In file F1, the theme *reinforced plastic* is constantly in quadrant 1, whereas in F2 it only appears between 1973 and 1978 as a non-isolated aggregate, over which period it remains in quadrant 3 (thus it is a developed, non-central theme). It then splits away from the academic network to become an isolated cluster that no longer plays a role in the dynamics of F2. This suggests the following hypothesis: despite the relatively weak centrality of this theme in academic research, it might be that it is its high development which is responsible for the central position of reinforced plastics research in the general network. Evidence to support this hypothesis can be found in analyzing the associated clusters.

As shown by Table 13, in file F2 (academic science) the study of reinforced plastics seems to be limited to the model of hardening by vulcanization or by adding charges, such as carbon black or glass balls.

Table 13
"Reinforced plastic" cluster changes over time (F2)



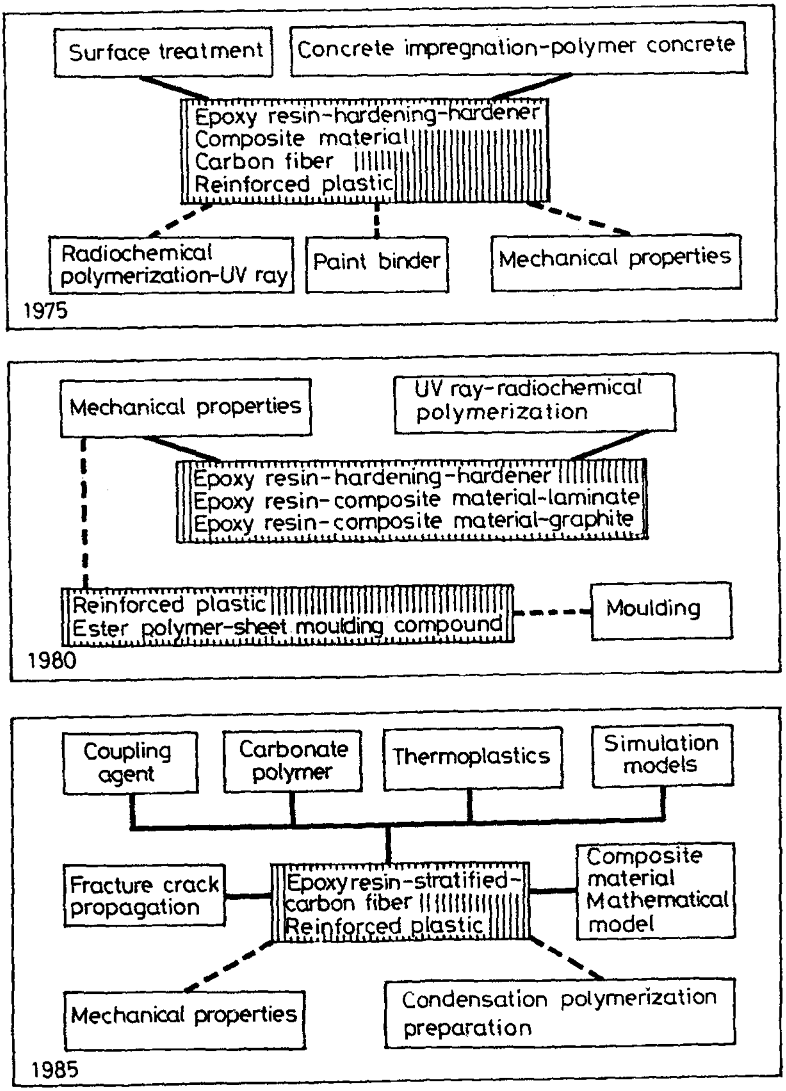
This is the strong contrast with file F1, where the theme of reinforced plastic develops nicely from 1978 to 1985. As we see in Table 14, there are two lines of evolution, which result in the splintering of the cluster in 1975 into two distinct clusters in 1980. In 1985, there is a fusion between these two lines of development.

Throughout this period, and despite the divergence in 1980, the theme *reinforced plastic* is closely tied to composite materials and to hardening attributable to epoxy resins. This cluster becomes more important over time. One is struck by the multiplicity of links that are formed and consolidated over time with such clusters as *preparation by polycondensation, radiochemical polymerization*, and also *mechanical property, surface treatment, fracture crack propagation, coupling agent, impregnation of concrete moulding-implementation* or again *mathematical modelling-composite materials.* In the course of time, the theme *reinforced plastic* creates itself a veritable ecological niche made up of techniques of polymerization, implementation modalities and modelling activity.

The contrast between the two files is striking. As regards F2, research is centered on a limited problem that remains rather marginal in the academic research network, despite signs (the constant improvement of the density measure) that interest for this theme is growing over time (quadrant 3). Concerning F1, new considerations constantly enrich this theme making it one of the more rapidly developing research contexts in the field over the period studied. Reinforced plastics constitute an applied rather than a basic research theme. It seems that academic science "unburdened" itself of this theme (towards the end of the 1970's) that became isolated within the field of basic research even if remaining the object of an important investment (quadrant 3). This fits with the fact that the theme achieved a stage of evolution allowing it to be taken up in a more applied network. This is what we learn from the strategical reinforcement of this cluster in F1, where it becomes the point of convergence and hybridization of multiple and heterogeneous concerns – ranging from modelling to applications by way of techniques for the development of reinforced plastics (Fig. 8).

Is academic research the driving force contributing to the structuration of the general research network? It is difficult to say given the short time period. However, this example shows that research problems which grow in strength in F2 can be "shipped out" of their immediate context and taken over by F1 for further development. In the case of reinforced plastics, academic research played a leading role (or at least actively participated) in the evolution of a research field which developed largely because of its contact with a context of application.

Table 14
"Reinforced plastic" cluster changes over time (F1)



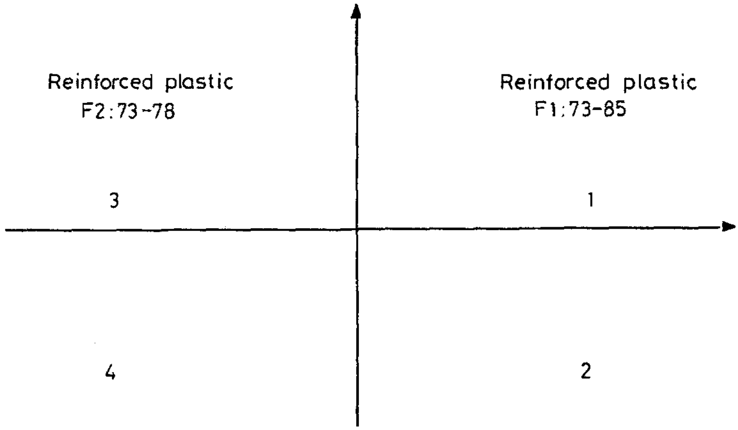


Fig. 8. "Reinforced plastic": strategic position in F1 and F2

(b) From applied research to the academic research

Let us take the case of the cluster "conductivity".

Table 15 gives a detailed characterization of this theme's trajectory within the network of academic science. In 1973-75 it is absent. In 1976-78 it is present but completely isolated. In 1979-80 it becomes crossroads and is linked to two secondary clusters (*fluorescence* and *polymer coordination*); it is located in quadrant 1. In 1981-83 it is once again isolated, and it becomes a principal cluster in 1984-86, located in quadrant 3. Up until 1981-83 there is not much change in its content. It deals with the study of the electrical properties of polymers taken as semi-conductors (doped polymers). In 1984-86 it is still concerned with semi-conducting polymers, but in a new way, probably corresponding to the adoption of a particular model – that of the doping of acetylene p. by arsenic V. fluoride. In parallel, a link is forged with NMR spectrometry in order to analyze these structures. Furthermore, in 1984-86 a major new concern appears at the very heart of the theme – that of the development of conducting polymers by electrolytic polymerization.

Table 16 sketches the evolution within F1 of clusters identified as belonging to the same family. A theme called electrolytic polymerization appears in this file in quadrant 1 in 1985. It thus occupies a central position. Its content gathers together ideas associated with the series of clusters on electrical properties of polymers in file S2 (most notably doping, electrolytic polymerization, iodine). This emergence as a strategical crossroads theme is the result of a long process of evolution which goes back to 1975. In 1975 there is an isolated cluster connected to the theme just discussed for 1985. It was not, however, totally isolated. The theme was linked in weak associations (less than three links) to several other clusters, most notably from "above" to the strategical theme *thermostability*. It is striking that at this date it included considerations of acetylene p. a substance only found in F2 in 1984-86 – that is to say almost ten years later. In 1980, the cluster was still isolated, but its content has been enriched by the inclusion of the triangle: acetylene p., doping, arsenic V. fluoride. A weak link appears associating this cluster with another, strategical, one called *mechanical property*.

Table 15
"Conductivity" cluster changes over time in file F2

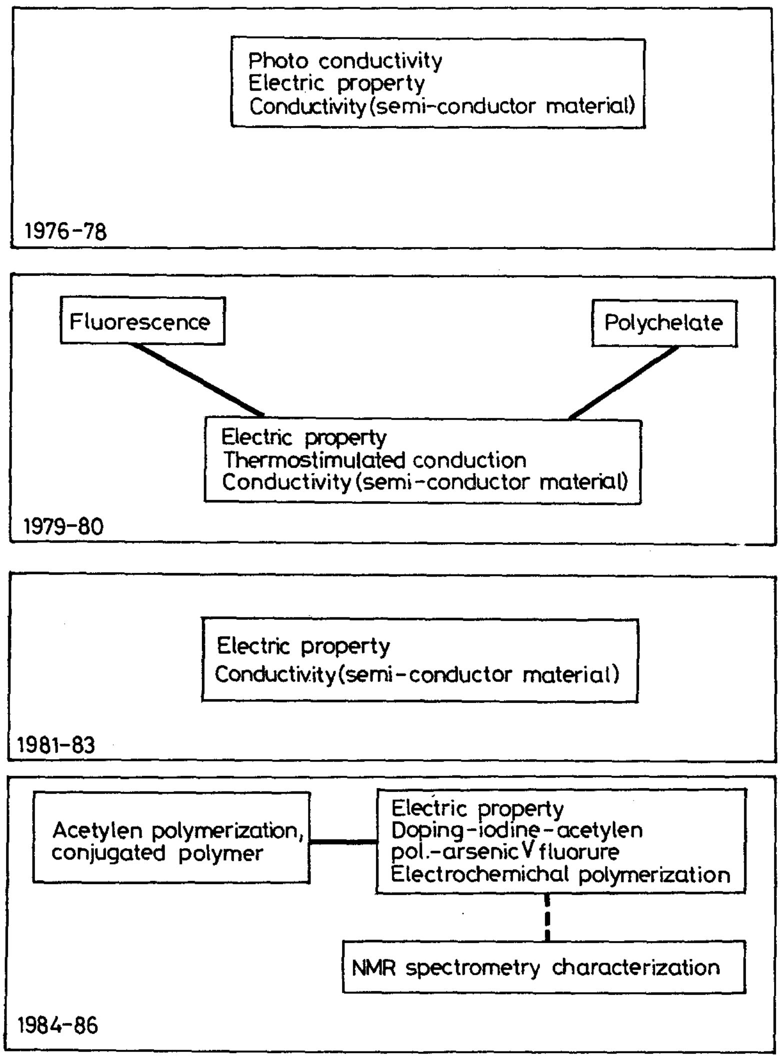
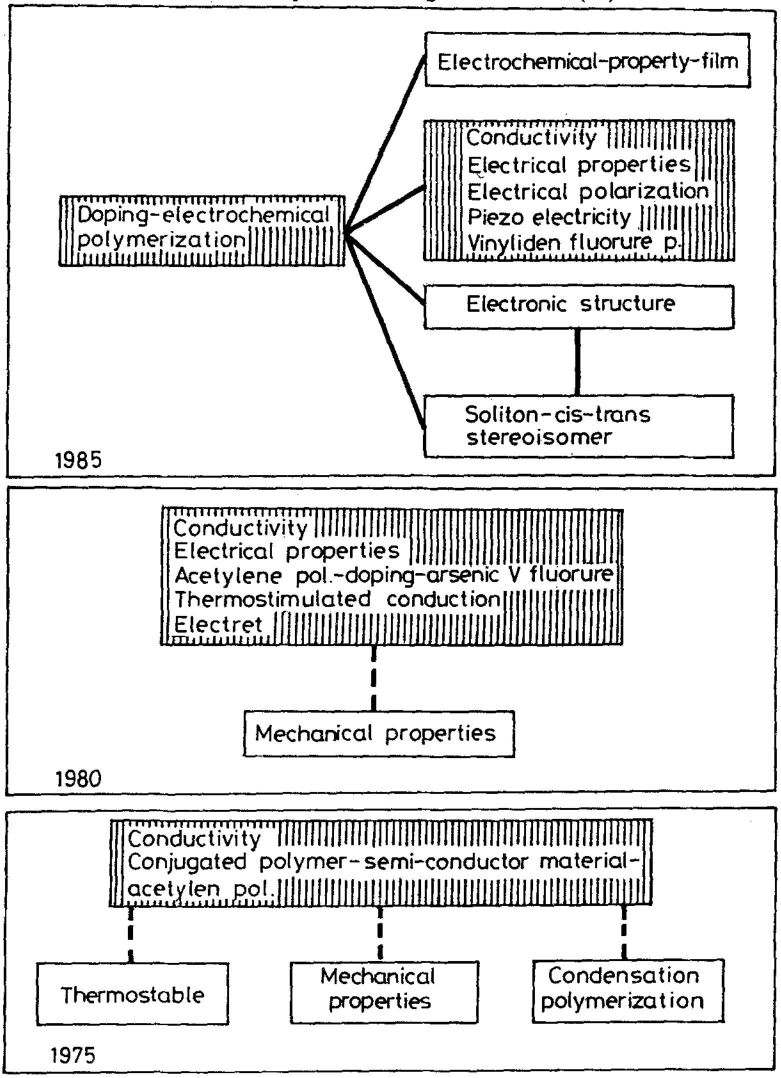


Table 16
"Conductivity" cluster changes over time – (F1)



It should be noted that in 1980 the content of this F1's theme is practically identical to that identified above in 1984-86 in F2. A dense and diversified sub-network emerged in 1985, one which included at least five themes. Apart from *conductivity*, and *doping* one finds the study of the *electrochemical properties of films*, the model *soliton-polaron* and the theme of the *electronic structure of polymers*. Here we see the whole context of research on conducting polymers being brought out. This research cannot be reduced, as was the case in F2, to the study of doped polymers alone. These interests are inseparable from the study of electrolytic polymerization and of films. The analysis of electrical properties seems to be inseparable from that of optical properties, heat-stimulated conduction and also mechanical properties.

A first level of interpretation, which we would need to confirm in a more detailed study taking in the articles themselves, can be made of the preceding observations. It seems that general research (F1) played an instigatory role in the formulation of a research problem that academic science took up as such at the end of the period. In F2, *conductivity* effectively appears as a recognized research problem over the entire period. It is stably defined until 1982-83. In 1979-80 the *conductivity* cluster is central in F2; in 81-83 it becomes isolated and cuts itself off from other academic research. It starts to look as if conductivity is a dead-end. After 1984-86, a new perspective emerges with work aimed at *doping acetylene p.* using *arsenic V. fluoride* which appeared in F1 in 1980. This new direction of research was no doubt a response to the growing weakness of the theme's strategic positionning in F2. By importing a subject that had previously been studied in an applied context, academic researchers succeeded in restoring the vitality of this line of the research. Although they developed its ties with other themes through an intense activity, the cluster remains peripherical in the academic research network. During the same period, the theme was also consolidated in F1. It became a highly developed sub-network linking a number of heterogeneous questions together: a wide variety of applications were being studied.

In this case, the impetus of the evolution that we have reconstructed for the period considered is found in the development of problems and models in the general research network. The academic research community latched onto them as a means of renewing its own inspiration and dynamic.

(c) Complex interactions between F1 and F2.

Tables 3 and 4 suggest that it might be interesting to look at *grafting.* In F1, it passes from quadrant 1 to quadrant 2, that is to say from a central and developed position to a still central but less developed position. Over the same period, it occupies a strategical position in F2 only to disappear in the end. A comparison of these contrasting evolutions suggests a progressive (and still uncertain) taking over of this theme by academic science. In order to confirm, rebut or qualify this assertion, we will look more closely at the trajectories of the clusters concerned within the two files.

Table 17
"Grafting" cluster changes over time – (F1)

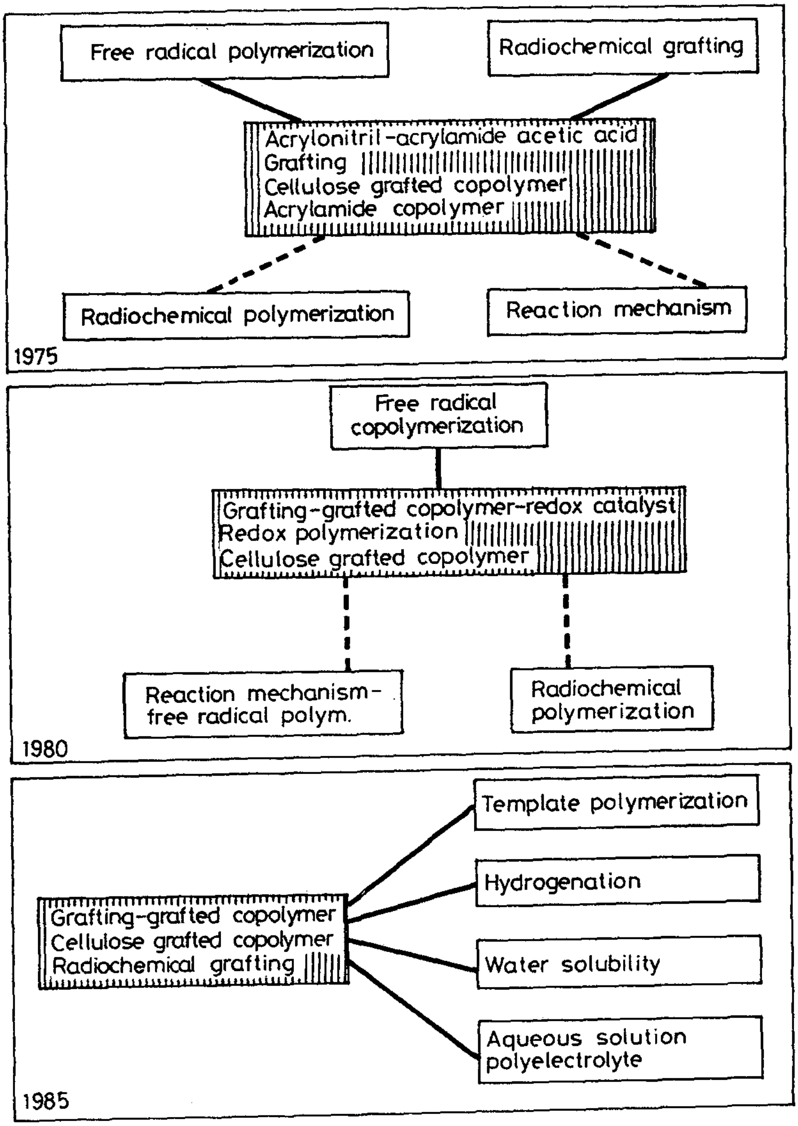
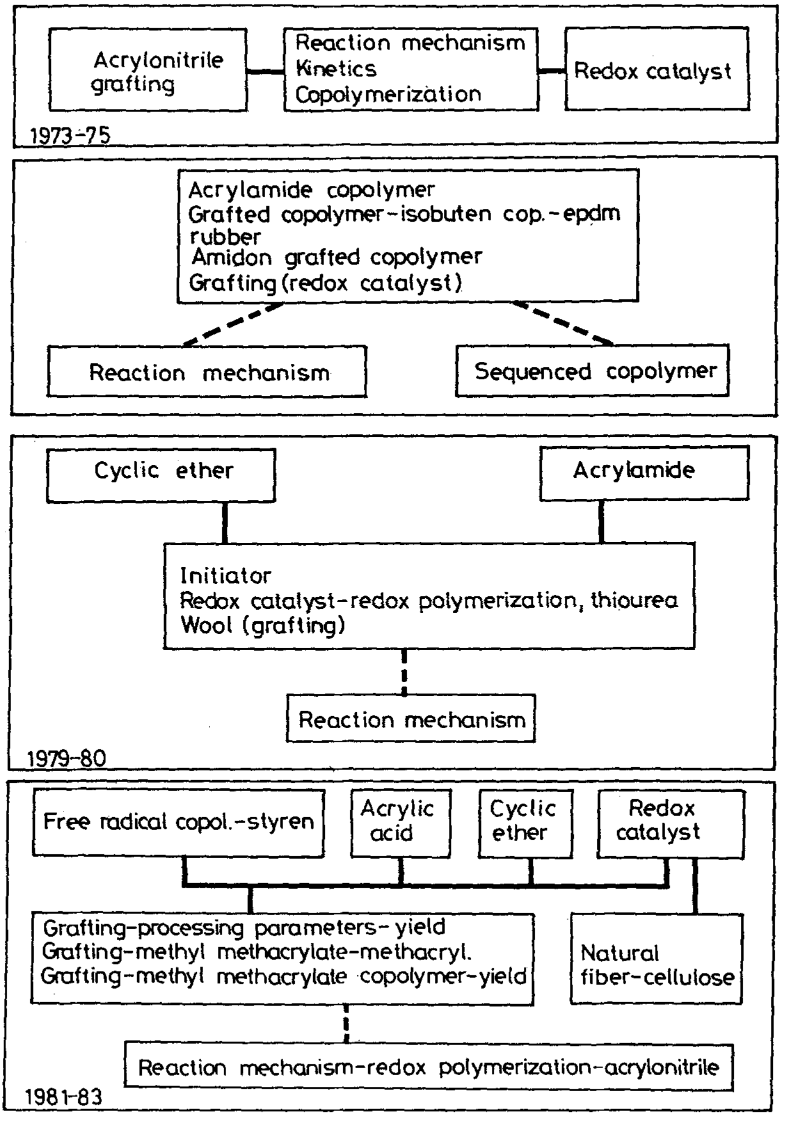


Table 18
"Grafting" cluster changes over time – (F2)



Let us start with the general file (F1) and Table 17. In 1973, the theme *grafting* was isolated. In 1975, a cluster organized around the notion of grafting appeared. It was associated with the crossroads cluster *radicular polymerization*, whilst the theme *radiochemical grafting* was connected with it from "below". In 1980, the theme was still there – it had become a secondary cluster of the crossroads cluster *radicular copolymerization* (which was also central). It should be noted that at this time the theme included treatments of the energizing of redox polymerization. In 1985 the cluster *grafting* remained present and its centrality is high. At this time, the association with redox polymerization seems to have disappeared. In 1985, on the other hand, the link with radiochemical grafting was consolidated, and the two themes fused.

There is comparable evolution of the theme of grafting in F2 (Table 18).

In 1973-75, the theme *acrylonitrite grafting* was secondary and linked to a crossroads cluster called *polymerization reaction mechanism*. Over the period 1976-78 integration with redox energizer occurred and the theme became isolated. In 1979-80, it became a principal cluster, with its content both reinforced and stabilized. In 1981-83 there was a fission. A new cluster called *grafting* became autonomous, as well as a cluster *redox energizer*, which is linked to it. Further, an independent cluster, *redox polymerization*, became autonomous. At the same time, the theme *grafting* became strategical in F2. It had many connections at this stage.

Overall, the following interpretation seems to impose itself. There was parallel evolution in F1 and F2, and this created room for reciprocal influences. The cluster *redox polymerization* appeared in F2 in 1976-78 and in 1980 in F1, then gave birth once more to two clusters in F2 in 1981-83. One could also stress, inversely, the practical concerns that appear in F2 at the end of the period and that provide evidence for the circulation of interests from general to academic research. There are many other possible examples. Grafting onto cellulose was studied in 1975 in the general network; and in 1979-80 researchers in the academic network took an interest in grafting onto wool and then more generally onto other natural fibres (cotton etc.).

(d) Some tools allowing a simpler analysis.

The examples that we have just given demonstrate the value of co-word analysis for the study of interactions between networks. It is clear, as we have shown, that the study of these interactions is rather complex, above all when one wants to go into any detail. If one is prepared to lose some information along the way, it is however possible to simplify the analysis by using diagrams that bring out the evolutions of the indices of centrality in the different networks theme by theme, series by series. Figure 9, which presents these graphs, shows that such diagrams can be used to draw conclusions that are not so different from those that our more detailed analysis led to.

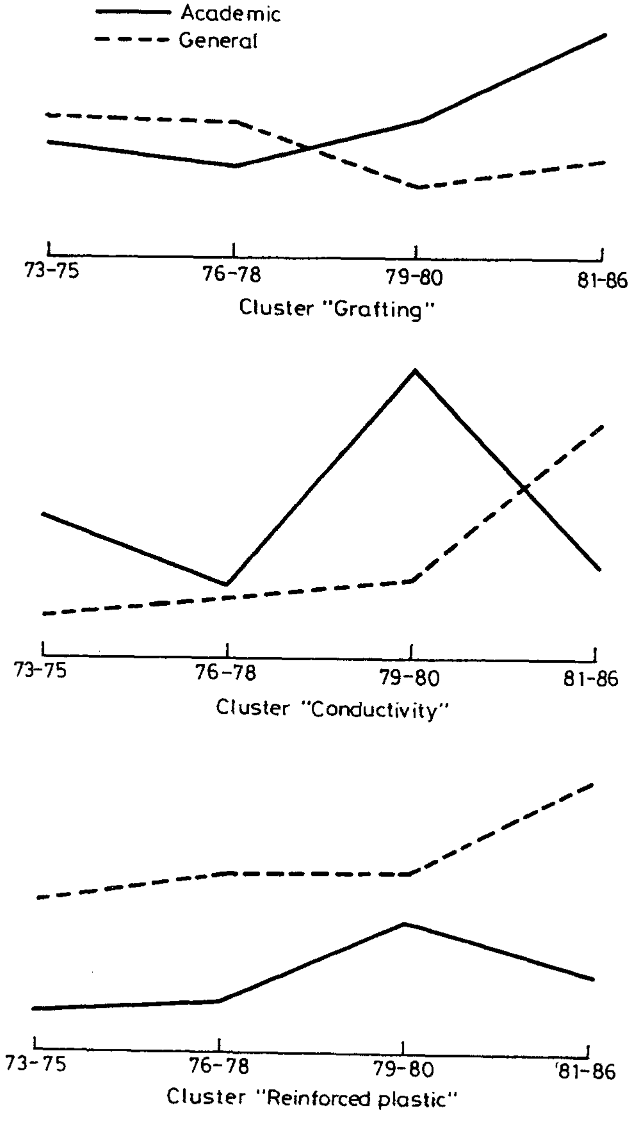


Fig. 9. Simplified representation of interactions between networks

For example, we can see right away that in F1 the centrality of the theme *reinforced plastic* increases regularly over time, with a strong increase between 1980 and 1985. At the same time, in F2, the index increases spectacularly in 1976 and 1980, and then falls away suddenly afterwards. This diagram suggests that the theme has become consistently more strategical in work on plastics, and that there has been a double transfer: first F1 → F2 and then F2 → F1.

The diagram covering the theme of grafting is equally clear. The graph representing the variation of the index of centrality in F1 and F2 shows a falling off and then regrowth of centrality. The theme id regenerated between 1976 and 1980. Looking at the curves for F1 and F2, we can see that this new life, no doubt initiated in F2, is amplified by a probable interaction between F1 and F2. This interaction is shown by the cluster's continued fall in the general network before it bounced back in 1980, after which time its development is comparable in the two networks.

The other diagram given in Fig. 9 presents some very varied types of interaction. The theme of *conductivity* seems to be a good case of academic research that either peters out or is quickly completed and leads to more applied research.

These various examples demonstrate that it is possible to follow the interactions between networks using just a few indicators that are easy to interpret. One could, of course, envisage more complete diagrams with index of density as well as centrality being given for each network.

6. Conclusions

In this article we have attempted to establish the usefulness of scientometric methods, and in particular co-word analysis techniques, for the study of relationships between academic and applied research. The method can be extended to include a patent analysis thereby making it an appropriate tool for empirically describing the network of interactions between science, technology and the market place. This is a central question to research underway on the economics of technical change. We can regroup our findings in the following way:

(1) Co-word analysis enabled us to identify the subject areas that characterized basic and applied research at different periods in time. Not only were we able to describe the content of the research underway, but two indicators were also developed and used in order to position each subject area in its corresponding research network (centrality) and to determine its degree of internal development (density). The links between themes were visualized in order to identify the axes of

research which remain constant over time, those which disappear and those which emerge. A coefficient was calculated and applied to measure changes in the stability of an axe's subject content over time. This transformation index is useful in describing research trajectories.

(2) Our analysis of these trajectories is quite different to the one that economists generally produce in connection with their discussion of technological trajectories. For us, a trajectory is defined as much by context as it is by content. Concretely, the analysis of a trajectory supposes that the three measures of centrality, density and content transformation be considered simultaneously. These three measures constitute powerful instruments for studying the dynamics of a research network. They enable us to characterize research themes given variations in their content (does this content change over time); their positions (are they obligatory passage points in the network); their degree of development (are they solidly constituted or not). They supply information on the theme and the network, two sides of the same coin which cannot be separated when trying to understand a trajectory. Two tools have been instrumental in synthesizing this information. The first is the strategic diagramme which gives a classification of themes in relation to their centrality and density and enables us to follow the changes in their network positions over time. The second retraces the life cycle of the themes by reproducing on the same schema the evolution of several variables: the coefficients of centrality, density and content transformation, but also the number of documents related to a theme which measures its relative importance. These tools serve to identify different trajectories given a particular combination of the above variables. An example of one type of life cycle is supplies by the trajectories of subject areas in which scientists are constantly obliged to redefine the content of their research programmes in order to maintain their position in a network.

(3) The dynamic of a network can be apprehended more simply. Two synthetic measures such as the degree of global integration or polycentrism supply a first approximation. As a means of following the evolution of a field, they enable us to go farther than a simple publication count (articles or patents). We showed for example that the academic research network is more open to outside influences than the applied research network and, because of its adaptability, its structure changes more rapidly than that of the applied research network.

(4) One of the goals of this article was to show that co-word analysis can be used to analyse the complex series of interactions which are typical of the network of innovation. Our demonstration was based on a discussion of relationships between

basic and applied research. While this is only a part of the total story, we have shown that co-word analysis allows us a rich variety of analysis which is not generally available when using other methods. Subject areas can be compared in terms of their content, their strategic character and their stage of development. Time lags between sectors of research can be detected and interpreted after examination of different interaction modalities. We found for example that certain themes are first developed in an academic context and then transferred towards more applied research. The content is transformed during this transfer as is the position of the themes in both research sub-networks: the progressive abandon of an academic research theme placed it on a trajectory which led to its marginalisation (quadrant 4) while, at the same time, it reinforced its position in the network of applied research to which it had been transferred. Conversely, other themes which were initially developed uniquely as an object of technological research came to be considered as a starting point for the renewal of a basic research programme in academic science. In another case, the evolution was simultaneous, with many cross-fertilizations. All these examples show that the reality gives rise to a large number of different types of interactions. They can be satisfactory understood with the help of the method.

(5) A last point should be underlined with respect to the method. Co-word analysis, and in this it is in line with the sociology of traduction which gave rise to it, does not rely on a priori definitions of research themes. The subject areas identified are those which are constructed by different actors (researchers, engineers,...) and which they define and transform in the course of their interactions. Co-word analysis considers the dynamic of interactions to develop as a result of actor strategies. If the content of a subject area changes over time, if its centrality and density vary, if it disappears from one network only to reappear in another, the explanation lies in the combined effect of a large number of individual strategies. Co-word analysis techniques should allow us in principle to identify the actors and explain the global dynamic: an analysis which was not carried out in this paper. However, the very possibility of being able to achieve this goal opens new perspectives particularly in connection (*Callon* et al, 1990; *Callon*, 1990) with the theoretical study of the techno-economic networks which establish the many fluctuating links between science and the economy.

References

BAUIN S., B. Michelet, *Comprendre la réaction chimique: étude des tendances internationales,* mimeo CNRS, 1988.

CALLON M., J.P. COURTIAL, F. LAVILLE, *Co-word Analysis as a Tool for Evaluating Public Programmes of Research,* CSI-NSF, 1990.

CALLON M., P. LAREDO, V. RABEHARISOA, Public technological programmes and the dynamics of techno-economic networks; submitted for publication (French version to appear in: D. FORAY (Ed.), *L'évaluation économique de la technologie.* Economica, 1990).

CALLON M., Some elements of a sociology of translation: Domestication of the scallops and fishermen of St. Brieuc Bay, in: J. LAW (Ed.), *Power, Action and Belief: A New Sociology of Knowledge?* Routledge and Kegan Paul, 1986.

CALLON M., Techno-economic networks and irreversibility; submitted for publication (French version to appear in: R. BOYER (Ed.), *Figures de l'irréversibilité en économie,* EHESS, 1990).

CALLON M., LAW J, RIP A. (Eds), *Mapping the Dynamics of Science and Technology*, McMillan, London, 1986.

COURTIAL J.P., Qualitative models, quantitative tools and network analysis, *Scientometrics,* 15 (1989) 527.

FREEMAN C., *The Economics of Industrial Innovation,* Frances Pinter 1982.

GRUPP H., B. SCHWITALLA, Technometrics, Bibliometrics, Econometrics and Patent Analysis: Towards a Correlated System of Science, Technology and Innovation Indicators, in: A. VAN RAAN et al. (Ed.), 1989.

MICHELET B., *L'analyse des mots associés,* unpublished dissertation, 1988.

NARIN F., E. NOMA, Is technology becoming science?, *Scientometrics,* 7 (1985) 369-381.

NARIN F., E. NOMA, R. PERRY, Patents as indicators of corporate technological strength, *Research Policy*, 16 (1987) 143.

NELSON R. R., S.G. WINTER, In search of a useful theory of innovation, *Research Policy,* 6 (1977) 36.

OEHLER K., W.E. SNIZEK, N.C. MULLINS, Words and sentences over time: How facts are built and sustained in a specialty area, *Science, Technology and Human Values,* 14 (1989) 258.

PAVITT K., Sectoral patterns of technological change: Towards a taxonomy and a theory, *Research Policy,* 13 (1984) 343.

TURNER W.A., CHARTRON G., LAVILLE F., MICHELET M., Packaging information for peer review: new co-word analysis techniques, in: A. VAN RAAN (Ed.) *Handbook of Quantitative Studies of Science and Technology,* Elsevier, North Holland, 1988.

VAN RAAN A., A. NEDERHOF, H. MOED (Eds), *Science and Technology Indicators,* DSWO Press, University of Leiden.

B.G. VAN VIANEN, H.F. MOED, A. VAN RAAN, The assessment of the science base of recent technology, *Research Policy,* 19 (1990) 61.

VON HIPPEL E., *The Sources of Innovation*, Oxford University Press, 1988.

WALSH V., Invention and innovation in the chemical industry: Demand pull or discovery push? *Research Policy,* 13 (1984) 211.