INFORMETRICS 87/88
L. Egghe and R. Rousseau (Editors)
Elsevier Science Publishers B.V., 1988

A METHOD FOR MAPPING BIBLIOMETRIC RELATIONS BASED ON FIELD-CLASSIFICATIONS AND CITATIONS OF ARTICLES

R.J.W. TIJSSEN*, J. DE LEEUW** & A.F.J. VAN RAAN*

*Science Studies Unit, LISBON-Institute, University of Leiden
Stationsplein 242, 2312 AR Leiden, The Netherlands

**Department of Data theory, University of Leiden
Middelstegracht 4, 2312 TW Leiden, The Netherlands

Abstract

We introduce quasi-correspondence analysis, a new method for mapping interrelations between publishing scientific entities, e.g. journals. It is based on the spatial configurations resulting from the analysis of the following bibliometric data: the relations in the citation-structure, which are examined for both the cited and citing-mode of scientific entities, and the distributions of the frequencies of different field-classification codes assigned to the documents published by each entity. The network of relations are visualized ('mapped') as separate configurations of points in a low-dimensional space. Each point represents a row or column of the corresponding data matrices. Distances between points represent the strenght of the relationship between scientific entities with respect to field-classifications of documents and citation-flows, respectively. Both configurations are then combined into a joint interpretation of the relations between the entities.

1. INTRODUCTION

1.1 Bibliometric studies of science

The explosive growth in basic and applied scientific research and in science-based technology over the last decades, and the even larger growth of the necessary expenditures, have given rise to the urge of scientists and S & T policy-makers to assess research performance or identify developments in scientific (sub)fields. It became necessary to develop and apply readily available 'objective' analytic tools, in addition to traditional expert-judgement, to assess policy relevant aspects of scientific research, in particular, output and impact of research activity. An emphasis was put on the analysis of quantitative data on scientific research, as embodied by manifest contributions to scientific knowledge, usually in the form of research documents, such as conference proceedings, patents and, mainly, publications in professional scientific journals. Contributions in refereed journals are considered to be the main communication-channel for presenting original scientific findings and dissemination of scientific knowledge. Quantitative information on scientific activity based on the coded contents of such scientific documents (i.e., field-classifications of the document) as well as the number of references to scientific documents are considered important bibliometric (i.e., literature-based) indicators of developments in scientific activity. In recent years, there is both an increased utilization of such quantitative descriptive indicators in science policy to assess publishing scientific entities, and a trend toward the assessment of the structure and dynamics of larger parts of science (the

so-called "mapping of science"). Scientific entities are, for instance: individual scientific documents (e.g., articles, patents, books), individual authors, or aggregates of these entities, such as: professional journals, academic departments, universities, industrial organizations, nations or (sub)disciplines of science.
In the present state of art in quantitative science studies, the increasing complexity of science as a structure of knowledge and the consequent need for more - and more detailed - information on the characteristics of science, places demands on bibliometric data analysis. The application of data-analysis methods has become a prerequisite for, in particular, more sophisticated science studies dealing with a large amount of (aggregated) quantitative data on a number of several aspects of scientific entities.
In this paper an explorative bibliometric method is discussed, based on a data-analysis technique, which maps two related structures of relations found within a set of (aggregates of) scientific entities. The method is based on information retrieved from scientific documents, such as journal articles. Two important quantitative features of documents are incorporated: (1) the number of citations to and from the documents, and (2) the frequencies in which coded classifications of the contents (e.g., keywords defining subject-fields) are assigned to the documents within the scientific entities. The frequencies which certain classification-codes are assigned to aggregates of scientific documents (like journals) presents us with a quantitative "contents-profile", and can thus serve as a bibliometric indicator of the contents. Field-classification codes are available in a number of data bases, each dealing with a particular scientific (sub)discipline[1].
The method is, in principle, suitable for various types of bibliometric data from any kind of publishing scientific entity. Given the amount of data available on scientific journals, it is particularly useful to assess structural relations between such scientific journals, because journal-to-journal relations have proved promising as a useful point-of-entry in various bibliometric applications such as the mapping of the structure of scientific (sub)disciplines (Small & Griffith, 1974), monitoring the dynamics of (sub) disciplines (Leydesdorff, 1986) or scientific library management (Dhawan, Phull & Jain, 1980). For these reasons we will discuss the method in terms of an application on bibliometric (inter)relations between journals.

1.2 Multivariate data-analysis methods

In our method we apply a multivariate data-analysis method (i.e., a method which can analyze the quantitative relations between a number of variables simultaneously) to both structures, and compare the analysis results. In the multivariate type of quantitative science studies, data are generally collected in a data matrix. Each row and column of the matrix represents a variable. The matrix elements are numericals, representing frequencies in which a row entity is found in combination with a column entity. In general, there are two major groups of possibilities to fill such scientometric matrices, depending on the form of the matrix (i.e., a square or rectangular matrix) and the form of the data it contains (i.e., symmetrical or asymmetrical data). For instance, one can opt for different row and column units (e.g., journals versus subfields) or similar sets of units (e.g., a group of articles versus the same group, as in co-citation analysis). Only the latter case creates a symmetrically filled matrix, i.e., the data structure located under the main diagonal of the matrix is similar to the upper-diagonal part of the matrix. Asymmetric data are in particular found when dealing with transaction data, such as citations. Each cell of such a transaction matrix contains values indicating the level of transaction between a row and a column entity, in general the observed number of citations. When one is interested in citation-interrelation- ships between sets of entities, the rows of a rectangular matrix classify the cited mode, and the columns the citing mode of an entity, or vice versa. If data sets contain the same (and identically ordered) entities, the elements in the main diagonal of the square matrix represent self-citations.

1.2.1 Correspondence analysis

There are a number of related multivariate data-analysis techniques available for analysis of different types of scientometric data matrices, discussed in the previous section. In this paper, we will discuss a data-analysis method known under many names, but mostly referred to as Dual Scaling (Nishisato,1980) or Correspondence analysis (Greenacre, 1984). Additional information on (particularly the geometrical properties of) Correspondence analysis (CA in the following) is given in Gifi (1981). More detailed information on the use of CA as a tool in citation analysis can also be found in Tijssen (1987).
In short, CA assigns scale values (scores) to each row and column and represents them simultaneously as a configuration of points in a multidimensional space with so-called chi-squared metric characteristics. Each dimension accounts for an independent part of the information (also referred to as 'variance') on the relations between matrix rows and columns. The first dimension accounts for most variance, the second dimension for the second largest amount of variance, etcetera. The distances between pairs of points in the common space is dependent on the values indicating the strength of their relationship (e.g., highly related points will be located in each other vicinity, while those which are apart are relatively unrelated). CA does not operate directly on the values of the observed data, but on deviations from expected values, calculated from an independence model; in fact, the expected numerical entries are the result of independent row and column contributions, and based on the row and column sums only. The independence model is only used as a baseline model. One is primarily interested in cross-structural relations between the participating row and column entities, whereas the marginal frequencies (the row and column sums) are of lesser or no importance. It is obvious that the expected values based on this model will generally not fit scientometric data. However, the differences between the observed data and the expected citation values based on the model will yield useful information on the cross-structural relations,because the 'size effects' of the rows and columns will then have been ruled out. The remaining structure of dependencies among the rows and columns, i.e. the values of the residuals after fitting the independence model to the matrix, can thus be analyzed to investigate the structure of the network of interrelations between the rows and columns.
When interpreting the relations between row scores (or column scores) in the chi-squared metric, one must bear in mind that row and column scores with profiles (i.e. the values in a row or column divided by their marginal value) which are (very) deviant from these mean-profiles are found in the periphery of the plot. The mean profiles, the profiles of the row and column sums, are always located in the origin of the space [denoted with the spatial co-ordinates (0,0,...,0)]. When interpreting distances between row and columns, one can roughly assume that, the distance between a row and a column is small if the observed values are (much) larger then the expected values. Conversely, if the observed values are (much) smaller then the expected values, the points are far apart, and will be presented as opposed points in a multidimensional space. The chi-squared distance of a point to the origin thus represents the extent of the difference between the row and column profiles and the mean profiles.
Note that the low-dimensional CA-configuration is primarily an approximation of the structure as a whole. One must therefore exercise care when interpreting relations between separate pairs of points; comparison with the original data is required.

1.2.2 Quasi-correspondence analysis

CA requires a completely filled data matrix, but in some cases one might encounter a data matrix which is incomplete, or in which the incorporation of certain matrix elements may seem inappropriate for an adequate analysis. In a citations-analysis context one can consider the relatively high values of the self-citations within journals one often has to deal with. First, one is often only interested in the inter-journal relations, whereas the self-citation frequency is typically intra-journal information. Secondly, when applying CA on data matrices

with high self-citation frequencies, the row and column profiles tend to become more similar as a result of the dominance of these diagonal elements over the off-diagonal elements. Consequently, relatively little differentiation is found between the row and column scores, thus obscuring the most interesting feature of CA: the representation of the asymmetry in the inter-journals citation relations. In order to discard these self-citations in CA one must thus eliminate these values in an appropriate way, without influencing the computation of the other row and column scores.
A solution to this problem can be found by introducing a so-called quasi-independence model; a generalisation of the independence model to incomplete matrices. Quasi-correspondence analysis (QCA) is an accordingly modified version of CA for incomplete matrices, yielding information on the relations between entities, of both the row and column mode, while ignoring (the values in) specified matrix elements. More details on QCA are given in De Leeuw & Van der Heijden (1987). An application of QCA on citation-transaction matrices is given in Tijssen et al. (1987).

2. AN APPLICATION TO JOURNAL CITATION AND JOURNAL FIELD-CLASSIFICATION DATA: MATERIALS SCIENCE

2.1 Introduction

Results will be presented on structural relations between materials science journals for two successive periods of time, thus monitoring recent developments in this subfield of physics. We chose to monitor the dynamics of this subfield, because, in general, developments in physics are relatively well documented through scientific publications in a large number of scientific journals, and because bibliometric information on these journals is present in several data bases. The subfield of physics described as materials science was chosen as an example of a subfield in which turbulent developments have taken place in the last decade, marked by the introduction of many new and useful materials and, only very recently, the explosion of research activities in search of superconductivity at increasingly higher temperatures.
Finally, it must be stressed that the primary aim of this investigation is not to assess the dynamics of the whole field of materials science, but to display developing relations between a set of journals which play an important role in this subfield.

2.2 Method

2.2.1 Collecting citation data

To assess developments in materials science by means of changes in citation-patterns between scientific journals, we constructed a so-called journal-to-journal citation matrix. We compared the relations between the cited and citing-"profiles" of the journals based on the citing years 1980 and 1985. Only citations to recently published work were taken into account, i.e., citations given in 1980 to publications in the period 1976-1980, and citations given in 1985 to publications in the period 1981-1985. This period of 5 years was chosen mainly to obtain a matrix with as much sufficiently filled matrixelements as possible, to enable to perform a reasonably valid correspondence analysis of a larger group of journals.[2] A period of 5 years was considered a minimum; a longer period would have been even better for this matter, but use of an extended period means less measurements based on longer periods of time, thus possibly loosing grip on the dynamics of a development.
The citation data were collected manually from the 1980 and 1985-Journal Citation Reports (JCR), compiled by the Institute for Scientific Information (ISI) containing aggregated citation data from the Science Citation Index.[3] Utilization of JCR-volumes for retrieval of this type of information has some limitations. First, in general not all journals active in a discipline are covered by ISI. In the

materials science case, this handicap particularly resulted in the exclusion of several Russian journals (or their English-translated versions) from our investigation. Secondly, the JCR-volumes only display information on the distribution of citations between journals if the total number of citations given, respectively, received in the full period of the journal existence, exceeds a certain lower bound. The height of the two thresholds depends on the journal and is on the average equal to about 5 citations. So, in our investigation this is not a serious problem; matrix elements with this type of "missing" data will always contain relatively low citation values, and thus have only a small influence on the total citation structure. Instead of treating these elements as missing, we placed estimates of the citation values for a 5-year period. These values were set equal to half of the value of lowest threshold (the lower bound is based on the total period of existence, in the case of most journals more than 10 years).

2.2.2 Collecting data on journal article classifications

Classifications of the contents of scientific documents dealing on subjects in the field of material science can be obtained from the Physics Abstracts' INSPEC-data base. This data base contains classification codes (based on the Physics Abstracts Classification Scheme-PACS) for the contents of documents published in a large number of professional physics journals. Each document is classified by experts (and in many journals the authors are asked to classify their articles themselves), resulting in the assignment of one or more PACS-codes to each publication, covering the essentials of the contents. An online search of the Physics Abstract database in the ESA hostcomputer (Frascati, Italy) was performed to obtain the lists of journals which contained most documents on material science (i.e., PACS-codes 8100-8190; in aggregated form referred to with the 2-digit PACS-code 81) during the periods 1976-1980 and 1981-1985. For each period, the ESA-database provided a frequency table of, in a descending order, journals containing publications to which PACS 81-code were assigned. These lists cover >95% of all journals with 81-coded documents.

The next step dealt with the question: which journals should one incorporate in the analysis to adequately cover the field of materials science? Assuming that the journals with most 81-codes are the important contributors in the field, we selected the ISI-covered journals which ranked highest on one or both of these lists.[4] This resulted in a total set of 18 journals: 9 USA journals, 6 West European journals, 2 Japanese journals and 1 Russian journal. These journals, representing about 4% of the journals with materials science-contributions covered by Physics Abstracts, cover approximately 30% and 40% of the total amount of 81-coded documents in the periods 1976-1980 and 1981-1985, respectively. An overview of these journals is given in table 1.

Subsequently, we collected quantitative data on the contents of these journals for each period via frequency distributions of the 4-digit PACS-codes. For each journal and period, the ESA-database provided frequency tables, of maximally 30 pages, with, in a descending order, the frequencies in which the different PACS-codes were assigned to the publications which appeared in the journal. These lists covered 95-100% of all codes assigned to each journal in each period. In table 2 we present the aggregated 2 digit-classification codes, relevant to materials science subjects, which were found to be most frequently present in these lists.[5]

For each period a so-called journal-to-field matrix can be constructed from the obtained PACS-code frequencies: assign the selected journals to the rows of the matrix and the selected PACS-codes to the columns. The elements of the matrix are filled with the percentages in which a specific 2-digit PACS-code is present in a journal, with respect to the total number of PACS-codes found for that journal. Each row can thus be seen as a field-classification, or "contents-profile" of a journal. The values in the columns yield an indication of the presence of a specific PACS-code over the selected set of journals.

The following background information on the PACS-code distribution per journal is given in table 3: an overview of (1) the proportion of the selected set of PACS-codes compared to the total number of assigned codes (codes 01-99); (2)

Table 1 - Scientific journals: Full title, title abbreviation, nation of origin

Acta Metallurgica (AM)-USA
Applied Physics Letters (APL)-USA
Fizika Metallov I Metallovedenie (FMM)-USSR
Journal of the American Ceramic Society (JACS)-USA
Journal of Applied Physics (JAP)-USA
Journal of Crystal Growth (JCG)-The Netherlands
Journal of the Electrochemical Society (JES)-USA
Journal of the Japanese Institute of Metals (JJIM)-Japan
Journal of Materials Science (JMS)-United Kingdom
Journal of Materials Science Letters (JMS-L)-United Kingdom
Journal of Vacuum Science & Technology A: Vacuum Surfaces and Films (JVST-A)-USA*
Journal of Vacuum Science & Technology B: Microelectronics Processing and Phenomena (JVST-B)-USA*
Japanese Journal of Applied Physics (JJAP)-Japan**
Materials Science and Engineering (MSE)-Switzerland
Metallurgical Transaction A: Physical Metallurgy and Materials Science (MT-A)-USA
Physica Status Solidi A - Applied Research (PSS-A)-Germany
Scripta Metallurgica (SM)-USA
Thin Solid Films (TSF)-Switzerland
*JVST was split-up in JVST-A and JVST-B in 1982 **In 1982, JJAP was devided into part 1, consisting mainly of normal articles, and part 2, which contains letters. Only part 1 is used in the analyses, under the heading JJAP, because not enough citation-information was found for JJAP-part 2.

Table 2 - The set of PACS-codes (cf. classification scheme Physics Abstracts-1983)

05	Statistical physics and thermodynamics
44	Heat flow, thermal and thermodynamical processes
46	Mechanics, elasticity, rheology
47	Fluid dynamics
61	Structure of liquids and solids; crystallography
62	Mechanical and acoustic properties of condensed matter
63	Lattice dynamics and crystal statistics
64	Equations of state, phase equilibria, and phase transitions
65	Thermal properties of condensed matter
66	Transport properties of condensed matter (nonelectronic)
68	Surfaces and interfaces; thin films and whiskers
71	Electron states
72	Electronic transport in condensed matter
73	Electronic structure and electrical properties of surfaces, interfaces and thin films
74	Superconductivity
75	Magnetic properties and materials
76	Magnetic resonances and relaxation in condensed matter; Mössbauer effect
77	Dielectric properties and materials
78	Optical properties and condensed matter spectroscopy and other interactions of matter with particles and radiation
79	Electron and ion emission by liquids and solids; impact phenomena
81	Materials Science
82	Physical chemistry
86	Energy research and environmental science

the proportion of the materials science PACS-code (81) in the total of codes, and, (3) another most frequently found PACS-code. The corresponding rankings are based on their percentwise presence in the selected set of codes for each journal.

A first glance at the table shows that most journals are relatively well covered by the selected set of PACS-codes; the broadly-oriented APL has the smallest percentage (45.6%), whereas SM is almost completely covered (99.2%). There are large differences with respect to the percentage of the specific materials science 81-codes per journal, ranging from only 6% (APL and JAP) to as much as 50.6% (MT-A). Futhermore, there are only a few substantial changes between periods within each journal, the most striking one is the change of the percentage of 81-codes in JACS.

2.3 Results of the analysis

2.3.1 Introduction

In the discussion of the results from the QCA of the journal-to-journal citation matrices, and the CA of journal-to-field matrices, we restrict our presentation to the analysis results in the first and second dimension. The two-dimensional analyses account for about 65-70% of the information in the data, of which approximately 50% is accounted for in the first dimension. Thus, this flat-space graphical display of the relations amongst the rows and columns of the matrices provides an approximation of the actual structures which is good enough for our explorative aim. For the sake of completeness, we mention that the third and forth dimension account for only 7-9% and 4-7% in the analyses, respectively.

The CA's were carried out with the computer program ANACOR (Gifi, 1985). The QCA's were computed with a newly developed program written in APL. All self-citations were eliminated in the QCA's of the citation matrices.

The scores assigned to the rows and columns by CA and QCA have a sign which is arbitrarly positive or negative, i.e. the results are invariant under simultaneous transformation of the signs of all scores. In some cases the combined geometrical

Table 3 - Overview PACS-code distributions per journal, for each period

	1976-1980						1981-1985					
	1	2		3			1	2		3		
	total	%-81	rank	code	%	rank	total	%-81	rank	code	%	rank
AM	98.8	38.8	1	62	19.1	2	97.6	38.4	1	62	20.7	2
APL	45.6	6.0	2	61	6.8	1	51.8	5.8	3	61	7.3	1
FMM	97.9	29.7	1	75	19.9	2	98.7	26.0	1	61	13.7	2
JACS	64.4	21.1	1	82	8.6	2	90.8	41.6	1	62	12.6	2
JAP	72.8	6.0	4	75	20.1	1	70.9	4.5	5	75	17.7	1
JCG	81.4	29.3	1	61	17.4	2	81.7	25.4	1	61	17.3	2
JES	66.2	17.0	1	82	8.7	2	68.9	13.9	1	82	7.7	2
JJIM	96.8	45.8	1	62	15.7	2	98.2	46.5	1	62	16.1	2
JMS	95.5	46.5	1	62	20.5	2	96.4	42.1	1	62	19.0	2
JMS-L							93.9	34.3	1	62	11.3	2
JVST-A	68.7	10.8	3	68	15.2	1	70.7	9.8	3	68	16.4	1
JVST-B							59.8	9.3	3	68	11.3	1
JJAP	69.1	9.7	2	61	10.4	1	55.0	5.5	3	72	8.0	1
MSE	97.4	42.1	1	62	24.5	2	97.5	45.9	1	62	23.2	1
MT-A	96.3	50.6	1	62	22.8	2	98.8	49.8	1	62	21.6	1
PSS-A	89.8	7.0	4	61	18.1	1	88.6	5.5	5	61	19.8	1
SM	99.2	35.9	1	62	24.0	2	99.4	41.3	1	62	20.3	1
TSF	82.5	17.6	1	68	16.4	2	79.4	16.0	2	62	16.1	1

interpretation of a more than one CA/QCA - configuration is thus simplified if an a posteriori transformation is carried out on one or more analysis results to obtain a similar sign for the scores. Such a transformation has also been carried out in the following applications.
For reasons of brevity, and lack of detailed knowledge of the materials science-field, we only discuss elementary features of the analysis findings.

2.3.2 Journal-to-field matrix 1980

The CA-solution with two dimensions accounts for 68% of the variance, with 46% in the first dimension. The configuration of the scores for journals and PACS-codes is given in figure 1. When interpreting the first dimension of the 1980-analysis solution (i.e., projecting the points on the first - horizontal axis - and comparing their positions), it immediately becomes clear that the set of journals can roughly be divided in two groups: at the right side of the figure applied physics journals with a broad scope, amongst them US journals such as JAP and APL, the German PSS-A and the Japanese journal JJAP, in addition to some specific journals such as JES, TSF and JVST; to the left of the figure, we find a group of journals more devoted to research on materials and metallurgy in particular, such as the US journals AM, MT-A, SM and the Russian FMM. In this group we find journals like the US journal JACS, and the European journals MSE and JMS. An intermediate position is taken by JCG, a journal focusing on crystals.

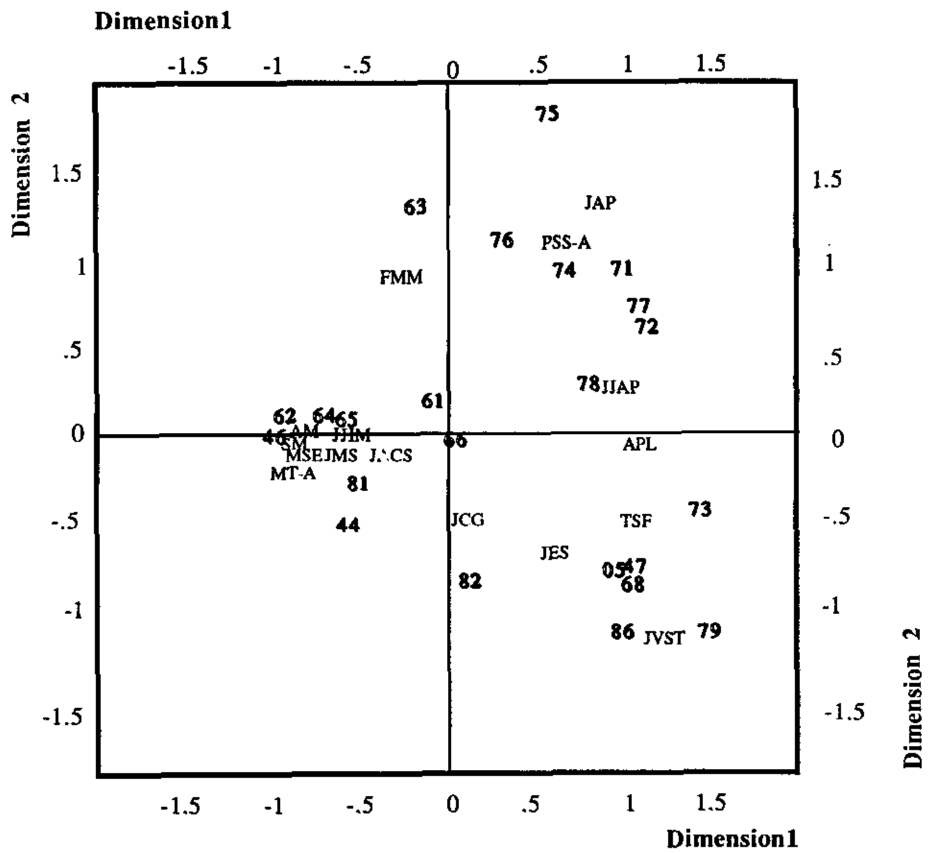

Figure 1: Correspondence analysis journal-to-field matrix 1980. Configuration of dimension 1 - dimension 2. Thin label: journal; shadowed label: PACS-code.

The second dimension, responsable for the (relatively) largest part of additional information, mainly accounts for a differentiation between the journals JES, TSF and JVST, focusing on instrumental aspects of experimental physics, versus JAP, PSS-A and FMM which are also more experimental oriented journals, but with an emphasis on models and theory. In the latter group FMM is the sole representative from the 'metallurgy group'.
What does this differentiation between journals mean in terms of their contents? For an answer we investigate the location of the PACS-codes. Starting at the left, we see that the journals in the metallurgy group have contents classified with the PACS-codes 44, 46, 62, 64, 65 & 81, which can be interpreted as dealing with a more general class of materials science-subjects. Two journals within this group are of special interest in the sequel: JACS is a journal dealing with the more traditional materials science (e.g., pottery), and JMS which is a broadly-oriented journal in the materials science.
There is a gradual transition to 'intermediate' codes presenting more common aspects of the contents of the selected journals (thus located in the centre of the figure-cf. section 1.2.1) such as 82, but also codes dealing with crystals: 61, 63 & 66, which are particlarly connected to JCG.
The contents of the other group of journals concentrates on subjects classified by 05, 47, 68, 71-76, 78, 79 & 86, which we roughly grouped under the heading - electronic and surface properties of materials.
The configuration in the second dimension is characterized by the differentiation of the latter group, in a group focusing on magnetic and associated properties of materials (i.e., 63, 75 & 76) versus a group of codes emphasizing surface aspects of materials (i.e., 05, 47, 68 and 79).

2.3.3 Journal-to-field matrix 1985

The two-dimensional CA-results of the 1985 journal-to-field matrix displays a structure (figure 2) which is largely the same as the 1980-configuration, but there are also some notable differences. In the first dimension (accounted variance 49%) we still find the macro-level differentiation between the 'metallurgy group' and the 'electronics group', but at the micro-level, seemingly small - but important - changes are to be found in the positions of journals in the coherent metallurgy group: first, JJIM has moved more to the left, toward 46 & 62 (inspection of the data reveals that 62 is mainly responsible, 62 in JJIM: '76-'80 - 15%; '81-'85 - 16%). Secondly, 81 has taken a position in the centre of the metallurgy group. The responsible journals are mainly: JACS ('76-'80 - 36%; '81-'85 - 41%) and SM ('76-'80 - 36%; '81-'85 - 41%).
The position of PACS-code 44 has changed considerably; it nows takes a position in the upper-centre of the figure, largely a results of the fact that JACS has less 44-codes (0.4% and 0.1%, respectively), while more 44-codes were assigned to JJAP (0.1%-0.3%) and JAP (0.0%-0.1%).
Further examination of the one-dimensional configuration reveals a notable shift to the right of 63 from the upper-centre near FMM, to a position amongst the 70 series of PACS-codes (reason: JAP, JJAP, APL and JVST-A have more 63 codes). In this part of the figure we see a moderately small move to the left of the codes 72, 77 (PSS-A is mainly resonsible) and 78 now between JAP (4.6%-5.2%) and APL (4.6%-5.2%).
In the lower-right part of the configuration there are three notable changes: 82 jumps from the lower-centre to a position in between the journals located in the lower-right section, this is particularly due to TSF (1.5%-2.0%), JGC (1.2%-0.8%) and JVST-A (5%-7.2%). Code 86 also moves to a position in the centre of this section of the configuration, coming from a more peripheral position. The responsible journal is JES (3.3%-6.1%). Finally, there is a counter movement of 47, away from this group of codes towards JCG (0.1%-0.8%), also indicating that these subjects are more (proportionally) equally spread over the journals for the periode '81-'85 (9 journals with 47-codes, whereas only 5 journals were found in '76-'80).
Going from the first to the second period, new journals have appeared in the arena of materials science. In this investigation we see the newcomer JMS-L,

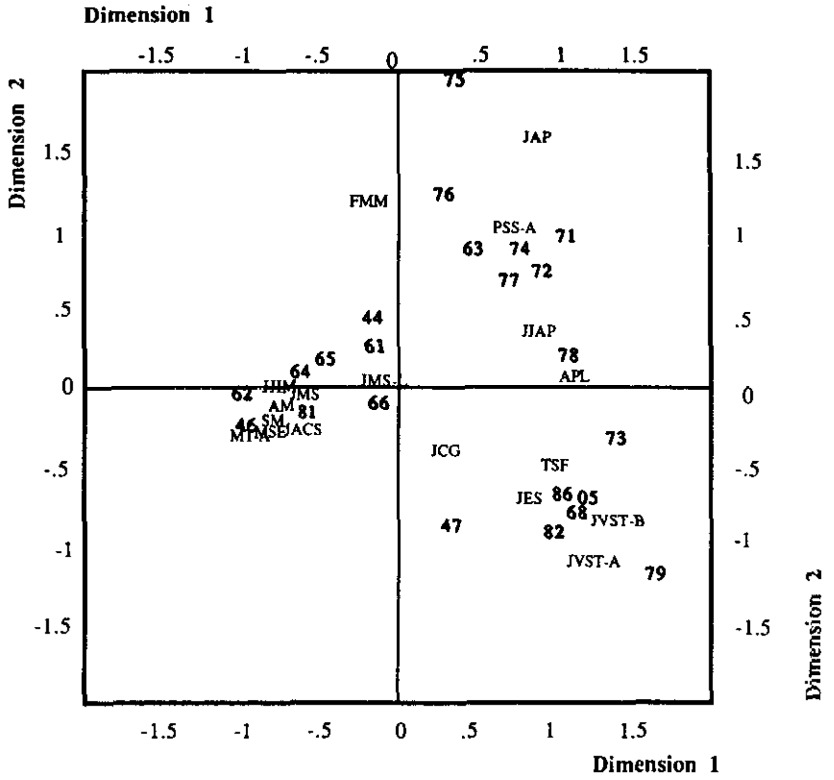

Figure 2: Correspondence analysis journal-to-field matrix 1985. Configuration of dimension 1 - dimension 2. Thin label: journal; shadowed label: PACS-code.

which has not taken a position in the vicinity of its normal articles-counterpart, but is located in the centre of the configuration (comparison of the contents-profiles of JMS and JMS-L shows that JMS-L contains more 71, 72, 73 & 82 coded letter-articles, and less 62 & 81 coded letter-articles). The separation of JVST has resulted in a clear split-up of JVST-A and JVST-B in the configuration, wherein JVST-B has taken a position away from 79 (JVST-B 4.0%; JVST-A 10.6%) and more toward 05 (JVST-B 7.1%; JVST-A 2.3%) and 73 (JVST-B 10.6%; JVST-A 6.4%)

2.3.4 Journal-to-journal citation matrix 1980

The two-dimensional QCA-solution of the 1980 citation structure yields a first dimension which accounts for 51% of the variance and a second dimension accounting for only 14%. The configuration of the scores for the journal modes is shown in figure 3.

When interpreting the first dimension of the 1980-analysis solution, it immediately becomes clear that we find a spectrum of journals which is highly comparable with the position of the journals in the journal-to-field configuration: to the centre-right of the figure the journals covering a broad scope of physics topics, like JAP, APL and JJAP; to the left of the figure, we find the same group of journals emphasizing on metallurgy, such as AM, MT-A and FMM. However, the intermediate position is now taken by PSS-A, which is located near the centre of the figure (i.e., this journal cites most other journals, and is also cited by them). The second dimension finds a differentiation between journals in the metallurgy group in particular cited JMS in combination with citing JACS, versus the other journal-modes. Apparently citations from JACS are directed mostly toward broad-scoped JMS, which contains publications on various aspects of materials science, including ceramics-subjects.

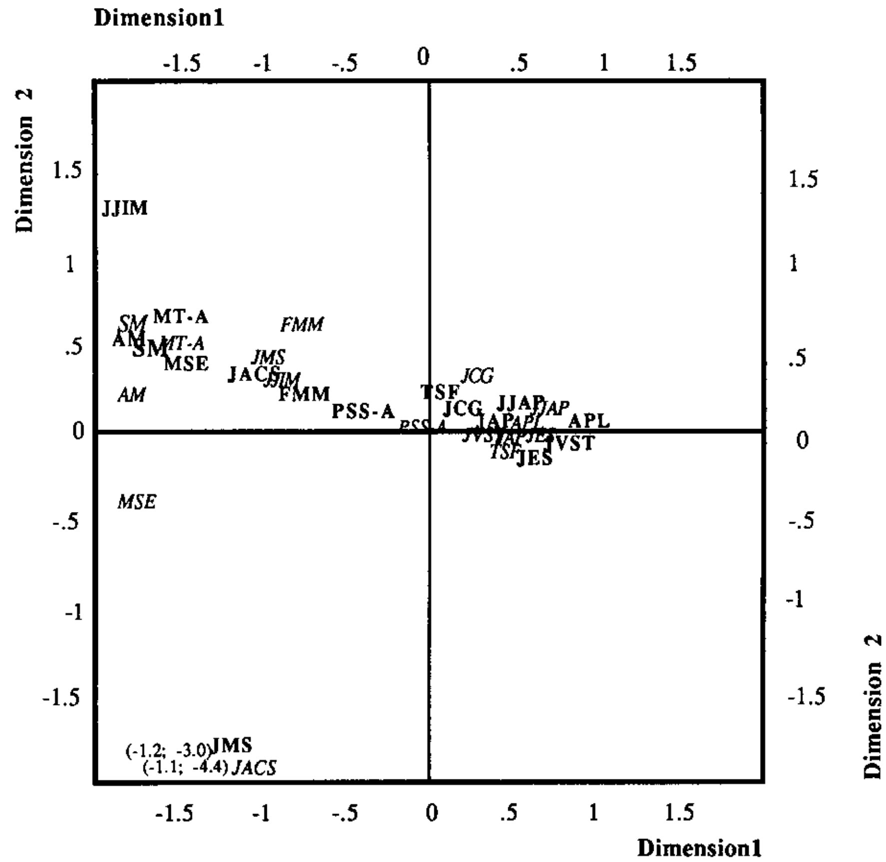

Figure 3: Quasi-correspondence analysis journal-to-journal citation matrix 1980. Configuration of dimension 1 - dimension 2. Italized label: citing journal; bold label: cited journal.

2.3.5 Journal-to-journal citation matrix 1985

The QCA-results of the citation matrix based on the 1985-articles of our set of journals yields a first dimension (accounted variance - 47%) which is highly comparable to the situation in 1980. This is completely in accordance with the journal-to-field result, from both periods, which also show a stability over time. Although only accounting for 16% of the information in the citation matrix, the configuration in second dimension reveals some interesting changes in the citation structure compared to the 1980-findings: a larger separation has taken place between JMS, its letter companion JMS-L and JACS versus the other metallurgy-oriented journals. It seems that the interrelations between these three journals has become increasingly tight, compared to the interrelations with the other metallurgy journals. Apparently, although the contents of JMS and JMS-L are slightly different (see section 2.3.3), they do have relatively large citation

flows between them. Furthermore, the position of cited JJIM and cited MT-A have changed within the 'metallurgy group'; at first they were located in the midst of the metallurgy-journals, but now take positions more towards the group with JACS.
Moreover, cited PSS-A and cited TSF have also moved closer to the JACS/JMS/JMS-L group. Inspection of the citation data indicate that these shifts in position are in fact solely due to their citation relation with JMS-L. In this respect, the central position of JMS-L between these four jurnals in the journal-to-field configuration (figure 2), seems appropriate.

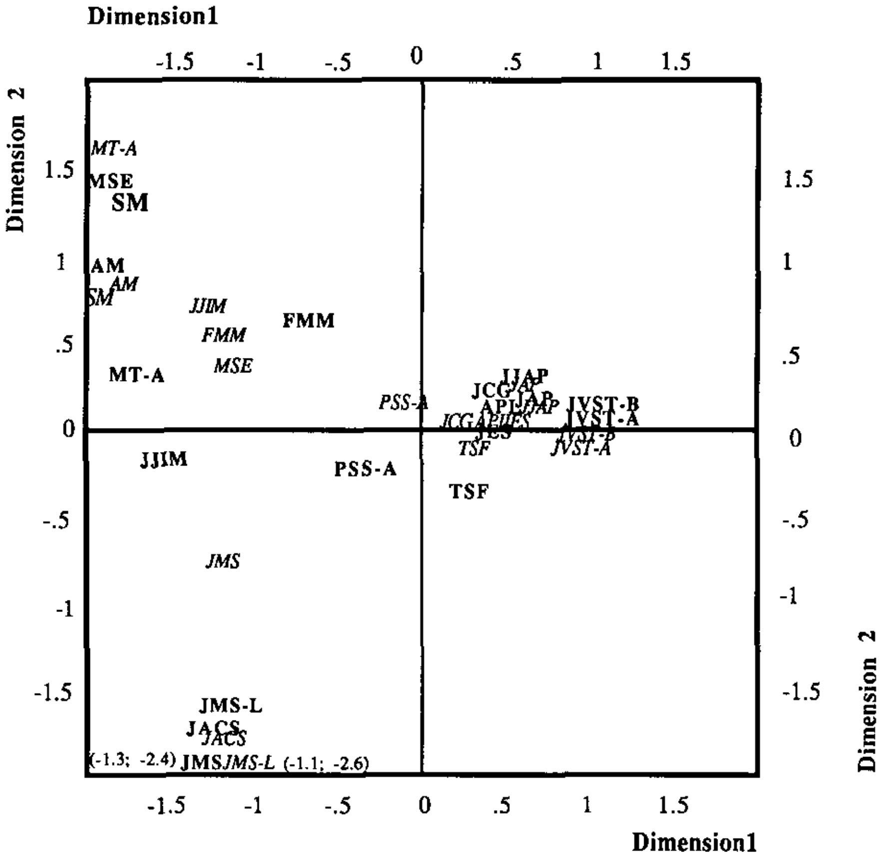

Figure 4: Quasi-correspondence analysis journal-to-journal citation matrix 1985. Configuration of dimension 1 - dimension 2. Italized label: citing journal; bold label: cited journal.

2.3.6 Final Remarks

The analysis findings of the journal contents has shown a more stable configuration, compared to the results from the journal-to-journal citation data; this is a quite plausible outcome, considering the fact that citation data are based on the subset of journal articles which actually received citations, with a relatively large variability in citation rates within and between periods. Comparing the analysis for both periods raises the question why different groups are differentiated; in the journal-to-journal citation configuration the journals in the metallurgy group are split up, whereas the journal-to-field configuration only divides the journals in the 'electronics and surface'-group. The probable answer lies within the relation structures for both groups: in the metallurgy group the journals apparently have more comparable field-classification "profiles", but at the same time they also have sufficiently large differentation in citation patterns to manifest itself in the second dimension. Of course, this does not mean that no differentiation in citation patterns exists within the other group; however, those citation patterns are apparently more diffuse, and can thus account for less information in the citation structure as a whole. As a result, CA will not 'recognize' those citation patterns as very significant.

This relatively larger diffusion in the other group can be explained by the finding that this journal group shows a relatively strong heterogeity with respect to their contents, a differentiation which manifests itself in the second dimension of the journal-to-field configuration. One may thus assume that this variety of contents (yielding 'an informative second dimension' in the journal-to-field analysis) is the cause of the more diffuse citation patterns in this group of journal (yielding 'a non-informative second dimension' in the journal-to-journal analysis).

3. CONCLUSIONS

In the foregoing we have succesfully combined analyses of journal interrelations, depending on citation flows and field-classification of the corresponding publications. The findings display a similar overall pattern for two periods of time. Moreover, we found that notable changes in the citation structure are related with shifts in the contents of journals. This application has thus shown that a combined structural analysis of bibliometric relations, based on citations as well as field-classifications, as dependent bibliometric indicators of scientific development, enables a fruitful explorative interpretation of dynamical aspects of science. The resulting journal configurations yield an excellent insight in the spatial position of a journal within a set of related journals. In general, the analysis method seems capable of detecting relatively small changes in specific relations, imbedded in the dynamics of larger system of relations between journals. Such macro-level analysis findings can also serve as a starting-point for a more fine-grained analysis of specific relations, e.g. on only a few scientific entities or contents-codes, or at a lower aggregation level of contents and/or entities. Finally, the method can also serve as a graphical tool for representing the position of new journals (e.g., JMS-L), or the relocation of the parts after a split-up of an existing journal (e.g., JVST-A and JVST-B).

In sum, the methodology presented in this paper seems a promising one. However, these findings must be seen as a preliminary assessment of the potentials of this method as a device to produce specific "maps of science". Further research is needed with respect to the possibilities to incorporate different, or more, types of bibliometric data and, eventually, to establish its validity as an analytic tool in, for example, science studies or evaluative research of journal standing by scientific publishers.

NOTES

1 Abstracting services such as Physics Abstracts, Chemical Abstracts and Biological Abstracts offer the text of abstracts of scientific publications and coded information on the contents of a publication.

2 Results from a correspondence analysis of a sparsely filled matrix (e.g., the citation data incorporating only 1 or 2 years) can be very unreliable; due to the creation of artifacts, a higher chance exists of emphasizing on specific relations, which are based on a very small number of citations.

3 The Science Citation Index (SCI), compiled by the Institute for Scientific (ISI), is a computerized bibliographic data base providing an annual source for

retrieval of bibliographic information in science, covering about 500.000 individual publications with roughly 8.000.000 references, appearing in over 3000 core professional journals (Garfield, 1979). In addition, aggregated citation information on the professional journals level is collected in the annually updated Journal Citation Reports (JCR). These reports include lists of all journals in the databases citing itself and other journals and, secondly, cited by itself and by other given journals.

4 Of course, this assumption is only partly true: possibly important, specialized - but small - journals will be left out when applying this selection criterion. However, such journals will generally have (very) small citation flows (if they are covered by ISI at all) and are thus mostly useless in explorative investigation of more large-scale features of citation structures.

5 With respect to this relevancy we consulted experts in the field of materials science.

REFERENCES

Callon, M., Courtial, J.-P., Turner, W.A. and Brain, S., From Translations to Problematic Networks: An introduction to Co-Word Analysis, Social Science Information, 22 (1983) pp. 191-235.

Dhawan, S., Phull, S.K., Jain, S.P., Selection of Scientific Journals: A Model, Journal of Documentation, 36 (1980) pp.24-41.

Gifi, A., ANACOR, Research Report UG-85-04, (Department of Data Theory, University of Leiden, 1985).

Gifi, A., Nonlinear Multivariate Analysis, (Department of Data Theory, University of Leiden, 1981).

Garfield, E., Citation indexing - Its theory and Application in Science, Technology and Humanities, (J. Wiley & Sons, New York, 1979).

Greenacre, M.J., Theory and Applications of Correspondence Analysis, (Academic Press, New York, 1984);

De Leeuw, J., Van Der Heijden, P.G.M., Correspondence analysis of incomplete tables, Psychometrika (in print).

Leydesdorff, L., The development of Frames of References, Scientometrics 9 (1986) pp.103-125.

Nishisato, S., Analysis of categorical data: dual scaling and its applications, (Toronto : University of Toronto Press, 1980).

Small, H., Griffith, B., The structure of Scientific Literature I: Identifying and Graphing Specialities, Science Studies, 4 (1974) pp.17-40.

Tijssen, R.J.W., De Leeuw, J., Van Raan A.F.J., Quasi-correspondence analysis on square scientometric transaction matrices, Scientometrics, 11 (1987) pp.347-361.

Tijssen, R.J.W., Scientometrics data-analysis methods: an introduction to correspondence analysis, (Research Report 87-1, LISBON-Science Studies, University of Leiden, 1987).