INFORMATION EPIDEMICS AND THE GROWTH OF PHYSICS

Albert N. Tabah

Graduate School of Library and Information Studies
McGill University
Montreal, Canada.

March 1996

A thesis submitted to the Faculty of
Graduate Studies and Research
in partial fulfilment of the requirements of the degree of
doctor of philosophy

0-612-56723-0

ABSTRACT

This study examined the prevalence of information epidemics in the physics literature. The primary interest was to find out whether outliers observed on time series charts of literatures are due to information epidemics, whether these epidemics are widespread occurrences in physics, whether literatures showing such rapid growth arise mainly due to the influence of an important work and, if so, what characterizes these literatures. Information epidemics were defined as spurts of growth in the literature of a field that reflect a sudden excitement and increase in activity. It was hypothesized that information epidemics are common occurrences in the growth of the physics literature and that outliers observed during the growth of a field are caused by influential works that attract new workers into it and cause them to publish extensively. Growth spurts where information epidemics lead to a permanent change and the emergence of a new subspecialty are termed knowledge epidemics.

The monthly number of abstracts indexed by each chapter of *Physics Abstracts* between 1977 and 1987 was plotted on a time series chart and an envelope of ± 3 standard deviations was fitted around the regression line. All spikes that crossed the envelope were considered to be outliers and thus potential information epidemics. The abstracts contained in each outlier were identified in the *Science Citation Index* and analyzed for spread (corporate sources of authors) and impact (citations).

Results show that information epidemics exist, but they are not widespread. Only four information epidemics were identified in the data. They are in chapters 2 (mathematical methods), 36 (clusters), 73 (heterostructures) and 74 (superconductivity). Only the growth in superconductivity can be considered to be a knowledge epidemic. All four arose due to new instrumentation and/or cheaper materials and are examples of puzzle-generating and enabling science. A second major result was that information epidemics are caused by as well as carried forward by groups of influential works. Third, increased activity in a given field is accompanied by an increase in conference papers. On the other hand, the journal literature of a given field is sufficient to represent the direction of literature growth accurately.

This work confirms and extends the epidemic model for the growth of literatures by demonstrating that not only does the contagion effect exist in physics but that there is also a catalyst effect present. It provides a statistical description for the growth and decline of fields of research.

RÉSUMÉ

Cette étude avait pour but d'examiner la fréquence des épidémies d'information dans la littérature de physique, pour comprendre si les points de poussée dans les séries chronologiques de croissance de cette littérature sont généralisés, si les publications qui montrent ces mécanismes de croissance résultent des travaux influents d'un groupe et si ces publications ont des caractéristiques particulières. Les épidémies d'information sont définies par rapport aux points de poussée dans la croissance de la littérature d'un domaine qui reflètent une phase d'excitation soudaine. L'hypothèse de base était que des épidémies d'information se déclarent souvent dans la littérature de physique et que les points de poussée observés pendant la croissance sont causés par un nombre de travaux influents qui attirent des nouveaux chercheurs et les poussent à publier abondamment. Des points de poussée qui finissent par transformer un domaine et mènent à la naissance d'une nouvelle spécialisation se nomment des épidémies de connaissances.

Les nombres mensuels de notices indexées dans chaque chapitre du *Physics Abstracts* entre 1977 et 1987 étaient tracés sur une carte de séries chronologiques, et la ligne de régression était encadrée d'une enveloppe de ± 3 d'écart-type. Tous les points qui dépassaient l'enveloppe étaient considérés comme potentiels pour des épidémies d'information. Les notices faisant partie de chaque point de poussée étaient identifiées dans le *Science Citation Index* et analysées en terme de diffusion (adresses d'auteurs) et d'impact (citations).

Les résultats démontrent que des épidémies d'information existent mais non de façon généralisée. Seulement quatre épidémies ont été identifiées dans les données. Elles se trouvent dans les chapitres 2 (méthodes mathématiques), 36 (agrégats), 73 (couches minces de semi-conducteurs) et 74 (supraconductivité) du *Physics Abstracts*. Seule la croissance en supraconductivité peut être considérée comme étant une épidémie de connaissances. Les progrès en instrumentation et les nouveaux matériaux sont des facteurs communs aux quatre épidémies et représentent des exemples favorisant l'épanouissement des sciences. Un deuxième résultat majeur est le constat que les épidémies d'information sont effectivement causées et stimulées par des groupes de travaux très remarqués. Troisièmement, l'activité croissante des publications dans un domaine est accompagnée par une croissance des communications à des conférences, bien qu'il soit suffisant de suivre la publication d'articles de périodiques pour représenter adéquatement la croissance d'un domaine.

Cette étude confirme et élargit le modèle des épidémies en démontrant qu'au delà de l'effet de contagion, il existe aussi l'effet catalyseur dans la croissance des littératures scientifiques. Elle fournit une méthodologie statistique pour décrire la croissance et le déclin des domaines de recherche.

For

Frances Groen

She opened the doors.

Was man nicht weiss, das eben brauchte man,
Und was man weiss, kann man nicht brauchen.

Goethe, Faust I[1]

---

1 What we don't know, that's just what we need,
And what we know, we cannot use.

ACKNOWLEDGEMENTS

First and foremost I wish to convey my deep sense of gratitude to my dissertation committee: my supervisor Prof. Lorna Rees-Potter, Prof. Andrew Large, Prof. Thomas Eisemon, Dr. Charles Davis, and Prof. Michael Mackey for having accepted to supervise me and for their many helpful suggestions. I am also indebted to Prof. Jamshid Beheshti for his numerous instances of encouragement and help.

The Social Sciences and Humanities Research Council honoured me with a doctoral fellowship and allowed me to start my studies. A bursary from the Faculty of Graduate Studies and Research of McGill University helped me obtain some of the data used in this study.

There is a number of staff members at McGill University who helped me in many ways during the last several years: Dorothy Carruthers and Kathryn Hubbard of the Graduate School of Library and Information Studies, Henry Lim Soo of the Physical Sciences and Engineering Library, Angella Lambrou and Fred Van Vliet of the Health Sciences Library, and Maria De Souza, Elizabeth Dunkley and Janice Simpkins of Interlibrary Loans of the McLennan Library, all at McGill University.

J'aimerai aussi exprimer mes sentiments de gratitude envers plusieurs collègues à l'Université de Montréal pour m'avoir soutenu, encouragé et aidé à travailler dans un environnement harmonieux. Les Profs. Marcel Lajeunesse et Gilles Deschatelets m'on engagé et guidé pendant trois ans de vie de chargé d'enseignement. Lyne Dufresne, Francine Gagnon-Benoit et My Luan Duong de la Bibliothèque de bibliothéconomie et des sciences de l'information ainsi que Mmes. Janine Cadet et

Francine Senez et M. Daniel Lefevre de la Bibliothèque de physique m'on encouragé et m'on aidé à obtenir des documents. Je suis aussi très reconnaissant à Lucie Carmel et à Lisette Morin-Jazouli pour leur soutien technique, financier et moral.

To my parents, Becky and Nesim Tabah, I am forever grateful for having brought me to Canada and continuously encouraged me to succeed and become better.

I also wish to acknowledge the contributions of a number of close friends who now form a family at large and who deserve a large amount of credit for providing me with many of the comforts and the high quality of life I now enjoy. They are my aunt and uncle Ida and Mony Azouz as well as Helen and Vladimir Kurgansky, Larry McGoldrick, Peter Illich, and Vladislav and Danielle Rovny-Bellrose. Helen and Vladimir put their "sanatorium" at my disposal and kept pushing me and haven't stopped. Larry helped me with his probing questions and needs for clarifications; he also spent many hours programming some of the tools used in the analysis of the data.

A number of those who would have been happy to witness the completion of this work are alas no longer present. I am particularly reminded of my friend and co-author Jaan Saber, my co-worker Helen Ellison and my old roommate Kay Hobbs, all of Concordia University. During occasional moments of reflection, Goethe's opening words to Faust keep coming back:

> Und manche liebe Schatten steigen auf ...
> Was ich besitze, seh ich wie im Weiten,
> Und was verschwand, wird mir zu Wirklichkeiten.[2]

---

[2] And certain beloved shadows rise up ...
What I possess, see I at a distance,
And what disappeared now becomes existence.

This thesis is dedicated to Frances Groen of McGill University. For over twenty years Fran has been my mentor, advisor and guardian. She introduced me to bibliometrics, encouraged me and helped me enter Library School and thus ensured that I could have a career. Quite literally, without her help in the late 1970s I would not have become a librarian and would not be at this juncture in life.

Finally, eternal dues are to my wife Margaret, my divine inspiration without whom there would be no life and no fun.

TABLE OF CONTENTS

List of Tables

Table | Page

List of Figures

List of Appendices

CHAPTER 1

INTRODUCTION

1.1. The Problem

This thesis seeks to study patterns of short-term rapid growth in the literature of physics and to identify the factors that characterize the rapid growth in its specialties.

Most models for the growth of science hold that scientific disciplines grow either exponentially or linearly and that, in any case, their literatures fully reflect that growth.[3] These models are based on the yearly numbers of articles abstracted in indexing journals such as *Physics Abstracts* or *Chemical Abstracts* and, as such, constitute a macro-level approach to growth. However, at a finer, micro-level, such as monthly patterns, growth is characterized by continuous ups and downs and occasional major spikes on a time-series chart.[4] The growth is at times linear, at times exponential and even, at times, negative.

This observation raises a number of questions, among them: 1. Is it possible that the yearly models force an "extreme smoothing" on the observations and that the occurrence of spikes is more prevalent than hitherto suspected; this especially in view of

---

[3]Derek J. de Solla Price, *Science since Babylon* (New Haven, CT: Yale University Press, 1961); Francis Narin, *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity* (Philadelphia: Computer Horizons Inc., 1976); M.R. Oliver, "The Effect of Growth on the Obsolescence of Semiconductor Physics Literature," *Journal of Documentation* 27 (1971): 11-17; and Donald W. King, Dennis D. McDonald, and Nancy K. Roederer, *Scientific Journals in the United States: Their Production, Use and Economics* (Stroudsburg, PA: Hutchinson Ross Publication Co., 1981).

[4] Albert N. Tabah, "Nonlinear Dynamics and the Growth of Literature," *Information Processing and Management* 28 (1992): 61-73.

the discontinuities and bursts of creativity that are prevalent in science?[5] 2. Is it possible that some of the spikes are caused by influential papers that attract a large number of contributors to publish intensively in a given specialty area for a given time and cause a dramatic increase in the growth pattern of that area?

These spikes are hereby termed "information epidemics" and are treated as part of a phenomenon that is a distinct entity within the process of knowledge growth. An information epidemic is regarded as a temporary occurrence, not a trend. It is a singular and highly visible event. There are several possible reasons for these phenomena, such as an influence from another specialty, a sudden influx of funding into a field or a surge of publications in it. In this work it is hypothesized that an information epidemic is caused by one or a set of influential publications that attract workers into a field, and draw them to work and publish for a certain duration of time. It is the result of work significant enough to get highly cited, provoke rapid growth in the knowledge base, and influence the publication of further important work in the field. Thus, *information epidemics* may be caused by influential papers that create temporary excitement. Certain influential papers have significant import, are cited heavily and contribute significantly to knowledge growth over time. A highly regarded paper that induces the publication of other influential and highly cited papers and ends up producing permanent knowledge growth would be deemed to have started a *knowledge epidemic*.

---

[5]Harriet Zuckerman and Joshua Lederberg, "Postmature Scientific Discovery?" *Nature* 324 (1986): 629-631.

Another motivation for this study comes from a paper entitled *Fashion in Science* by Diana Crane, who in the last paragraph states that

> The empirical study of fashion phenomena in science should begin with the location of exceptionally rapid growth, as indicated by the number of publications appearing per year.[6]

It needs to be pointed out here that the expression "fashion" carries a pejorative connotation and denotes something that is not only short-lived and of passing interest but also lacking seriousness. In contrast, an epidemic to a scientist in the middle of it is of the utmost seriousness because it is that person's "raison-d'être" and his/her entire motivation for undertaking the research of the moment.

The distinction being drawn here is an analytic rather than an empirical one. One would expect to find many more information epidemics where what seemed exciting and promising at first did not meet with success, and the publication volume in the field declined subsequently. Once in a while, though, an exciting idea gives rise to a high publication rate that is sustained for a long time, and the field is transformed permanently. The idea not only spurs on new research but also brings changes to the core knowledge of the field. The field attracts new workers and research funds, and the publication volume remains high. In that case one can speak of a knowledge epidemic, because the excitement not only gave rise to a high publication volume but also to a high participation rate that kept that volume high. In fact, results of this thesis will show if the empirical distinction is also valid. The diagram below illustrates the model at hand.

---

[6]Diana Crane, "Fashion in Science: Does it Exist?" *Social Problems* 16 (1968): 433-41.

INFLUENTIAL WORK.

Widespread excitement

INFORMATION EPIDEMIC (temporary event)

Ideas are rejected

Ideas are accepted

DIES OUT

KNOWLEDGE EPIDEMIC

KNOWLEDGE GROWTH

(Permanent change)

NEW SUBSPECIALTY

The focus of this thesis is the initial identification of information epidemics in a systematic manner. The above presentation is made for illustrative purposes only, and is not a model to be tested. The actual testing could be part of a future project once significant and influential works have been identified. For example, one could take a sample of citation classics identified in Eugene Garfield's previous writings and see whether their appearance was followed by an influx of new workers into their fields and an increase of new publications.[7] Citation classics, by definition, are works that have largely influenced their fields and the authors' coworkers. Garfield calls them *milestone papers*.[8] If a given citation classic is followed by an upsurge of new publications and it becomes the start of a new subspecialty, the work would be judged to have started a knowledge epidemic. If, on the other hand, it was followed by an initial expansion in the volume of publications that quickly subsided, then the work would be judged to have caused an information epidemic. Even though the work received a very high number of citations and became a citation classic, one would have to conclude that it was a temporary success on the part of the author but did not necessarily have a lasting influence on its field. Thus, the difference between an information epidemic and a knowledge epidemic in this context depends on whether the influence is temporary and fleeting or whether it is of sufficient significance to change the future course of its specialty.

---

[7]Eugene Garfield, *Essays of an Information Scientist*, 17 volumes. (Philadelphia: ISI Press, 1977-1994).

[8]Ibid., 5:124.

Recent literature provides several useful examples. A surprising paper in 1986 changed the future course of the field of superconductivity with the discovery of an unexpected family of materials that exhibit superconductivity at higher temperatures than ever attained before.[9] This paper not only led to an information epidemic of high-temperature superconductivity publications, but also started a knowledge epidemic by influencing the publication of a number of other works that have become highly cited themselves.[10] For example, in its first two years after publication (i.e. 1987-1988), the Bednorz and Muller paper obtained some 2124 citations (as reflected in *Science Citation Index*). During that same period, 11,088 items were abstracted by *Physics Abstracts* in superconductivity compared to 2594 in the previous two years (1985-1986). In a way, the notion of a knowledge epidemic is somewhat akin to the idea of *explosive papers* described by Garfield, Malin and Small.[11]

The literature of cold fusion provides the counter example. The information epidemic on cold fusion started following a press conference by Fleischmann and Pons, several months before the publication of their article in a refereed

---

[9] J. G. Bednorz and Kurt A. Muller, "Possible High Tc Superconductivity in the Ba-La-Cu-O System," *Zeitschrift fur Physik B - Condensed Matter* 64 (1986): 189-93.

[10] Albert N. Tabah, "Growth Patterns Following Sudden Discovery: The Case of Superconductivity Literature" (Graduate School of Library and Information Studies, McGill University, Montreal, 1991).

[11] Eugene Garfield, Morton V. Malin, and Henry Small, "Citation Data as Science Indicators," in *Toward a Metric of Science: The Advent of Science Indicators*, ed. Yehuda Elkana et al. (New York: Wiley, 1978), 195.

journal.[12,13] Their initial results and methodology were circulated by facsimile and put to scientific scrutiny. Soon problems with their work started circulating and the majority of the specialists concluded that their work had serious shortcomings and that their results were untenable. The field stopped growing, interest in the topic waned, most workers left the field, and the number of publications dropped dramatically.[14] While cold fusion caused an information epidemic, it never progressed toward a knowledge epidemic. A cursory online search in *Science Citation Index* and *Chemical Abstracts* showed that in the two years following the press conference (1989-1991) the authors' initial paper had obtained 386 citations while some 662 items had been published on the topic. The publication output went from forty-four in May 1989 to twelve in May 1990 to less than five in May 1991.[15]

To date no one has examined short-term dynamic phenomena that are here termed information epidemics. For one, all time series work has been done on a yearly basis. The only exception, that of Sullivan and Koester in the late 1970s, followed the growth of the electroweak interactions field by analyzing the co-citation

---

[12]Martin Fleischmann and Stanley Pons, "Electrochemically Induced Nuclear-Fusion of Deuterium," *Journal of Electrochemical Chemistry and Interfacial Chemistry* 261 (1989): 301-8.

[13]Frank Close, *Too Hot to Handle: The Story of the Race for Cold Fusion* (Toronto: Penguin Books, 1992).

[14]Bruce Lewenstein, "Cold Fusion and Hot History," Osiris 7 (1992): 135-63.

[15]Ibid., 161.

patterns among a chosen group of authors.[16] Second, spikes on time series charts, technically termed *outliers* have hitherto been regarded as anomalies by statisticians. Various methods have been devised to either eliminate them or find ways to reduce their impact on subsequent data - all that with the purpose of making reliable forecasts.[17] Instead, the emphasis in this thesis is on outliers themselves because they represent data that may be hiding significant truths about the progress of science (as evidenced by its literature). The only exception, again, has been Eugene Garfield who called them *shooting stars*.[18] His work, though, concentrated on the yearly citation patterns of influential papers. Thus, a systematic study of outliers in short-term publication data is entirely lacking.

There are few comprehensive surveys of the growth of physics literature. One is that by Keenan and Atherton who conducted a journal survey of the 1961 issues of *Physics Abstracts*. They supplemented it with a second survey in 1965.[19] Another is

---

[16] Daniel Sullivan et al., "Understanding Rapid Theoretical Change in Particle Physics: A Month-by-month Co-citation Analysis," in *Information Choices and Policies: Proceedings of the 16th ASIS Annual Meeting, Minneapolis, Minnesota, 14-18 October 1979*, (New York, Learned Information, 1979), 276-85.

[17] R. Douglas Martin, Alexander Samarov, and Walter Vandaele, "Robust Methods for ARIMA Models," in *Applied Time Series Analysis of Economic Data: Proceedings of a Conference* (Arlington, VA: U.S. Dept. of Commerce, 1983), 153-69; Helmut Thome, "A Box-Jenkins Approach to Modeling Outliers in Time Series Analysis," *Sociological Methods & Research* 23 (1995): 442-78.

[18] Eugene Garfield, "The Most-Cited Papers of All Time, SCI 1945-1988, Part 4: The Papers Ranked 301-400," *Current Contents* 21 (27 May 1991): 5-16.

[19] S. Keenan and Pauline Atherton, *The Journal Literature of Physics* (New York: American Institute of Physics, 1961); S. Keenan and F.G. Brickwedde *Journal Literature Covered by Physics Abstracts in 1965* (New York: American Institute of Physics, 1968).

the review by Anthony, East and Slater who compared several indexes and abstracts covering physics, among them *Physics Abstracts*.[20] A more recent paper by Vlachy compares the growth of physics subfields between 1955 and 1980 as obtained from *Referativnyi zhurnal - fizika*.[21] While they all give information on physics subfields at different periods of time, none have done time-series analyses nor have looked at short-term developments.

In terms of extensive time series work, D.H. Hall has been following the parallels in growth between the world literature on radioactive minerals and ores between 1935 and 1980, the state of the uranium industry, the rate of discovery of uranium ores and the production of uranium for various periods ranging between 1935 and 1985.[22] He has analyzed the time series for growth as well as time and frequency domain correlations and has found good correlation between literature production (as a reflection of the state of geoscience) and industry production figures. While his work is based on yearly figures, he does acknowledge the need for more detailed studies:

---

[20]L.J. Anthony, H. East and M.J. Slater, "The Growth of the Literature of Physics," *Reports on Progress in Physics* 32 (1969): 709-67.

[21]Jan Vlachy, "Publication Output in Physics Subfields," *Czechoslovak Journal of Physics* B 29 (1979): 829-36.

[22]D.H. Hall, "The Interface Between Geoscience and Industry: A Case Study of the Interaction Between Research and the Discovery and Mining of Ores for Nuclear Fuels," *Scientometrics* 11 (1987): 199-216; "Rate of Growth of Literature in Geoscience from Computerized Databases," *Scientometrics* 17 (1989): 15-38; and "The Science-industry Interface: Correlation of Time Series of Indicators and their Spectra, and Growth Models in the Nuclear Fuel Industry," *Scientometrics* 24 (1992): 237-80.

> What is required of these [science] indicators are longer time series and finer breakdowns of science statistics into specific fields.[23]

The only recent work directly focused on rapidly growing literatures is that of John Budd and C.D. Hurt who have counted the number of citations obtained by an influential paper on superstring theory over five years and tested the distribution against growth models suggested by Price and Kuhn.[24] They conclude that the distribution does not fit theoretical models and suggest that a different mechanism is at work for what they call *fast literature*.[25] They define a fast literature as:

> denoted by its very rapid citation impact, citation frequency, and its concomitant swift diffusion into the literature of the specialty.[26]

The same definition of a fast literature will be adopted for this thesis.

Unfortunately, Budd and Hurt have followed only one article, that of Schwarz in 1982.[27] However, it is not clear why they chose this one when the Green and

---

[23]Hall, *The Science-industry Interface*, 240.

[24]J.M. Budd and C.D. Hurt, "Superstring Theory: Information Transfer in an Emerging Field," *Scientometrics* 21 (1991): 87-98; C.D. Hurt and J.M. Budd, "Modelling the Literature of Superstring Theory: A Case of Fast Literature," *Scientometrics* 24 (1992): 471-80; Derek J. de Solla Price, "Networks of Scientific Papers," *Science* 149 (1965): 510-5; and Thomas S. Kuhn, *The Structure of Scientific Revolutions*, 2nd ed. (Chicago: University of Chicago Press, 1970).

[25]Budd and Hurt, *Superstring Theory*, 97.

[26]Hurt and Budd, *Modelling the Literature of Superstring Theory*, 472.

[27]John H. Schwarz, "Superstring Theory," *Physics Reports* 89 (1982): 223-322.

Schwarz article of 1984[28] is the one that started the first real excitement in superstring theory and by the end of 1992 had garnered more citations than the older Schwarz article of 1982 (932 for Green and Schwarz of 1984 versus 763 for Schwarz of 1982). Their analysis follows yearly citation rates which in the case of fast growing literatures is a rather unsatisfactory approach. What makes fast-growing literatures special is that their growth reaches exponential proportions within a few months of the publication (or impact) of the first influential paper. They show explosive characteristics that attract the attention of a large number of workers both within the field and without. In addition, citations to one paper alone are not very helpful in sensing the direction in which a field is moving. Most epidemics are caused by a group of works. Budd and Hurt's concentration on one paper alone may be telling only part of the story.

Citations take place in a social context.[29] No paper can form an isolated and unique field or subfield. That social context is provided by the population of other publications and cited works on that subject. From this, the need becomes even stronger to look at how a field develops as a whole and how citations to important works grow along with this development. The development of a field can be measured

---

[28]Michael B. Green and John H. Schwarz, "Anomaly Cancellations in Supersymmetric d=10 Gauge-theory and Superstring Theory," *Physics Letters B* 149 (1984): 117-22.

[29]In the context of this thesis, the words *reference* and *citation* are used in their technical context: A reference is an acknowledgement of someone's influence on a given work. A citation is the acknowledgement one receives from someone else's work.

by the growth in the number of publications in it and the increasing number of authors participating in the work.

Thus, the methodology adopted in this thesis draws a combination of ideas from the two sets of authors mentioned above: finer and more detailed time series of fast growing literatures. If a phenomenon transpires very rapidly over time, its detailed understanding demands data on short-term occurrences. Fast growth as a result of the influence of a significant work becomes an epidemic occurrence. Thus, it becomes easy to accept the term information epidemic to denote a rapid and significant increase in the number of papers published following the appearance of a significant work.

Based on the above discussion, the purpose of this thesis is to:

1. characterize the sudden short-term growth patterns,
2. establish the prevalence of information epidemics, and then
3. characterize the literatures of information epidemics.

1.2. The Study of Growth

There are several reasons for studying the growth of science and its subspecialties. For one, it can be considered a branch of epistemology. Understanding why we know what we know is akin to trying to understand how things come about, and that is the fundamental curiosity driving science. Second, the study of growth leads to understanding the advancement of knowledge and the conditions that surround discovery. Third, growth and the specialization that follows it is the hallmark of modern science. There is a need for deeper understanding of the basic mechanisms of growth, fragmentation and reinterpretation of science in the light of new discoveries. Fourth, the volume and rate of innovation in certain fields of science continues to accelerate despite the overall slowdown in science itself.[30] The idea of progress and development is synonymous with the notion of growth.

The theoretical basis of this thesis is drawn from two well-accepted domains:

1. Philosophical: Major changes in ideas and concepts in a given field bring renewed interest into that field. This is in part connected with Kuhn's model of paradigm shifts.[31]

2. Sociological: When a major shift occurs in a field and it results in the birth of a new subspecialty, new people are attracted to that subspecialty.

There are, in the main, four approaches to the study of the growth of science and its specialties[32]:

1. Historical: to follow the movement of ideas and people in the relatively distant past, concentrating on the internal development of specific fields. One example is Gerald Holton's *Thematic Origins of Scientific Thought*.[33]

---

[30]John Ziman, *Science in a Steady State* (London: Science Policy Support Group, 1977).

[31]Kuhn, *The Structure of Scientific Revolutions*, 160.

[32]Adapted from G. Lemaine et al., eds. *Perspectives on the Emergence of Scientific Disciplines* (Paris: Mouton, 1976), 1-23.

[33]Gerald Holton, *Thematic Origins of Scientific Thought: Kepler to Einstein* (Cambridge, MA: Harvard University Press, 1973).

2. Sociological: to follow the social processes associated with the activities of scientists and to concentrate more on the structural and networking aspects within given fields. The works of Merton, Lemaine and Knorr-Cetina[34], just to give a few examples, belong to this category.

3. Philosophical: to follow the truth claims of scientific knowledge and obtain insights into the process of scientific discovery. Recent significant examples include works by Popper, Kuhn and Rescher.[35]

4. Information science: to follow the published literature and infer from the growth of the literature the movement of ideas and associations between scientists. The best known proponent of this line of work is Derek Price.[36] There are also contributions from Menard and Kochen.[37]

Those trying to get at a deeper and more holistic picture of the development of a scientific field may want to somehow combine all four approaches in the hope of arriving at more reliable conclusions. However, such an undertaking would be

---

[34]Robert K. Merton, *The Sociology of Science* (Chicago: University of Chicago Press, 1973); Lemaine, *Perspectives*, 1-23; and Karin D. Knorr-Cetina, *Science Observed: Perspectives on the Social Study of Science* (London: Sage Publications, 1983).

[35]Karl R. Popper, *Conjectures and Refutations: The Growth of Scientific Knowledge* (New York: Harper, 1963); Kuhn, *The Structure of Scientific Revolutions*, 1970; and Nicholas Rescher, *Scientific Progress* (Pittsburgh, PA: University of Pittsburgh Press, 1978).

[36] Price, *Science since Babylon*, 1961; and *Little Science Big Science* (New York: Columbia University Press, 1963).

[37]Henry W. Menard, *Science: Growth and Change* (Cambridge, MA: Harvard University Press, 1971); Manfred Kochen, *Integrative Mechanisms in Literature Growth* (Westport, CT: Greenwood Press, 1974).

impractical. The richness of the data would be more than offset by the volume and complexity of interpretative possibilities.

The emphasis in this thesis, in terms of its framework, methodology and interpretation of results, will rely on the principles of information science. The strategy involves following the process started by Derek Price and using publication indicators to obtain a deeper understanding of short-term growth patterns in the literature of physics. Wherever possible, useful principles will be borrowed from the other three above-mentioned approaches to bring a more balanced and complete perspective into view. The literature review will concentrate on the information science background and occasionally draw from the other three to supplement relevant and appropriate information.

In addition, this study is data-descriptive. That is, rather than starting from an established and well-described problem in information science - be it theoretical, be it empirical - it starts with an interesting observation and focuses on the discovery of new patterns with the eventual purpose of generating hypotheses on the behaviour of fast growing literatures. This is somewhat analogous to the use of a new instrument (such as a telescope, a microscope or nowadays a computer) to uncover new patterns among old observations.

The notion of growth reflects the development of ideas conveyed by an efficient communication system in science. The communication system in science is public. Therefore, it is accepted here that the evolution of ideas giving rise to the growth of a scientific field is fully reflected in the growth of its published literature.

By the same token, one needs to discount ideas that never find success as measured by numbers of publications as well as those that remain unpublished (by choice or by rejection).

1.3. Definitions

Given that there are several new concepts introduced in this thesis, an appropriate definition of these concepts is useful. *Growth* (although not a new concept) is taken to be the increase in the quantity of articles published over time, accompanied by an increase in the size of the field of interest and/or the number of participants in the field in which the growth is taking place. An *epidemic* is a sudden increase in the incidence rate of an occurrence (such as a disease) well above the value regarded as normal. In health related terminology, it affects large numbers of people and is spread over a large area. In this thesis "well above" is defined at a level surpassing the 3.0 standard deviations envelope surrounding the regression line of a time series chart. This is further explained under Section 4.4 in the Methodology chapter.

An *information epidemic*[38] denotes a rapid and significant increase in the number of papers published in a field of study. It is hypothesized that an information epidemic takes place following the appearance of a significant work, but the effect is short lived and the publication volume soon comes down again. A *knowledge*

---

[38]The origin of the term owes its birth to 1991 when this author's advisor, Charles H. Davis, then at the Science Council of Canada, during a conversation exclaimed "Ah, you mean information epidemics!"

epidemic, in turn, is caused by an influential work that produces permanent changes in the core knowledge of a given area of endeavour and is followed by an upsurge of new publications that bring about the start of a new subspecialty. As mentioned above, the difference between an information epidemic and a knowledge epidemic is one of permanence. Once an influential work is followed by the publication of other influential works that produce a chain reaction and give rise to the birth of a new specialty, a new plateau is reached in terms of literature growth.

With respect to the spread of ideas, there is a number of other terms such as *diffusion, spread*, and *adoption* that are used here in parallel fashion. Spread is the most general term and refers to the distribution or scatter of an idea or concept over an area of interest or study. Diffusion is similar to epidemics, except that it is slower and lacks a time vector. Van Vianen and Van Raan define diffusion as

> the use of field-specific knowledge in other fields of science and technology. If this use leads to a further generation of knowledge on the original subject, we may speak about 'knowledge expansion.'[39]

The suggestion here is that if the further generation takes place very rapidly we have an information epidemic at hand. Finally, adoption refers more to technological adoption and is concerned with the economic significance of technological discoveries. Some of these differences will be further discussed below (page 34).

---

[39]Ben Van Vianen and Anthony Van Raan, "Knowledge Expansion in Applied Science: A Bibliometric Study of Laser Medicine and Polyimide Chemistry," in *Dynamics of Science- Based Innovation*, ed. H. Grupp (New York: Springer, 1992), 227.

With respect to areas of study, a *specialty* is a loosely bound set of work groups linked informally and working on closely related problems. It comes as a result of the division of labour among occupations. The number of groups in a specialty worldwide is small, ranging from fifty to a hundred. Specialties exist usually within disciplines represented by university departments and scientific societies.[40] A discipline is a:

> body of specialized knowledge and skills but also political institutions, demarcating areas of academic turf and structuring claims on resources. ... Disciplines are scientists' chief reference group. [They] prescribe a division of labour and channel communication between different groups of specialists.[41]

In turn, a *field* is a sphere of influence or interest, a domain embraced within a given study, whereas an *area* is a set of phenomena having some common and unifying characteristic.[42] Thus, in terms of stratification, science can be represented as consisting of disciplines which are divided into specialties which themselves consist of fields and areas. Although, by the above definition, an "area" is a smaller entity than a "field", the differences between the two are small enough that they will be used interchangeably.

Finally, the identification of epidemics takes place through the analysis of points that stand out in time series charts. In general, a point or value that stands out well beyond the average values surrounding it and thus is inconsistent with them, is

---

[40]*Encyclopedia of Sociology*, s.v. "specialty."

[41]*Dictionary of the History of Science*, s.v. "discipline."

[42]*Dictionary of Sociology*, s.v. "field."

called a *spike* on a time series chart. In statistical parlance it is also called an *outlier*. In general, outliers may be due to recording errors, calculating errors or unusual events. Care has been taken, and the data have been rechecked several times, to ensure that all outliers analyzed in this thesis are intrinsic to the data and that they are due to unusual events. In this thesis a clear and significant differentiation will be made between a spike and an outlier. An outlier is defined as any point that crosses the ± 3.0 standard deviation envelope around a regression line (for further discussion see page 88 below). Anything within the ± 3.0 standard deviation envelope will be considered a spike and will be ignored. As mentioned above, in standard time series analysis, outliers are considered to be aberrations and are the subject of intervention analysis to eliminate their influence on subsequent data. On the other hand, the interest in this work bears on the very analysis of outliers and the reasons for their occurrence.

CHAPTER 2

BACKGROUND

2.1. Theoretical Background

2.1.1. Growth of Literature

The impetus for recent studies on the growth of literature originates with Derek Price. In 1951 he proposed that science in general grows exponentially, doubling its size every ten to fifteen years.[43] He popularized his ideas in 1961 but also added that the exponential rate could not be sustained for a long time and that the growth of science was about to slow down and enter a steady-state level.[44] While May and Line agreed with his overall thesis, Oliver and King argued that scientific literature more likely grows linearly.[45]

Menard found differing growth rates in different literatures. He found linear growth in optics and acoustics between 1900 and 1960. However, in two subfields of physics (nuclear and solid state) he found exponential growth, though at different rates.[46] He classified subfields according to three patterns of growth: slowly growing

---

[43]Derek J. De Solla Price, "Quantitative Measures of the Development of Science," *Archives internationales d'histoire des sciences* 14 (1951): 85-93.

[44]Price, *Science Since Babylon*, 113.

[45]K.O. May, "Quantitative Growth of the Mathematical Literature," *Science* 154 (1966): 1672-3; Maurice B. Line, "The Half-life of Periodical Literature - Apparent and Real Obsolescence," *Journal of Documentation* 26 (1970): 46-52; Oliver, *The Effect of Growth on the Obsolescence*, 11-17; and King, McDonald and Roederer, *Scientific Journals in the United States*, 1981.

[46]Menard, *Science: Growth and Change*, 50.

- old and large fields that grow constantly but relatively slowly (e.g. economic geology or geomorphology), rapidly growing - showing consistent growth trends with doubling times of only 5-10 years (e.g. continental drift, geochemistry) and cyclical subfields where support or interest appear to fluctuate (e.g. petroleum geology or structural geology).[47]

Throughout the 1960s several Soviet workers took exception to the exponential model but their work hardly received any mention in the western literature.[48] Moravcsik tried to explain exponential growth in relation to research fronts growing at *the epidermis of science*.[49] Some others, such as Bottle and Rees found a zig-zag pattern in the literature on liquid crystals where growth, once it resumed, took on exponential characteristics.[50] In their survey, which spanned 1910-1972, they also found that the core journals changed considerably between periods of growth. Hall found a similar mechanism in the geoscience literature for 1940-1980 where overall the literature doubled every 8 years.[51]

---

[47]Ibid., 54-5.

[48]Leonard N. Beck, "Soviet Discussion of the Exponential Growth of Scientific Publications" in *The Information Conscious Society: Proceedings of the Thirty-Third ASIS Annual Meeting, Philadelphia, 11-15 October 1970*, ed. J.b. North (Philadelphia: American Society for Information Science, 1970), 5-17.

[49]Michael J. Moravcsik, "Phenomenology and Models of the Growth of Science," *Research Policy* 4 (1975): 80-6.

[50]R.T. Bottle and M.K. Rees, "Liquid Crystal Literature: A Novel Growth Pattern," *Journal of Information Science* 1 (1979): 117-9.

[51]Hall, *Rate of Growth of Literature in Geoscience*, 35.

Despite extensions and disagreements over the last three decades, the exponential paradigm has nevertheless remained the best known and most publicized mechanism of literature growth.[52]

Hawkins, in part influenced by the work of William Goffman, examined the literature of noble gas compounds.[53] He found patterns of sudden start and rapid growth as well as movements of large numbers of investigators in and out of subspecialties. Most of those remained active in the field for only short periods of time. He characterized the overall growth of the literature of noble gas compounds as "a sudden spurt of interest following initial discovery, followed by a decline, and then a moderate growth".[54]

Current work on growth has been featuring either curve-fitting exercises to available data or trying to develop informetric models to simulate mechanisms of growth. Wolfram et al. have used 20 years of bibliographic data from some 20 databases and tried to fit three growth models to the data: linear, exponential and their

---

[52]Jean Tague, Jamshid Beheshti, and Lorna K. Rees-Potter, "The Law of Exponential Growth: Evidence, Implications and Forecasts," *Library Trends* 30 (1981): 125-50; L. Egghe and I.K. Ravichandra Rao, "Classification of Growth Models Based on Growth Rates and its Applications" *Scientometrics* 25 (1992): 5-46.

[53]Donald T. Hawkins, "The Literature of Noble Gas Compounds," *Journal of Chemical Information and Computer Science* 18 (1978): 190-9; William Goffman and Vaun A. Newill, "Generalization of Epidemic Theory: An Application to the Transmission of Ideas," *Nature* 204 (1964): 225-8.

[54]Hawkins, *The Literature of Noble Gas Compounds*, 199.

own power model.[55] They have found that the linear and power models fit the data well whereas the exponential model showed the poorest fit. They conclude that "the breakdown in exponential growth is well underway and is giving way to linear growth".[56]

Egghe and Rao have reexamined Wolfram's data and have found that the exponential model never occurs and that only power models and Gompertz models are applicable to the data.[57] Power models fit scientific and technical online databases better whereas the Gompertz S-shaped distribution shows better fit to social sciences and humanities online databases.

Czerwon examined publication and citation indicators in the growth of a new specialty, that of "Monte Carlo methods in lattice field theory".[58] He analyzed factors such as publication counts, impact factor, relative citation rate, aging and scattering among 668 articles indexed by INIS Atomindex between 1979 and 1984. Among others, he concluded that the specialty provides a characteristic example of the growth of a *new science subfield from a core body of seminal literature*.[59] While his

---

[55]D. Wolfram, C.M. Chu and X. Lu, "Growth of Knowledge: Bibliometric Analysis Using Online Database Data," in *Informetrics 89/90: Selection of Papers Submitted for the Second International Conference on Bibliometrics, Scientometrics and Informetrics, London, Ontario, 5-7 July, 1989* (Amsterdam: Elsevier, 1990), 355-64.

[56]Ibid., 362.

[57]Egghe and Rao, "Classification of Growth Models," p. 5.

[58]H.J. Czerwon, "Scientometric Indicators for a Specialty in Theoretical High-energy Physics: Monte Carlo Methods in Lattice Field Theory," *Scientometrics* 18 (1990): 5-20.

[59]Ibid., 18.

work is valuable and provides a useful model for future scientometric investigations of growth areas, his data are based on a 'time-series' of six points of annual data and he does not compare how the growth of the specialty stands out in comparison to other specialty areas in physics. As well, while there was exponential growth over 6 years, the total numbers involved are rather small in that he is not dealing with a fast growing literature.

Thus, while the exponential growth mechanism is the one most people will cite, it is no longer valid in all cases. The growth of a given scientific literature depends on the specialty and the time period at hand and may at different times exhibit different mechanisms. As De Mey has pointed out:

> However, despite important practical considerations and suggestions implied by this growth model of science, its importance for our understanding of paradigms lies not so much in the characteristics of the growth of science on a global scale as in the analysis of the detailed mechanisms of this growth.[60]

In agreement with this assertion, the emphasis in this study is on short-term and sudden spurts (outliers) on a time series chart that give rise to the question: Are these outliers meaningful in terms of the growth of knowledge and the evolution of ideas?

---

[60]Marc De Mey, *The Cognitive Paradigm: An Integrated Understanding of Scientific Development* (Chicago: University of Chicago Press, 1992), 114.

2.1.2. Theory of Epidemics and the Diffusion of Ideas

2.1.2.1. Epidemics

Comparisons between the transmission of ideas and epidemic processes date back to the mid-nineteenth century. In 1856 Sir Bernard Brodie was the first to state that *there are epidemics of opinion as well as disease.*[61] In 1915 Sir Ronald Ross developed the first generalized algebraic formulation of epidemics adaptable to the description of a variety of situations and coined the phrase "A Theory of Happenings".[62] He pointed out that his results could also be applied to fields such as economics and sociology. Alfred Lotka revised Ross' work and, expressing his model in differential terms, obtained a logistic equation to describe population growth.[63]

A similar approach was used by W.O Kermack and A.G. McKendrick to develop a mathematical epidemic theory.[64] They devised a set of three coupled differential equations to reflect the relationships between the number of susceptibles, the infection rate, the recovery rate, and the removal (death) rate.

William Goffman used these equations in the 1960s to develop his own epidemic theory on the diffusion of ideas and the growth of scientific specialties.

---

[61]B. C. Brodie, *Psychological Inquiries: In a Series of Essays, Intended to Illustrate the Mutual Relations of the Physical Organisation and the Mental Faculties*, 2nd. ed. (London: Longman, Brown, Green, and Longmans, 1855), 26.

[62]Ronald Ross, "Some a Priori Pathometric Equations," *British Medical Journal*, (March 27 1915): 546-7.

[63]Alfred J. Lotka, "Contributions to the Analysis of Malaria Epidemiology," *American Journal of Hygiene* 3 (Suppl.1, 1923): 1-121.

[64]W.O. Kermack and A.G. McKendrick, "A Contribution to the Mathematical Theory of Epidemics," *Proceedings of the Royal Society A* 115 (1927): 700-21.

Goffman regarded the transmission of a scientific idea within a population as being analogous to the spread of a communicable disease in society. He took the published article as the agency of transmission (vector) and characterized the transmission as an idea moving from an *infective* (author of a paper) to a *susceptible* (*reader of a paper who will be infected given effective contact*).[65] The analogy to the standard model of disease spread indicates that a susceptible person may become infected as a result of contact with an infected person.[66] The newly infected individual then passes through a latency period (publication lag time) and then through an infectious time (time during which the work is cited) before becoming removed by no longer being active in the field. Thus, the spread of ideas within a scientific community can be seen as a variation of an infection process.[67]

Goffman tried to apply the theory of epidemics to the problem of growth in science. By using mast cell research and the growth of symbolic logic in mathematics as case studies, he tried to demonstrate that it was possible to see growth and development as sequences of overlapping epidemics.[68] He also demonstrated that

---

[65]Goffman, *Generalization of Epidemic Theory*, 225.

[66]Kermack and McKendrick, *Mathematical Theory of Epidemics*, 701.

[67]A. Siegfried, *Germs and Ideas* (London: Oliver and Boyd, 1965).

[68]William Goffman, "Mathematical Approach to the Spread of Scientific Ideas: The History of Mast Cell Research," *Nature* 212 (1966): 449-52; and "An Application of Epidemic Theory to the Growth of Science (Symbolic Logic from Boole to Godel)," *Progress of Cybernetics* 3 (1970): 971-84.

once a producer left a given subspecialty in science, he/she was never expected to return to it.[69] He summarized his work by stating that:

> this approach makes it possible to establish, quantitatively, the relative importance of past lines of inquiry within a given area of scientific activity and to predict the future behaviour of existing lines of investigation as well as the emergence of new ones within the area.[70]

Unfortunately, Goffman's model of three simultaneous differential equations is full of simplifying assumptions and parameters that are difficult to estimate or derive empirically.

It should be pointed out that generalized mathematical treatments of epidemics have not only been developed for contagious diseases but also for the spread of rumours, the spread of riots, the diffusion of innovations, the propagation of consumer goods as well as the dynamics of technological progress.[71] For example, Burbeck et al. describe the spread of a riot in the following terms:

> A riot attracts and infects individuals, many of whom are originally merely detached and indifferent spectators and bystanders. At first,

---

[69]William Goffman, "A Mathematical Method for Analyzing the Growth of a Scientific Discipline," *Journal of the Association for Computing Machinery* 18 (1971): 173-85.

[70]Ibid., 173.

[71]Klaus Dietz, "Epidemics and Rumours: A Survey," *Journal of the Royal Statistical Society, Ser. A* 130 (1967): 505-528; Stephen L. Burbeck, Walter J. Raine, and J.Abudu Stark, "The Dynamics of Riot Growth: An Epidemiological Approach," *Journal of Mathematical Sociology* 6 (1978): 1-22; Everett M. Rogers, *Diffusion of Innovations*, 3rd ed. (New York: Free Press, 1983); Vijay Mahajan and Yoram Wind, *Innovation Diffusion Models of New Product Acceptance* (Cambridge, MA: Ballinger Publishing Company, 1986); and Louis A. Girifalco, *Dynamics of Technological Change* (New York: Van Nostrand Reinhold, 1991).

> people may be merely curious about the given behaviour, or mildly interested in it. As they catch the spirit of excitement and become more attractive to the behaviour, they become more inclined to engage in it.[72]

> Riots, although more spectacular and destructive, are indeed close relatives of the more common collective behaviors; ... the fact that the time course of riot behaviour is similar to the familiar growth of fads or rumours indicates a similarity of dynamics.[73]

In fact, the approach of this thesis is based on the acceptance that the dynamics of spread in society involves the diffusion of information and that diffusion works by a self-consistent mechanism. This mechanism may show slight variations depending on whether the spread involves a rumour, the growth of knowledge or the spread of action in a given population. Information epidemics also belong to this category with the special characteristic that they are dramatic events that occur within short periods of time.

This approach is reflected by David Fan who advocates shifting the emphasis away from infectives and susceptibles and instead paying more attention to the content of messages transmitted between people.[74] According to him:

> Such a structure permits predictions about the rate of change of ideas in society based solely on the information available to the population.[75]

---

[72]Burbeck et al., *The Dynamics of Riot Growth*, 4.

[73]Ibid., 21.

[74]David B. Fan, "Ideodynamics: The Kinetics of the Evolution of Ideas," *Journal of Mathematical Sociology* 11 (1985): 1-23.

[75]Ibid., 1.

Increasingly, with the availability of electronic forms of communication, it is becoming possible to pay attention to the messages transmitted and their change over time, with the purpose of gaining insight into the evolution of ideas. Fan terms this process *ideodynamics* and states that it is possible to codify information content into units he calls *infons*, although he offers no information on its units of measurement.[76] The advantage of ideodynamics is that it emphasizes the time course in the spread of an idea regardless of the inherent values of the messages transmitted. He claims his version of differential equations can be used to predict the time course in the development of an idea or a habit based on the information content transmitted in a population.[77] The veracity of his claims has yet to be demonstrated.

The difficulty with Fan's idea is that it has remained in the theoretical domain. While he has developed all the necessary tools to do ideodynamic analysis, the operational definition and quantification of an "infon" is as yet unrealized. To date, no metric for an "idea" has been developed. In other words, there is no satisfactory way to define an idea "unit" and to measure it:

> The lack of a basic metric of information means that much of our research ... rest[s] upon questionable premises and approximate measures.[78]

---

[76]Ibid., 1.

[77]David B. Fan, "Ideodynamic Predictions for the Evolution of Habits," *Journal of Mathematical Sociology* 11 (1985): 265-81.

[78]Blaise Cronin, "When is a Problem a Research Problem?" in *Applying Research to Practice: How to Use Data Collection and Research to Improve Library Management Decision Making*, ed. L.S. Estabrook (Urbana-Champaign, IL: University of Illinois Graduate School of Library and Information Science, 1992), 121.

To date, the most practical measuring unit of an idea remains, despite all its imperfections, the published article - the same unit used to measure the spread of information and its growth in science as well as the epidemic growth of an idea and its ramifications.

From a sociological point of view, epidemic theory closely resembles Herbert Menzel's *contagion theory*.[79] Menzel examined the transmission of ideas in a medical community with the purpose of finding out what it was that influenced the diffusion of an idea or an innovation. His results indicated that the closer an individual was to the medical community and the more integrated he/she was into it, the sooner that person was likely to adopt a given product. This is not very different from established epidemic models in that the more often a susceptible individual is exposed to a disease, the faster that person will become an infective. The main difference between the two is that Goffman's is a mathematical model whereas Menzel's is a sociometric model; otherwise they are equivalent in terms of modelling an information process.[80] Menzel's model is influenced by a two-step communication model and reflects changes in a network of potential adopters. Goffman's model is influenced by a mathematical model and, when it is possible to apply it, attaches a

---

[79]Herbert Menzel and Elihu Katz, "Social Relations and Innovation in the Medical Profession: The Epidemiology of a New Drug," *Public Opinion Quarterly* 19 (1955): 337-52.

[80]Dennis B. Worthen, "The Epidemic Process and the Contagion Model," *Journal of the American Society for Information Science* 24 (1973): 343-6.

numeric value to the evolution of an epidemic and tries to bring prediction to the number of people that will be *afflicted* with an idea.

In sum, epidemic theory provides a model (and explanation) for the rapid spread of a given collective behaviour in a given population. As such, it conveniently summarizes a large amount of data and provides insight into the mechanisms of spread. With the epidemic model in mind, it is much easier to see that a sudden growth point in the progress of a field may have arisen as a result of a swift information epidemic within a group of researchers who may have been influenced by a new and exciting piece of work.

In this context of epidemics as suddenly occurring and fast moving phenomena, two interesting pieces of insight need to be mentioned here: one comes from John Pierce who considers that the purpose of science is to surprise, to come up with unexpected ideas and results whereas the purpose of technology is to avoid surprise, by making reliable products.[81] Thus, when studying science, one needs to remain alert to surprising discoveries and to popular rushes toward exciting possibilities. The second insight comes from Magyar who states that in science experimental facts do not change; it is scientific ideas and the re-interpretations of experimental results that change and bring on revolutions in ideas and concepts.[82]

---

[81]John R. Pierce, "Research and Surprise," in *International Symposium on International Cooperation and Competition in Science and Technology, Tokyo, 12-13 April, 1988* (Tokyo: Engineering Academy of Japan, 1988), 1-5; John R. Pierce, letter to author, 30 July 1991.

[82]George Magyar, "How Trustworthy are Experimental Facts?" *European Journal of Physics* 2 (1981): 244-9.

Experimental facts have an accumulative character whereas it is scientific hypotheses and theories that undergo revolutionary changes. Again, sudden discoveries or surprising publications may influence a large multitude to see old facts and rework old theories in a new light and thus offer surprising solutions to old problems (a classical case of Kuhnian paradigm shift). The rush to rework and to publicize may in fact result in the information epidemics that are the focus of this thesis. Given this framework, the discovery of information epidemics in scientific specialties should not be surprising events.

It should be pointed out, however, that there are two kinds of epidemics and two separate models to pay attention to. The difference depends on the presence of a vector of transmission. The first can be called a *Contagion Model.* It is vector-based and follows the transmission of an infectious agent in a population. It is largely grounded in the Kermack-McKendrick equations.[83] This is the one that has largely influenced the models of communication and diffusion of ideas. The second can be called a *Catalyst Model.*[84] It is not vector-based. It pays attention to factors affecting the incidence of non-communicable diseases in a population. A number of or a cumulation of factors are involved. The causes of the epidemic are either unknown or at least questionable.

Both types of models are applicable to this thesis. The ability of an influential work to infect researchers with an important idea or result clearly belongs to the

---

[83]Kermack and McKendrick, *Theory of Epidemics*, 701.

[84]Ole S. Miettinen, *Theoretical Epidemiology* (Albany, NY: Delmar, 1985).

Contagion Model. However, if it is assumed - as a working approach - that information epidemics are prevalent and occur frequently in science, and if one is looking for factors that bring on such epidemics, one also needs to take the Catalyst Model into consideration. The medical equivalents of these two are measles and heart disease, respectively. The first is transmitted by a carrier (a virus). The second, although devoid of a carrier, is nevertheless present in epidemic proportions in society due to the confluence of a number of factors (such as smoking, high cholesterol diet and bad lifestyle habits) that when present in an individual often brings on (or contributes to) heart disease. Thus, during the course of this work when information epidemics come into question, they will refer to the Contagion Model. On the other hand, when the discussion involves factors that predispose a work to bring on an information epidemic, then the attention will be on the Catalyst Model.

2.1.2.2. Diffusion

The difference between an epidemic process and a diffusion model is that an epidemic is an entirely time-dependent model with a contagion vector whereas diffusion has no vector of communication. In other words, both involve time but diffusion lacks a vector and usually proceeds more slowly. Everett Rogers, in his *Diffusion of Innovations* defines diffusion as:

> The process by which (1) an innovation (2) is communicated through certain channels (3) over time (4) among members of a social system.

> The four elements ... are identifiable in every diffusion research study, and in every diffusion campaign or program.[85]

However, not every writer makes time an explicit variable in his/her model.

Diffusion refers more to a process of dispersal. It is akin to a random walk, either geographically or through a gradient and over a whole population. It does not change individuals but affects the characteristics of the whole population. It is a random process. In contrast, an epidemic changes the behaviour of certain individuals in the population and it is not random. It is a rapid phenomenon, and it is a directed process. In diffusion, the rate between events is constant whereas in epidemics it accelerates in the exponential phase and slows down at the saturation phase. Strang states that in epidemic models rates increase monotonically with time whereas in diffusion models they are globally constant over time.[86]

It should be pointed out that the literature reflects a general state of confusion over appropriate terminology, and several authors talk about epidemic models when they need to say *diffusion* and vice versa. One even uses the term *an epidemic model of diffusion* and uses the two terms interchangeably.[87] The purpose of the above discussion was to point to some of the available literature and to the different possibilities for visualizing information epidemics. Both are the manifestation of the communication process in science. The reason why the expression *epidemic* was

---

[85]Rogers, *Diffusion of Innovations*, 10.

[86]David Strang, "Adding Social Structure to Diffusion Models: An Event History Framework," *Sociological Methods & Research* 19 (1991): 324-53.

[87]Ibid., 324.

chosen in the title of this dissertation is because of existing precedent in the information science literature with Goffman. As well, an epidemic process often involves an agent of transmission (in this case, the influential work), and that the time factor is used explicitly to reflect the dynamics of the process. Irrespective of the preferred model, an information epidemic will exhibit itself through rapid growth over time followed by an equally rapid decline.

There are several models of information diffusion. Avramescu has likened the mechanics of information dissemination in science and technology to a heat diffusion process.[88] Le Coadic assumes that the diffusion of scientific ideas is a dynamic process and follows the same logistic curve as the diffusion of innovations and of rumours.[89] He thus comes closer to Goffman's epidemic model or Menzel's Contagion Model.[90]

Bartholomew has suggested that the difference between models is that if the diffusion is modelled as being propagated by an external source, the growth curve will look like an inverted j-curve. If the model includes interpersonal contact, the curve will look like an S-curve.[91] However, more recent work in the area of

---

[88]A. Avramescu, "Modelling Scientific Information Transfer," *International Forum on Information and Documentation* 1 (1975): 13-9; and "Coherent Informational Energy and Entropy," *Journal of Documentation* 36 (1980): 293-312.

[89]Y.F. Le Coadic, "Information Systems and the Spread of Scientific Ideas," *R&D Management* 4 (1974): 97-111.

[90]Yves F. Le Coadic, "Modelling the Communication, Distribution, Transmission or Transfer of Scientific Information," *Journal of Information Science* 13 (1987): 143-8.

[91]D.J. Bartholomew, "Continuous Time Diffusion Models with Random Duration of Interest," *Journal of Mathematical Sociology* 4 (1976): 187-99.

telecommunication innovation demonstrates that the patterns of diffusion almost always take the form of modified logistic (S) curves, the differences being a function of the environment (presumably the subject area) and the target group and whether the speed of diffusion is slow or rapid.[92] Thus, it may be difficult to develop a mathematical diffusion model that will take the myriad possibilities into consideration at the same time.

2.1.3. Paradigm Shifts and Literature Growth

The expression *paradigm shift* owes its prominence to the influence of Thomas Kuhn and his *Structure of Scientific Revolutions* published in 1962 (revised 1970).[93] In the book he challenged the then prevalent view of the accumulative mechanism of scientific growth where development comes as a slow, piecemeal process. Instead, he suggested, progress follows the usual pace of 'normal science', occasionally interrupted by a revolutionary process where an accepted set of views (a paradigm) is challenged by a new set of data and observations and undergoes a reformation. During the revolutionary period the scientific community is divided into various camps of theories and explanations. There is an increased level of activity (cogitation), the different camps lack communication and a full understanding of each

---

[92]Christopher J. Easingwood and Simon O. Lunn, "Diffusion Paths in a High-Tech Environment: Clusters and Commonalities," *R&D Management* 22 (1992): 69-80.

[93]Kuhn, *The Structure of Scientific Revolutions*, 1970.

other's work, and thus they are said to be incommensurable.[94] During this time it is expected that the literature will show an increased level of growth and that some of this growth will be coming from influential work that obtains a large number of citations.

It should be pointed out that during times of *normal science*, the stage of steady accumulation of knowledge, one would expect to find a steady cohort of people working in the specialty domain. It is when a breakthrough occurs and a significant piece of work is published that many more, including outsiders from other specialties, will be attracted to it. This is when the literature is expected to show concomitant growth and exhibit spikes on a time series chart.

The implications of Kuhn's ideas on information science, on whether paradigm shifts at the micro-level are reflected in the literature, were tested by several workers. Moravcsik and Murugesan examined the implications of paradigm shifts on citation patterns in two cases of the physics literature - superconductivity and non-conservation of parity.[95] They were able to substantiate Kuhn's claims but also added that the idea of a simple paradigm shift was too unsophisticated to account for all the citation patterns observed.

---

[94]Thomas S. Kuhn, "Second Thoughts on Paradigms," in *The Structure of Scientific Theories* ed. Frederick Suppe (Urbana, IL: University of Illinois Press, 1974), 589-612.

[95]Michael J. Moravcsik and P. Murugesan, "Citation Patterns in Scientific Revolutions," *Scientometrics* 1 (1979): 161-9.

Diana Crane explored how the elements of a paradigm shift, as defined by Kuhn, fitted past experience in theoretical high energy physics.[96] She found the presence of the different elements but not always in constant conjunction with one another (as posited by the idea of cogitation), meaning that established and successful fields are not always characterized by normal science. She also added that rather than following Kuhn's theories of revolutionary growth, some fields came closer to Holton's idea of branching in science and growth by leapfrogging.[97] Similarly, Nadel found that during times of theory competition the degree of incommensurability varies over time and that Kuhn is only partially right.[98]

2.1.4. <u>The Growth and Development of Scientific Specialties</u>

There are several works bearing on the sociology of scientific specialties.[99] Kuhn's *The Structure of Scientific Revolutions* models revolutionary changes in the sciences with his idea of new perspectives (paradigms) bringing about sudden shifts

---

[96]Diana Crane, "An Exploratory Study of Kuhnian Paradigms in Theoretical High Energy Physics," *Social Studies of Science* 10 (1980): 23-54.

[97]Gerald Holton, "Models for Understanding the Growth of Research," *Daedalus* (1962): 94-131.

[98]Edward Nadel, "Commitment and Co-citation: An Indicator of Incommensurability in Patterns of Formal Communication," *Social Studies of Science* 13 (1983): 255-82.

[99]Kuhn, *The Structure of Scientific Revolutions*, 1970; Price, *Science since Babylon*, 1961; Belver C. Griffith and Nicholas C. Mullins, "Coherent Social Groups in Scientific Change," *Science* 177 (1972): 959-64; and Jonathan R. Cole and Harriet Zuckerman, "The Emergence of a Scientific Specialty: The Self-exemplifying Case of the Sociology of Science," in *The Idea of Social Structure: Essays in Honor of R.K. Merton* ed. Louis A. Coser (New York: Harcourt-Brace-Jovanovich, 1975), 139-74.

from *normal science* to *revolutionary science*. Derek Price, with his *Science since Babylon* has popularized the notion of the invisible colleges, the informal network of researchers at the core of a specialty, and its influence on the growth of a scientific specialty. Griffith and Mullins have pointed to the importance of *coherent social groups* in the development and success of a scientific discipline. Cole and Zuckerman outline three aspects of focus in the development of specialties: growth in authorship and literature, cognitive development, and the development of organizational infrastructures. Though variable in their extent and influence, all three are interrelated and play a role in studies of specializations.

Meadows and O'Connor have approached the subject from an information science point of view and examined the respective publications at different *growth points* in the development of a specialty.[100] They start their work with a simple question:

> Is [it possible], purely from a statistical analysis of scientific research papers, to identify the appearance of a new growth area and, if so, how soon after its first appearance can such an area be identified?[101]

and conclude with a wish:

> It would obviously be of value if a growth area could itself be discovered purely by a statistical analysis of the literature in the general field.[102]

---

[100]A.J. Meadows and J.G. O'Connor, "Bibliographical Statistics as a Guide to Growth Points in Science," *Science Studies* 1 (1971): 95-9.

[101]Ibid., 95.

[102]Ibid., 99.

This thesis provides a novel methodology toward the fulfilment of this wish.

In addition to these, there are several sociologists who have outlined and defined detailed stages in the growth of a scientific specialty. Diana Crane, in her well-known *Invisible Colleges*, proposed a four-stage model of the development of knowledge.[103] The sociologist Nicholas Mullins described the four stages in the development of a *theory group*.[104] Mulkay, Gilbert, and Woolgar have outlined a three-stage model.[105] Yet another four-stage classification of the life cycle of scientific specialties comes from Marc De Mey.[106] Whatever the approach:

> It has become clear that an understanding of how new research areas come into being is central to the sociological study of scientific development.[107]

The detailed explication of the stages is outside the focus of this thesis and will not be attempted here. However, for the purpose of providing a synthesis of the different approaches in the context of this thesis, here is a unified model of the different versions of specialty development:

---

[103]Diana Crane, *Invisible Colleges: Diffusion of Knowledge in Scientific Communities* (Chicago: University of Chicago Press, 1972).

[104]Nicholas C. Mullins, "Model for the Development of Sociological Theories," in *Theories and Theory Groups in Contemporary American Sociology* ed. N.C. Mullins (New York: Harper & Row, 1973), 17-35.

[105]Michael J. Mulkay, G.N. Gilbert and Steve Woolgar, "Problem Areas and Research Networks in Science," *Sociology* 9 (1975): 187-203.

[106]De Mey, *The Cognitive Paradigm*, 150.

[107]Mulkay, Gilbert and Woolgar, *Problem Areas and Networks*, 187.

1. **Birth of a notion:** the announcement of a discovery in the literature. It is usually generated by an individual and transmitted very rapidly. If the work is influential, it brings about a flurry of activity.

2. **Spread:** The scientific community reacts to the announcement by an increase in participation, in collaboration and in the diversity of publication sources. There is also an increase in the number of citations obtained by the first work. The number of self-citations decreases, at first high due to a vacuum created by the first work. This is accompanied by a shortening half-life in citations.

3. **Widespread acceptance**: As the field grows, it gains new adherents and starts fragmenting into subspecialties. There is also an increase in efforts toward popularization.

4. **Codification and institutionalization**: At this stage, there is an increased reorganization of the new knowledge into a coherent whole. The field is absorbed into the knowledge base of society.

Information epidemics come to being during the first two stages. Knowledge epidemics reflect the second and third stages.

There is an additional impetus in the case of fast growing literatures: the influential work that brings on a burst of activity refocuses people and catalyses a coordinated effort to look at things from a new point of view (angle). Those who accept this new point of view form a separate grouping that then (if successful) easily breaks off to form a new subspecialty grouping. All this is expected to be reflected in the literature. The burst of excitement remains an isolated incident and remains an information epidemic. However, the one that is followed by several spikes shortly thereafter and clearly reflects a significant rise in publication activity, may end up producing a knowledge epidemic that then may pass from the birth of a notion to its spread and even to its widespread acceptance. While an information epidemic remains a burst in the past, a knowledge epidemic would be expected to give rise to a new specialty or subspecialty.

To restate the points mentioned in the introduction above, the purpose of this thesis is to:

1. characterize the sudden short-term growth patterns,
2. establish the prevalence of information epidemics, and then
3. characterize the literatures of information epidemics.

The literature cited provides adequate support for this undertaking.

2.2. Methodological background

2.2.1. Publication Counts

The cornerstone of the methodology used in this study is composed of publication counts. Although popularized by Derek Price in the 1960s, bibliometric studies of growth based on publication counts have provided the standard approach since earlier this century.[108] However, the idea of quantifying knowledge growth as publication growth goes back to Machlup.[109]

On the one hand, publication counts can be regarded as an unrefined measure because of the complexities involving the communication system in science and all the filtering process that goes on before submitting work to print. On the other hand, their simplicity and ease of comparison with past activities make them impossible to ignore. Especially in physics, publication is the primary product of physicists' work and is at the heart of the reward system.

Although counting publications is simple and relatively straightforward, the interpretation of the data can create difficulties, and it is these difficulties that have in

---

[108]Price, *Science since Babylon*, 1961; F.J. Cole and N.B. Eales, "The History of Comparative Anatomy," *Science Progress* 11 (1917): 578-96; E. Wyndham Hulme, *Statistical Bibliography in Relation to the Growth of Modern Civilization* (London: Butler & Tanner, 1923); P.L.K. Gross and E.M. Gross, "College Libraries and Chemical Education," *Science* 66 (1927): 385-9; P.W. Wilson and E.B. Fred, "The Growth Curve of a Scientific Literature: Nitrogen Fixation by Plants," *Scientific Monthly* 41 (1935): 240-50; and Derek J. de Solla Price, "Measuring the Size of Science," *Proceedings of the Israel Academy of the Sciences and the Humanities* 4 (1969): 98-111.

[109]Fritz Machlup, *The Production and Distribution of Knowledge in The United States* (Princeton, NJ: Princeton University Press, 1962), 182.

the past led to severe criticisms of bibliometric methodology.[110] The main ones concern the problems of the least publishable unit (LPU), disciplinary variance, quality of work and variance in journal quality.[111] The problems with LPUs are serious because they inflate counts by fragmenting the presentation of data and increasing the number of co-authorships.[112] Particularly with the evaluation or comparison of small publishing units (such as individuals or research groups) the problems can become very acute.

There is, on the other hand, sufficient evidence to defend the use of raw publication counts. Several studies have concluded that there is a high correlation between quality and quantity. In other words, an increasing numbers of publications is parallelled by high quality work. Given that the issue is at the heart of the methodology employed in this thesis, some of these studies are worth mentioning.

In the 1950s Clark found a high correlation between a person's publication volume and his/her eminence in the field of psychology.[113] Eminence was measured as a function of his/her having held high offices in the American Psychological

---

[110]David Edge, "Quantitative Indicators of Communication in Science: A Critical Review," *History of Science* 17 (1979): 102-34; Michael J. Moravcsik, "Measures of Scientific Growth," *Research Policy* 2 (1973): 266-75.

[111]William J. Broad, "The Publishing Game: Getting More for Less," *Science* 211 (1981): 1137-9; Frances Anderson, *New Approaches to Research Policy Using Bibliometrics* (St. Foy, Québec: Conseil de la science et de la technologie, 1987), 24; and Susan E. Cozzens, *Literature Based Data in Research Evaluation: A Manager's Guide to Bibliometrics* (London: Science Policy Support Group, 1990).

[112]Broad, *The Publishing Game*, 1137.

[113]K.E. Clark, *America's Psychologists: A Survey of a Growing Profession* (Washington, D.C.: American Psychological Association, 1957), 46-56.

Association, the National Academy of Sciences of the United States or their own institutions, their presence in *Who's Who in America* and *American Men of Science*, and citations in *Psychological Abstracts* and *Annual Reviews of Psychology*. Derek Price in his *Little Science Big Science* claimed that:

> Flagrant violations there may be, but on the whole there is, whether we like it or not, a reasonably good correlation between the eminence of a scientist and his productivity of papers. It takes persistence and perseverance to be a good scientist, and these are frequently reflected in a sustained production of scholarly writing.[114]

Harriet Zuckerman showed that Nobel laureates (whose work is decidedly outstanding) publish more than other scientists at every stage of their working lives.[115] Cole and Cole reported a high positive relationship between quality and quantity for 120 university physicists.[116] Recently, Stephen Cole supported this finding:

> We would expect that scientists who have produced the most work in the past and whose work has been the most frequently cited would be the most eminent scientists in their fields. We would also expect that

---

[114]Price, *Little Science Big Science*, 40.

[115]H.A. Zuckerman, "Nobel laureates in Science: Patterns of Productivity, Collaboration, and Authorship," *American Sociological Review* 32 (1967): 391-403.

[116]Stephen Cole and Jonathan R. Cole, "Scientific Output and Recognition: A Study in the Operation of the Reward System in Science," *American Sociological Review* 32 (1967): 377-90; Jonathan R. Cole and Stephen Cole, "Measuring the Quality of Sociological Research: Problems in the Use of the Science Citation Index," *American Sociologist* 6 (1971): 23-9.

> these scientists should get higher ratings than scientists who have produced fewer papers and have been less frequently cited.[117]

Similarly, Lawani found a high correlation between the number of papers published and the citations obtained by the authors in Nigerian entomological literature.[118] Lawani also found that *in cancer research, the quantitative productivity of a country is positively correlated with the quality of her productivity.*[119] Stahl and Steger found a high correlation between innovation and productivity in 154 U.S. Air Force scientists and engineers.[120] They argued that innovation is a measure of quality and one of original and useful contribution, whereas productivity is one of quantity or output, without regard to innovativeness.[121] This nevertheless needs to be qualified. Martin and Irvine found that although the correlation between quality and quantity for any given person over a short period of time is tenuous, over longer time periods and for larger functional units (such as research groups, departments or institutes) there is some correlation between peer judgements and the quantity of

---

[117]Stephen Cole, *Making Science: Between Nature and Society* (Cambridge, MA: Harvard University Press, 1992), 142.

[118]Stephen M. Lawani, "Citation Analysis and the Quality of Scientific Productivity," *BioScience* 27 (1977): 26-31.

[119]Stephen M. Lawani, "On the Relationship Between Quantity and Quality of a Country's Research Productivity," *Journal of Information Science* 5 (1982): 143-5.

[120]Michael J. Stahl and Joseph A. Steger, "Measuring Innovation and Productivity - a Peer Rating Approach," *Research Management* 20 (1977): 35-8.

[121]Ibid., 36.

output.[122] Wallace and Bonzi provided evidence for the hypothesis that journals that constitute the nucleus in a Bradford distribution will be more frequently cited than non-nucleus journals.[123] Thus, nucleus journals which publish the majority of studies in a given area were shown to be of superior quality. Stephan and Levin showed that prolific scientists write in the more prestigious journals and that they do not *trade quality for quantity by publishing in journals which have a lower impact.*[124] Recently, Frumau found *a high correlation between the number of US patents, the number of international published articles (as found in the INSPEC database) and R&D expenditures.*[125] Given that patents are an accepted measure of quality, the finding also supports the link between quality and quantity.

In addition, there is a high degree of correlation between publication counts and other measures of scientific excellence such as funding and peer ranking. McAllister and Narin found a 0.95 correlation *between the amount of NIH (National Institute of Health) funds received and the number of biomedical publications from*

---

[122]B.R. Martin and J. Irvine, "Assessing Basic Research: Some Partial Indicators of Scientific Progress in Radio Astronomy," *Research Policy* 12 (1983): 61-90.

[123]Danny P. Wallace and Susan Bonzi, "The Relationship Between Journal Productivity and Quality" In *ASIS '85: Proceedings of the Forty-Eighth ASIS Annual Meeting, Las Vegas, 20-24 October 1985* ed. C.A. Parkhurst (White Plains, NY: Knowledge Industry Publications, 1985), 193-6.

[124]Paula E. Stephan and Sharon G. Levin, "Inequality in Scientific Performance: Adjustment for Attribution and Journal Impact," *Social Studies of Science* 21 (1991): 351-68.

[125]Coen C.F. Frumau, "Choices in R&D and Business Portfolio in the Electronics Industry: What the Bibliometric Data Show," *Research Policy* 21 (1992): 97-124.

[120] medical schools.[126] There is also a high degree of correlation between a university's size (by the number of papers it produces) and the citation quality of its publications (influence per paper).[127] Given this evidence that has been building consistently over the last four decades, one can state unequivocally that there is in fact a close parallel between quality and quantity in publication.

The best way to avoid criticisms of publication counts is to refrain from comparing disciplines (such as mathematics versus biological sciences) and to use reliable publication sources (e.g. major international indexes such as *Physics Abstracts*). It should be possible to choose groups publishing papers in the same universe of journals. These could be works produced in the same specialty or in similar specialties that tend to publish around a narrow group of journals.

In fact, the use of other science indicators is even more faulty: patents mostly apply to technology and manufacturing, reliable funding and manpower statistics (such as the number of PhDs graduating) are difficult to obtain and have a big time lag, and very often researchers obtain support from many different sources at the same time.

---

[126]Paul R. McAllister and Francis Narin, "Characterization of the Research Papers of U.S. Medical Schools," *Journal of the American Society for Information Science* 34 (1983): 123-31.

[127]Richard C. Anderson, Francis Narin, and Paul McAllister, "Publication Ratings versus Peer Ratings of Universities," *Journal of the American Society for Information Science* 29 (1978): 91-103.

Indicators based on funding present an additional circular argument in that to work scientists need money but to get money they must demonstrate adequate work.[128]

The use of publication counts in this study sidesteps the above mentioned problems by limiting itself to one field only, physics, and to using publication counts to follow short-term changes in a given specialty. In that respect, changes in publication numbers are driven by the dynamic characteristics of the fields of study themselves. The statistics are obtained for the work of many groups of authors, not for individuals. Thus, the influence of any one person with an unusual publication activity is minimized.

The current availability of a large number of bibliographic databases has considerably eased the acquisition of publication information. Several recent studies have established an adequate methodology for online bibliometric studies.[129] Lancaster and Lee used online databases to track the growth of acid rain literature and

---

[128]S. Cohn, "The Effect of Funding Changes Upon the Rate of Knowledge Growth in Algebraic Topology, 1955-75," *Social Studies of Science* 16 (1986): 23-59.

[129]Donald T. Hawkins, "Unconventional Uses of On-line Information Retrieval Systems: On-line Bibliometric Studies," *Journal of the American Society for Information Science* 28 (1977): 13-8; H.F. Moed, "The Use of Online Databases for Bibliometric Analysis," in *Informetrics 87/88: Select Proceedings of the First International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval, Diepenbeek, Belgium, 25-28 August 1987* ed. Leo Egghe and Ronald Rousseau (Amsterdam: Elsevier, 1988), 133-46; O. Persson, "Measuring Scientific Output by Online Techniques," in *Handbook of Quantitative Studies in Science and Technology* ed. A.F.J. Van Raan (Amsterdam: North-Holland, 1988), 229-52; and H. Dou, P. Hassanaly and L. Quoniam, "Infographic Analytical Tools for Decision Makers: Analysis of the Research Production of the Sciences," *Scientometrics* 17 (1989): 61-70.

to follow the diffusion of the topic through databases of various kinds.[130] The same methodology can be used to uncover new specialties in a given field of science.

It is always possible that an indexing service for some reason or another (death of an indexer, holidays, job changes) could publish materials from several issues of a journal at the same time, thus artificially inflating publication counts. It is also not unusual to find duplicate citations during the course of an online search. However, some verifications during the course of this study have shown that such events are unusual in *Physics Abstracts*, amounting to about one percent. This number, divided among 58 chapters of *Physics Abstracts* over 11 years comes to about 1.3 duplicate items per month - a number that can be safely disregarded.

Changes in fast growing literatures and information epidemics occur too rapidly for any one new journal or special issue to affect the picture significantly. It is outside the potential of any one serial or monographic publication to bring about a change of such a magnitude unless one is dealing with a topic or subclass where the monthly productivity is so low (1-5 items) that the publication of a special issue will make a big difference. One such possibility is with Chapter 96 and the sections dealing with the planets of the solar system. The average productivity in them is so low that when an occasional special issue discusses results from a recent satellite flyby (such as Jupiter and the Voyager expeditions), the numbers abstracted produce an outlier, even though there is no information epidemic taking place.

---

[130]F.W. Lancaster and Ja-Lih Lee, "Bibliometric Techniques Applied to Issues Management: A Case Study," *Journal of the American Society for Information Science* 36 (1985): 389-97.

2.2.2. The Importance of the Time Factor in Studies of Growth

The very notion of growth is synonymous with change, and change can only be understood in dynamic terms, with the help of concepts and models that make time an explicit variable. While the literature on growth is voluminous, there are comparatively few studies in information science in which the time element is taken as the key variable. For most, it is peripheral to the focus of interest and is expressed in periods of years. The most recent review of the time factor dates back to 1977.[131] A review by Zunde briefly summarizes equations in information science, among them some involving time-related growth phenomena.[132]

So far no one has addressed the problem of an adequate metric for time and what the proper period of duration should be in scientometric studies. The general attitude has been that it is easier to work with annual publication data and that such a large time period somehow removes concerns about the submission date of articles versus the cover date of publications versus the indexing lag time of the data sources.[133]

The model governing most studies of growth is the law of exponential growth $P(t) = ke^{at}$ where P(t) is the total number of publications at time t, k is the initial size

---

[131]A. Neelameghan, "Expressions of Time in Information Science and Their Implications: An Overview," *Annals of Library Science and Documentation* 24 (1977): 13-33.

[132]Pranas Zunde, *On Empirical Foundations of Information Science* (Atlanta, GA: School of Information and Computer Science, 1981), NTIS, PB82-125998.

[133]Henry Small, conversation with author, 2 July 1993.

at time zero and a is the growth rate.[134] However, the equation is a static one because it assumes a constant rate of growth, and, given enough time, the number of publications would grow to unrealistically high numbers.

With regard to the growth of scientific fields over time, the notion of *research fronts* was introduced by Derek Price to describe the dynamic aspects of literature growth.[135] However, time was not mentioned distinctly. Another well-known study that is lacking the time variable is Goffman's epidemic model mentioned above.[136] Griffith and Small demonstrated the existence of maps of scientific literatures in terms of literatures central to a subject area and subject clusters with papers in the periphery.[137] Instructive as they were, the maps were developed by static techniques, with no explanations as to causes and no place for a time factor. Griffith and Mullins' *coherent groups* suffer from the same lack.[138]

Recently, growth processes in various subfields have been receiving attention from a variety of workers, both theoretical and applied.[139] The majority of the

---

[134]Tague et al, *The Law of Exponential Growth*, 125.

[135]Price, *Networks of Scientific Papers*, 510.

[136]Goffman, *Generalization of Epidemic Theory*, 225.

[137]Belver C. Griffith et al., "The Structure of Scientific Literatures II: Toward a Macro- and Microstructure for Science," *Science Studies* 4 (1974): 339-65.

[138]Griffith and Mullins, *Coherent Social Groups*, 959.

[139]S.D. Haitun, "Stationary Scientometric Distributions, Part II: Non-Gaussian Nature of Scientific Activities," *Scientometrics* 4 (1982): 89-94; H.F. Moed, et al. "The Application of Bibliometric Indicators: Important Field- and Time-dependent Factors to be Considered," *Scientometrics* 8 (1985): 177-203; S.H. Sichel, "The GIGP distribution Model with Applications to Physics Literature," *Czechoslovak Journal of Physics B* 36

theoretical studies have involved probabilistic distributions whereby growth occurs by chance phenomena, randomly, and, with the exception of Burrell, the time element is considered to be linear. Bruckner adopted a theoretical evolutionary model to demonstrate growth, but the applications of the simulation models to growth trends in subfields remain to be shown empirically. Discipline-oriented studies have typically adopted a yearly approach to counting publications. The majority of the disciplines under study have indeed shown relatively slow and steady rates of growth that can safely be represented in yearly terms. However, even papers dealing with recent fast moving literatures have adopted yearly counts for their studies and as such have missed much of the dynamics available from monthly illustrations.[140]

(1986), 133-7; M. Kunz, "Time Spectra of Patent Information," *Scientometrics* 11(1987): 163-73; A.P. Trofimenko, "Scientometric Analysis of the Development of Nuclear Physics during the Last 50 Years," *Scientometrics* 11(1987): 231-50; Quentin L. Burrell, "Predictive Aspects of some Bibliometric Processes," in *Informetrics 87/88: Select Proceedings of the First International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval, Diepenbeek, Belgium, 25-28 August 1987* ed. Leo Egghe and Ronald Rousseau (Amsterdam: Elsevier, 1988), 43-63; E. Bruckner, W. Ebeling and A. Scharnhorst, "The Application of Evolution Models in Scientometrics," *Scientometrics* 18 (1990): 21-41; and Henry Small, "Macro-level Changes in the Structure of Co-citation Clusters: 1983-1989," *Scientometrics* 26 (1993): 5-20.

[140]M.P. Rebrova and V.V. Komarov, "Some Aspects of the Scientometric Analysis of the Development of Research in the Area of Superconductivity," *Nauchno-Tekhnicheskaya Informatsiya, Ser. 1* 16, no.8 (1989): 23-7; N.M. Builova, E.K. Zakharova and N.V. Akshinskaya, "How Information Publications Reflect Studies on High-Temperature Superconductivity," *Nauchno-Tekhnicheskaya Informatsiya, Ser. 1* 17, no. 8 (1990): 22-8; and Hurt and Budd, *Modelling the Literature of Superstring Theory*, 475.

The sole study of monthly (as opposed to yearly) changes in a subfield remains that by Sullivan and co-workers.[141] They followed the month-by-month changes in the development of electroweak theory by means of co-citation studies and were able to demonstrate the evolution of thinking and the relative success of opposing theories as reflected in the co-citation of star publications. The most probable reason why their studies have not been replicated is that one needs to establish a specialized database at the outset, and that their type of studies are very expensive to put together.

With the exception of Sullivan and his co-workers, the approaches mentioned above can be seen as reflections of an *equilibrium process* and represent the dynamic, moving and growing world of communication (publications) only in very crude terms. They better portray slow moving closed-system models and do not at all accord when it comes to the dynamic world of fast moving literatures. This thesis emanates from the idea that the growth of knowledge is a dynamical process and that time-dependent phenomena should be explained in terms of time-dependent models befitting the characteristics and magnitudes of the phenomena at hand. The advantage of observing fast growing literatures in a short-term and time-dependent manner is that changes can be tracked on a short-term basis, understanding and insight can be obtained quickly, and the results can in turn be used in a timely fashion.

[141] Daniel Sullivan et al. *Understanding Rapid Theoretical Change*, 309.

2.2.3. Converging Indicators

Given some of the above-mentioned criticisms of using publication counts alone to get a feel for the evolution of a specialty, it would be valuable to use supporting indicators to add certainty to the interpretation of the findings. Most of the literature has used single indicators to operationalize research variables. In other words, each variable until recently has been measured by a single indicator at a time.

Decreasing budgets for basic research, the concentration of resources in a few central facilities and continuing doubts about the effectiveness and objectivity of the peer review system have in the 1980s given rise to the search for new methods of evaluating research performance.[142] One of the more influential teams, Ben Martin and John Irvine at the University of Sussex, has developed a set of *converging partial indicators* that are based on the number of publications and citations obtained as a portion of the group total for various research groups and major research centers.[143] Martin and Irvine have concentrated on large research centers and major facilities (such as the Centre européen de recherche nucleaire (CERN) or astronomical observatories) rather than on specialties and have combined peer review with bibliometric measures of productivity and impact. As such, they have developed measures to tell how good the relative performances of different research centers have been.

---

[142]Daryl E. Chubin and Edward J. Hackett, *Peerless Science: Peer Review and U.S. Science Policy* (Albany, NY: SUNY Press, 1990).

[143]Ben R. Martin and John Irvine, "Assessing Basic Research: Some Partial Indicators of Scientific Progress in Radio Astronomy," *Research Policy* 12 (1983): 61-90.

The methodology adopted in this thesis is an extension of Martin & Irvine's work and is also based on the principle of triangulation used in sociology and communication work. Triangulation basically refers to the comparison of several types of data gathered about a single social phenomenon:[144]

> The triangulation approach allows an investigator to obtain a more integrated view of the communication that occurs within interpersonal networks in science, and of how such communications leads to the development of new knowledge.[145]
>
> Once a proposition has been confirmed by two or more independent measurement processes, the uncertainty of its interpretation is greatly reduced. The most persuasive evidence comes through a triangulation of measurement processes. If a proposition can survive the onslaught of a series of imperfect measures, with all their irrelevant error, confidence should be placed in it.[146]

Here, if an outlier on a time series chart reflects significant growth of publication activity, that increase in output can then be correlated with an influential work that many cite and the number of authors participating in the growth of the spike. If all three indicators are pointing in the same direction, that is the increase in output is parallelled by an increase in the number of authors and a concentration of citations to the same influential work(s), then that outlier can be called an information epidemic and signifies an important growth point in that scientific specialty.

---

[144]E.T. Webb et al., eds. *Nonreactive Measures in the Social Sciences*, 2nd ed. (Boston: Houghton Mifflin Company, 1981).

[145]Leah A. Lievrouw, "Triangulation as a Research Strategy for Identifying Invisible Colleges Among Biomedical Scientists," *Social Networks* 9 (1987): 217-48.

[146]Webb et al., *Nonreactive Measures*, 35.

Thus, the influence of a paper can be quantified by three factors:

1. **Output:** an increased number of papers following the publication of the influential work.

2. **Spread:** an increased number of research groups or institutions working on the same specialty area and citing the influential paper.[147] In the initial stages of a new growth area only a few research groups are likely to participate in the work. However, as the field grows and new scientists are attracted to the area, the proportion of the new literature due to any one group will decrease and the diversity of the groups will be expected to increase.[148]

In their discussion of the emergence of the sociology of science, Cole and Zuckerman note:

> Although a growing literature may indicate increased scholarly effort, it is not necessarily evidence for a shared intellectual focus among those at work in the specialty. There are however other reasons to think that such a focus was emerging: among them a growing consensus on the usefulness of particular publications, a consolidating research front in which new papers built directly upon those just published, and increased rates of collaborative publication.[149]

In the context of this study, the productivity of individual scientists or countries is not that important; it is more relevant to follow the productivity of

---

[147]Henri Dou et al. "La Physique a Aix-Marseille: Indicateurs et évolution à partir de l'analyse automatique de la base de données INSPEC," *Revue française de bibliométrie* 3 (1988): 49-85; A.F.J. Van Raan, "Evaluation of Research Groups," in *The Evaluation of Scientific Research* ed. D. Evered (NY: Wiley, 1989), 169-87.

[148]Meadows & O'Connor, *Bibliographical Statistics*, 97.

[149]Cole and Zuckerman, *The Emergence of a Scientific Specialty*, 146.

specific groups of scientists. According to Moed and Van Raan, it is groups of researchers that form the basis of progress and are the most interesting and most policy relevant entities to evaluate.[150] Different groups will tend to work on different problems. However, when one finds several dissimilar groups working and publishing on the same problem, this convergence confers the field an added importance, especially when compared to single efforts coming from here and there.

Finally, as a field evolves, it draws more cooperation and the number of co-authorships increases concomitantly. Indeed, Builova and her co-workers found that when high-temperature superconductivity work became widespread, the number of co-authorships, when compared to previous work in superconductivity, also increased dramatically.[151] Although co-authorship will not be examined in this work, it remains a valid measure of spread.

3. **Impact**: an increased number of citations obtained by the influential work immediately after publication.[152] Citation frequency is an accepted indicator of the importance of a work as judged by those working in the same field. The usefulness of a certain publication can be traced by its citation patterns whereas increased rates of

---

[150]H.F. Moed and A.F.J. Van Raan, "Critical Remarks on Irvine and Martin's Methodology for Evaluating Scientific Performance," *Social Studies of Science* 15 (1985): 539-47.

[151]Builova et al., *How Information Publications Reflect Studies on High Temperature Superconductivity*, 56.

[152]Julie Virgo, "A Statistical Procedure for Evaluating the Importance of Scientific Papers," *Library Quarterly* 47 (1977): 257-67; Eugene Garfield, "Is Citation Analysis a Legitimate Evaluation Tool?" *Scientometrics* 1 (1979): 359-75; and *Citation Indexing* (New York: Wiley, 1979).

collaboration can be traced by counting authors. In the beginning of a new specialty area, there will be a higher proportion of self-citations (due to the small number of workers in the field) which tapers off as the field gets crowded. The intensity of spread and impact will reflect the fact that a certain cohesion exists in the field under study and that workers are not turning out trivial and disconnected works.

Cole and Zuckerman continue to discuss the emergence of a cognitive consensus:

> A growing consensus among specialists on the usefulness of certain publications is a prime indicator that a specialty is developing distinctive problematics and thus a cognitive identity. The extent of convergence of citations to particular papers and to the work of particular authors is a rough measure of such consensus. ... Converging citations do not mean that all agree on the significance of cited research or that all highly cited authors have a common orientation but only that the cited work is influential in some respect.[153]

The only way to get proof of the degree and orientation of influence would be to interview every author and ask why they cited a particular piece of work. If a large majority gives the same (or similar) reasons, then that would be considered conclusive. Short of that, all this work is built on evidence.

The contention here is that if a strong burst of publications is accompanied by a growth in the number of authors and if a large number of them cite the same publications, then that constitutes sufficient evidence that a consensus exists on the influence of a given work or group of works and its (their) influence in giving rise to an information epidemic. Citation behaviour is somewhat akin to territorial claims

---

[153]Cole and Zuckerman, *The Emergence of a Scientific Specialty*, 146.

between competing species in ecology: each author or group of authors tries to expand its territory at the expense of the other's. In the work undertaken here, this is the concept of *acceptance*, measured by citation counts and the number of research groups working on the side of one approach to problem solving or another. In that respect, bibliometric methodology depicts the ongoing struggle and provides valuable insights into the conflicts.

All the foregoing is not intended to dismiss potential problems with citations and citation analysis. They have been well summarized before.[154] One set of criticisms comes from the perceived weaknesses of the *Science Citation Index* itself. Its coverage, especially of non-English literatures, is limited, its citation indexing covers first authors only, and there are large numbers of errors in the data with respect to spelling, publication names, and citations. The second set of criticisms deals with authors' citation practices. It is not always clear why an author cites a work. Not all citations are conferred in a positive light. Some works are cited all so often as to lose their relative merit; some others, due to the obliteration phenomenon, are never cited. The problem of the least publishable unit is liable to inflate citation

---

[154]Garfield, *Citation Indexing*, 240-4; Linda Smith, "Citation Analysis," *Library Trends* 30 (Summer 1981): 83-106; Diana Hicks and Dave Crouch, "Can Bibliometrics Measure Up?" *Physics World* 3 (September 1990): 27-8; and Sidney J. Pierce, "Disciplinary Work and Interdisciplinary Areas: Sociology and Bibliometrics," in *Scholarly Communication and Bibliometrics* ed. Christine L. Borgman (Newbury Park, CA: Sage, 1990), 47-58.

counts for some authors.[155] Others go to secondary sources rather than citing original documents.

The use of citations in this thesis bypasses most of these criticisms in the sense that the search here is for a large number of citations (often in the hundreds) to a work or a group of works that have influenced and inspired a large number of other workers due to their seminal nature. The criticisms above are largely valid, and citation analysis must be approached with considerable prudence. On the other hand, if an author or a work obtains hundreds of citations within a short period of time (such as a year or two), then that author or work must have been awarded with considerable endorsement, approbation and recognition no matter what the problems may be with citation analysis. As well, citation comparisons here remain within well-defined disciplines and do not stray to comparisons between disciplines. Thus, the approach taken in this thesis is defensible on the basis of a large volume alone.

The possible damage from the absence of non-English material in databases can be addressed by examining coverage of non-English materials in the INSPEC database. Given that *Physics Abstracts* covers the literature of physics worldwide, one would expect the percentage of works in any one language to respect the worldwide productivity in physics in that very language. In other words, *Physics Abstracts* rather well reflects worldwide activity in physics. Thus, an analysis of the language fields of the *Physics Abstracts* portion of the INSPEC database should rather well reflect the worldwide contribution of different languages in the physics literature. Of the

[155] Broad, *Publishing Game*, 1137.

approximately three million items in the database, 90.3% (2.7 million) are written in English, 3.8% in Russian, 1.5% in French and 1.4% in German. Thus, 97% of the physics literature analyzed is written in four of the most popular Western languages. While the absence of certain non-English items may be serious by themselves, overall, for the purposes of this study, the gaps are of no consequence.

2.3. Restatement of the problem

The fundamental question addressed in this thesis is: Are spikes seen on time series charts reflective of fast growing literatures? The issue is how an influential work affects the subsequent growth of its discipline and whether any measures can be developed to decide whether the popularity of an influential work has resulted in a significant change in direction for that specialty or whether it has been of passing interest.

The evidence for the usefulness of the approach in this study can be summarized as follows:

A. Theoretical:

1. Growth: mechanisms depend on the time period in the study and the field at hand,
2. Epidemics and diffusion: rapid growth comes as a result of the transmission and spread of an exciting idea,
3. Paradigm shifts: revolutionary science gives rise to a renewed vigour in a given field, and

4. Growth of specialties: a fast growing literature reflects a quickly developing area of study.

B. Methodological:

1. Publication counts: they provide the most convenient way to measure growth,

2. Time factor: fast moving and dynamic processes must be reflected in micro-level and short-term time series, and

3. Converging indicators: they give unequivocal corroboration for the existence of a growth area and its direction.

2.4. Hypotheses

Based on the above, the following hypotheses will be tested:

1. **Sudden growth patterns termed information epidemics are widespread in the growth of the physics literature.**

2. **Outliers observed during the growth of a field are caused by influential papers**. They are not aberrations.

3. Within the observed outliers:

3a. **Changes in the growth patterns are reflected in the number of works published, the duration of the growth and the number of research groups active in it.** Influential papers lead to increased activity in their specialty areas. Their influence can be temporary, in terms of months, or else they can considered to be revolutionary works, changing the future direction of their research fields and generating a considerable growth of publications. In either case, they would be expected to have obtained a high number of citations from a large number of research groups of diverse origins. However, this confers no judgement on the quality of the works under consideration.

3b. **The increased activity in a given specialty area is reflected in the increased proportion of conference papers abstracted** (when compared to journal articles). A higher activity in a given field will lead to increased communication and thus to a larger number of meetings and the publication of a larger number of conference proceedings.

2.5. Assumptions

Despite the evidence brought forward above, a number of assumptions need to be made to streamline this study. First, growth patterns in the formal literature of a scientific field fully parallel the advances and the conceptual developments of that field.[156] This premise has in fact been accepted for a number of years. Must of current research in scientometrics is largely based on the use of publications to measure the growth of a field. In addition, the growth of the individual parts (such as

---

[156]Price, *Science since Babylon*, 97; and "Measuring the Size of Science," *Proceedings of the Israel Academy of Sciences and Humanities* 4 (1969): 98-111.

chapters or sections) of *Physics Abstracts* is due to the growth of of their literatures and not due to editorial or selective indexing policies. On the one hand, this is experientially self-evident and needs not be stated explicitly. On the other hand, given that it is an important point and that it has not been tested specifically, it needs to be included as an assumption.

It doesn't matter whether the mechanism of growth of science follows an exponential or linear pattern. What is certain is that it fluctuates in time with the activity in a given field. The growth can at times be linear, at times exponential or even at times negative. The more dramatic the fluctuations, the higher the activity in a given field and the greater the likelihood that it will be subject to information epidemics.

It is possible to develop a set of scientometric measures with the purpose of identifying fields of intensive growth. These include not only publication counts but also citation patterns, authorship and co-authorship patterns as well as co-citation maps. The results can be based on work from information science or can be based on sociological evidence (such as the growth of social networks).

Citations are an acceptable measure of a work's influence on the advancement of knowledge. While no status as to quality is inferred, a work's importance and influence is reflected in its citation patterns.

There is no indexer effect; all keywords and classifications are assigned evenly and correctly. Even though *Physics Abstracts* adopted a new classification scheme in 1977, this scheme represented only a modification of a system used for much of the twentieth century and not a new system *ab initio*. Therefore, it is safe to assume that, in order to absorb the new classification system into *Physics Abstracts*, indexers will have needed to make only slight modifications to their practice

All journals are equally available and are equally timely in arrival. In practice, it is well known that this is not necessarily so. European and North American journals would certainly be more timely in arrival than some of their Third World counterparts. On the other hand, given the preponderance of European and North American, and especially English-language, journals in the physics literature, the variations introduced by the different arrival times of a small percentage of journals is negligible.

Finally, with respect to the language breakdown of physics publications (see pages 60-61 above), it is assumed that *Physics Abstracts* perfectly reflects the world literature in physics. The important distinction is that, even if 90.3% of publications processed by the index are in English and the rest is in various other languages, this doesn't mean that the language breakdown for *world output* will be the same. Due to various inclusion policies, certain non-English physics materials might not make it into *Physics Abstracts*. The assumption adopted here is that even if a difference exists, it is negligible and can be ignored.

2.6. Concluding thoughts

Finally, there is the question of whether highly cited authors start information (and then knowledge) epidemics by themselves or even occasionally over their career.

Can one trace the connections to highly cited people, journals, institutions as well as hot topics? But if one decides to pursue hot topics, that becomes a circular argument. How does one systematically find a hot topic in the first place? Popular magazines carry articles on seemingly exciting discoveries but the selection of the topics is arbitrary. With the exception of the Institute for Scientific Information and its publication *Science Watch*, that question has not been dealt with in the literature.[157] The purpose of *Science Watch* is to list the most cited 10 articles in different disciplines (such as the physical or biological sciences or engineering) and draw attention to their authors and the relative standings of the subspecialties:

> essentially, holding up a mirror to the scientific literature in order to reflect what the scientific community itself is signalling as noteworthy.[158]

The way the growth of literature has been tracked so far has not enabled the discovery of information epidemics.

Despite their simplicity, publication numbers provide the most rapid and practical indicators available to signal growth. One cannot trace funding alone: superconductivity was developed with very cheap materials. One cannot trace researchers and authors alone, because at times great ideas come from small laboratories with limited staff and money. Co-citation maps are expensive to chart and require voluminous data that may not be available or may be beyond many

---

[157]Henry Small and David Pendlebury, "Introducing Science Watch: A New Kind of Science Intelligence," *Science Watch* 1 (January 1990): 2.

[158]Ibid., 2.

researchers' financial means to obtain. As well, one needs a considerable lag time before undertaking a co-citation study. On the other hand, the triangulation methodology mentioned above provides a clear, direct method to obtain essentially the same information at less cost and effort. The three simple measures of publication counts, authorship and citation counts are simple to obtain, simple to interpret, and when used together provide powerful indicators of intensive areas of growth in scientific specialties.

In the main, there are two ways to identify epidemics. One is locally (within a given field) where a given area of interest is growing much more rapidly than neighbouring areas. This is the approach adopted here. The other is with "hot papers", that is by following highly cited works (as mentioned above with *Science Watch*) under the assumption that a series of hot papers makes a hot field - an occurrence not demonstrated so far.

It should be pointed out that hot fields are not necessarily the most active or productive fields in a given scientific specialty. The difference is that the way ISI determines *hot fields* is contingent on a large number of citations obtained by a group of papers in the same year in which they are published and their ranking compared to other groups of papers from other specialties.[159] Thus, presumably, a highly populated field with a high output would nevertheless be absent from the rankings if its publications were not being cited substantially in the same year that they were published. One could have an information epidemic with a high publication output

---

[159]*Science Watch*, Sample issue, [January 1990], 3.

without owing its inspiration to a recently published paper but a publication that was, for example, late in obtaining recognition.

It should also be pointed out that in ISI's parlance "hot" means very current and highly cited, but not necessarily belonging to the most productive specialty, whereas by the approach adopted in this thesis "hot" means very active, irrespective of the immediacy of the work(s) influencing the increased activity. Presumably, a very active scientific field will also be a very productive one, one capable of exhibiting an information epidemic and able to garner a large number of citations. This contention has yet to be tested systematically.

It is easy to agree that highly cited papers have to be central to an information or knowledge epidemic in the first place, but how one recognizes those works and goes about systematically identifying them in a given field of science depends on one's focus. Since the word *epidemic* conveys the idea of rapid spread, an information epidemic here means an event showing rapid growth of output, irrespective of the validity of the ideas (initially, at least), their success or citation impact. Ultimately, from an information science point of view and from the main interest in this thesis, the focus is on high output; thus one wants to know why that comes about. Is it from influential articles or is it accidental? Is it due to an indexing problem (since the data are obtained from a major index) or is there a different reason?

Useful as their works have been, none of the writers mentioned in this literature review have considered the implications of how the publication of an influential work alters the immediate (short-term) patterns of output in the literature.

If a work is deemed important and highly relevant, it should be expected to diffuse rapidly. Under exceptional circumstances it is possible to obtain epidemic spread. It thus becomes necessary to look at important works and find out whether the spread of their influence has been epidemic. Are those epidemics reflected in publication rates? Not every highly cited article or even citation classic necessarily gives rise to epidemics.[160] But how can one know if a work has had a major influence on the growth of a field in the first place? The number of citations it has received is one way of assessing it, but that only reflects its reception by others, not necessarily how it has affected the growth of the field. Therefore, the most reasonable course of action remains a survey of the published literature in a given field and the search for epidemics and the influential works that have brought on those epidemics. This is the purpose of this thesis.

---

[160]Daryl E. Chubin, A.L. Porter, and F.A. Rossini, "'Citation Classics' Analysis: An Approach to Characterizing Interdisciplinary Research" *Journal of the American Society for Information Science* 35 (1984): 360-8.

CHAPTER 3

SIGNIFICANCE

Developments in certain scientific specialties have given rise to an explosive growth of their literatures within a very short time span. The growth and success of these fields (such as superconductivity) have been well publicized. In the area of physics at least, the rise of some of these is attributable to the influence of one or a small number of important and highly cited papers. The question is whether these sudden explosive growth patterns are commonplace in physics or whether they are truly special and unusual occurrences.

What turns an *information epidemic* into a *knowledge epidemic* and contributes to knowledge growth is the enduring value of the information in the publications characterizing it. It is the nature of that endurance that is at the heart of this thesis. The attributes that make a work lasting and valuable, such as its ability to influence other work in its field, its dispersion and adoption by other workers as well as its use (citations) by them is the object of analysis. Stated differently, the purpose is to discover whether the transfer of knowledge from a given research front into core knowledge is accompanied by dramatic fluctuations.

Finding a large number of epidemics will mean that there are fields in physics that surge very quickly and that there are indeed growth points in science. If so, the identification of growth points may help shape strategy in a series of areas from science policy to library acquisition decisions. In addition, they may provide a focus for further historical and sociological studies.

The absence of epidemics in the data will mean that science, physics at least, grows slowly and routinely and that the sudden growth patterns that contributed to the development of this research question are indeed unique phenomena. They are temporary disturbances that quickly dissipate, after which the regular equilibrium is reestablished. Despite temporary and exceptional anomalies, "normal science" (in the sense of Kuhn) still dominates recent work in physics. If so, the evidence forces one to accept that science moves ahead slowly (irrespective of the underlying mechanism), as if by an invisible hand, as claimed by Michael Polanyi in the 1960s.[161]

In terms of using publicly available information in a systematic manner, the approach promoted in this thesis represents several advantages. Admittedly, the use of publication numbers alone is too coarse an indicator to use for individual evaluation purposes, but so far as the tracking of short-term growth phenomena in scientific fields is concerned, it is still the simplest measure.

In view of this discussion, this study hopes to make the following contributions:

1. **Establishing the usefulness of micro-level studies.**

A short-term approach to tracing information production and dissemination has clearly become not only possible but also necessary. There is currently a very

---

[161]Michael Polanyi, "The Republic of Science: Its Political and Economic Theory," *Minerva* 1 (1962): 54-73.

large amount of information published in a variety of scientific fields. As an example, *Physics Abstracts* published 79,830 abstracts in 1970. Ten years later in 1980 it published 109,577 (up 37.3%), and in 1990 this number increased to 157,051 (up 43.3%). Over the twenty years between 1970 and 1990 the physics literature, as abstracted by *Physics Abstracts*, doubled. The same range of increases has been true for chemistry, geology, biology and computer science, among others. All this volume provides the amount of data necessary for undertaking detailed short-term studies that are lacking in scientometrics. The level of detail provided by the systematic compilation of monthly publication numbers is simply not available from other studies. Future updates of the data obtained from this study are also certain to provide a rich source of data for further scientometric work.

2. **Determining the prevalence of epidemics in physics.**

   Goffman brought the epidemic concept into information science, but no one knows how common or widespread epidemics are. The results of this thesis will indicate how common epidemics are in the physics literature.

3. **Contribution toward a theory of fast growing literatures.**

   The establishment of a theory of fast growing literatures is still needed. This work is an initial response to Budd and Hurt's call for such action.[162] The mechanisms of the underlying patterns of short-term rapid growth are still not clear. More work is needed to examine the influence or usefulness of other

---

[162]Budd and Hurt, *Superstring Theory*, 97.

bibliometric indicators such as journals with high impact factors and literature doubling times. To quote Budd and Hurt:

For certain types of seminal literature, the generic models of scientific literature transfer do not hold. There appears to be a class of literature which ignores the standard models and presents a much steeper slope in terms of citation frequency than would otherwise be expected. The salient characteristic of this literature is the speed with which it is recognized and used by those in the scientific enterprise. We suggest that modelling these literatures will be much more descriptive of information transfer in science than will modelling the generic citation frequency of a paper at random.[163]

The results of this thesis will provide empirical support for this contention. It will state clearly how large this "special class of literature" is in the area of physics.

4. **Establishing information epidemics as a useful concept.**

Helped along by epidemic models, an understanding of fast growth phenomena in dynamic terms provides a new level of analysing the factors driving anomalous states of science literature - a paradigm that is lacking in the literature of information science. In addition, characterizing dynamic phenomena in a time-dependent manner confers on the discussion a richness that is not available from static descriptions.

5. **Contributing to the development of more finely tuned science indicators.**

Short-term studies of the state of science should in turn bring in the capability of more finely tuned and timely science indicators. Tracing disciplines with

---

[163]Ibid., 97.

short-term indicators would allow for a better understanding of the current state of science by quickly identifying areas of rapid growth and thus better fine tune science policy decisions. ISI's *Science Watch* already accomplishes this to a certain extent with its lists of *hot papers*.

At a finer level of trying to stay current with the growth of scientific disciplines and in comparing them one to another, no comparable approach exists so far. Sullivan et al.'s monthly co-citation technique[164] was based on a specially constructed database; it would be too cumbersome and expensive to adapt it to a whole series of disciplines. ISI's *Science Watch* is useful but only provides a list of lists of recent hot papers, and its source of information belongs to a private database. The Science Citation Index does provide monthly (and now even weekly) data, but because of a lack of a detailed classification system it is not possible to use it systematically. The availability of research fronts in it is useful so long as a given front is known a priori. However, the fronts are very unstable and lack the continuity necessary for such studies. They change their contents and their names as well as their numbering from year to year so as to make the tracking of a specialty very difficult, if not impossible, over several years.

---

[164]Sullivan et al. *Understanding Rapid Theoretical Change in Particle Physics*, 276.

It is possible to study the evolution of fields and ideas through co-citation studies which do provide reliable indicators.[165] However, these belong to a more retrospective historical approach. It would be faster and simpler to first pinpoint areas of growth through information epidemics after which more elaborate scientometric techniques could be used to obtain deeper insights. A complete survey of physics as attempted in this study would point to important case studies to work on.

---

[165]Henry Small and Eugene Garfield, "The Geography of Science: Disciplinary and National Mappings," *Journal of Information Science* 11 (1985): 147-59; H. Small and E. Greenlee, "Collagen Research in the 1970s," *Scientometrics* 10 (1986): 95-117; and H. Small, "Macro-level Changes in the Structure of Co-Citation Clusters: 1983-1989," *Scientometrics* 26 (1993): 5-20.

CHAPTER 4

METHODOLOGY

4.1. Physics Abstracts:

Physics Abstracts is one of the largest abstracting and indexing services in the world. It has been published since 1903 by the INSPEC (Information Services for the Physics and Engineering Communities) system which is the Information Division of the Institution of Electrical Engineers (IEE) in Great Britain. In 1941 the Science Abstracts were divided into three parts: Part A: *Physics Abstracts*, part B: *Electrical and Electronics Abstracts* and part C: *Computer and Control Abstracts*. It is Part A that is used as a source of data for this work.

All three sections of INSPEC's *Science Abstracts* are organized along a classification system:

> The purpose of [the] classification is to arrange documents by subject according to a logical scheme which reflects the way in which workers in the subject view the field, its subdivisions, and its relationships with other subjects. The principal use of the INSPEC classification codes and the purpose for which the classification schemes were first developed is to arrange the entries in subject order in the printed publications.[166]

Physics Abstracts was initially divided into six major topical divisions: general physics, light, heat, sound, electricity & magnetism, and chemical physics. In the 1940s the classification was enlarged to eight divisions: general physics, astronomy &

---

[166]Institution of Electrical Engineers. *INSPEC User Manual*. (London: Institution of Electrical Engineers, 1983), 3.1.

astrophysics, geophysics, radioactivity, radiation, heat, acoustics, and electricity & magnetism. The 1962 classification was enlarged to encompass fifteen divisions. In 1969 the divisions became subdivided into chapters, for a total of twenty-three divisions and 197 chapters. In 1973 the system was enlarged, the numbering system was changed and a number of new chapters and sections were adopted in order to better reflect major developments in physics. In 1977 INSPEC embraced the *International Classification System for Physics* of the International Council of Scientific Unions (ICSU). The scheme carried ten major divisions subdivided into sixty-one chapters and several hundred sections. This scheme was also adopted by *Bulletin signalétique* in France, *Physikalische Berichte* in Germany and *Referativnyi Zhurnal* in Russia.

The classification has since kept its structure of divisions and chapters stable, except for some revisions in 1988. In 1985 Section D, *Information Technology* was added but the experiment lasted only two years and was abandoned. The most recent classification for physics is listed in Table 1 below, the same scheme that remained unchanged between 1977 and 1987 which is the focus of this thesis. The stability in the classification scheme after 1977 constitutes a significant influence on the sample selected in this research.

Table 1
Physics Abstracts Main Chapters

No.	TITLE
A00	GENERAL
A01	Communication, education, history, and philosophy
A02	Mathematical methods in physics
A03	Classical and quantum physics; mechanics and fields
A04	Relativity and gravitation
A05	Statistical physics and thermodynamics
A06	Measurement science, general laboratory techniques, and instrumentation systems
A07	Specific instrumentation and techniques of general use in physics
A10	THE PHYSICS OF ELEMENTARY PARTICLES AND FIELDS
A11	General theory of fields and particles
A12	Specific theories and interaction models; particle systematics
A13	Specific reactions and phenomenology
A14	Properties of specific particles and resonances
A20	NUCLEAR PHYSICS
A21	Nuclear structure
A23	Radioactivity and electromagnetic transitions
A24	Nuclear reactions and scattering: general
A25	Nuclear reactions and scattering: specific reactions
A27	Properties of specific nuclei listed by mass ranges
A28	Nuclear engineering and nuclear power studies
A29	Experimental methods and instrumentation for elementary-particle and nuclear physics
A30	ATOMIC AND MOLECULAR PHYSICS
A31	Theory of atoms and molecules
A32	Atomic spectra and interactions with photons
A33	Molecular spectra and interactions with photons
A34	Atomic and molecular collision processes and interactions
A35	Properties of atoms and molecules; instruments and techniques
A36	Studies of special atoms and molecules
A40	CLASSICAL AREAS OF PHENOMENOLOGY
A41	Electricity and magnetism; fields and charged particles
A42	Optics
A43	Acoustics
A44	Heat flow, thermal and thermodynamics processes
A46	Mechanics, elasticity, rheology
A47	Fluid dynamics
A50	FLUIDS, PLASMAS AND ELECTRIC DISCHARGES
A51	Kinetic & transport theory of fluids; physical properties of gases
A52	The physics of plasmas and electric discharges

Table 1--Continued

No.	TITLE
A60	CONDENSED MATTER: STRUCTURE, THERMAL AND MECHANICAL PROPERTIES
A61	Structure of liquids and solids; crystallography
A62	Mechanical and acoustic properties of condensed matter
A63	Lattice dynamics and crystal statistics
A64	Equations of state, phase equilibria, & phase transitions
A65	Thermal properties of condensed matter
A66	Transport properties of condensed matter (nonelectronic)
A67	Quantum fluids and solids; liquid and solid helium
A68	Surfaces and interfaces; thin films and whiskers
A70	CONDENSED MATTER: ELECTRONIC STRUCTURE, ELECTRICAL, MAGNETIC, AND OPTICAL PROPERTIES
A71	Electron states
A72	Electronic transport in condensed matter
A73	Electronic structure and electrical properties of surfaces, interfaces, and thin films
A74	Superconductivity
A75	Magnetic properties and materials
A76	Magnetic resonances and relaxation in condensed matter; Mossbauer effect
A77	Dielectric properties and materials
A78	Optical properties and condensed matter spectroscopy and other interactions of matter with particles and radiation
A79	Electron and ion emission by liquids and solids; impact phenomena
A80	CROSS-DISCIPLINARY PHYSICS AND RELATED AREAS OF SCIENCE AND TECHNOLOGY
A81	Materials science
A82	Physical chemistry
A86	Energy research and environmental science
A87	Biophysics, medical physics, and biomedical engineering
A90	GEOPHYSICS, ASTRONOMY AND ASTROPHYSICS
A91	Solid earth geophysics
A92	Hydrospheric and lower atmospheric physics
A93	Geophysical observations, instrumentation, and techniques
A94	Aeronomy, space physics and cosmic rays
A95	Fundamental astronomy and astrophysics, instrumentation and techniques and astronomical observations
A96	Solar system
A97	Stars
A98	Stellar systems; galactic and extragalactic objects and systems; Universe

Physics Abstracts is now published twice a month, and includes 5500-6000 abstracts per issue. It indexes and abstracts over four-thousand publications annually. In 1992 INSPEC indexed 4257 titles. Journal articles normally make up about eighty percent of the total items abstracted, conferences comprise fifteen percent, and books, monographs, technical reports and dissertations make up the other five percent of the annual coverage.[167] While INSPEC normally indexes the majority of journals selectively, there are a number of journals considered important enough that they are abstracted completely. In 1992 there were 776 such journals.[168] There is a cumulative index every six months to subjects, authors, conference proceedings, and monographs. The coverage for online versions of the index goes as far back as 1969 and is made available by most online vendors worldwide. In 1990 a CD-ROM version was launched, starting with data from 1989.

Physics Abstracts represents the worldwide research effort in physics comprehensively, and evidence corroborates this contention. For example, Chang and Dieks compared the Dutch effort in physics to several countries in the world and found that *Physics Abstracts* is indeed suitable for such purposes.[169] Similarly, data from *Physics Abstracts* has been used extensively to make national comparisons in

---

[167]Ibid., 2.3-2.3.1.

[168]Institution of Electrical Engineers. *INSPEC List of Journals and Other Serial Sources, 1991/2* (London: Institution of Electrical Engineers, 1991), 349.

[169]Hans Chang and Dennis Dieks, "The Dutch Output of Publications in Physics," *Research Policy* 5 (1976): 380-96.

output.[170] Studies that use other indexes such as the *Science Citation Index* (SCI) often compare its physics coverage to that of *Physics Abstracts* as a standard to aspire to. For example, Narin and Carpenter found the correlation in physics between the two to be 0.97.[171] A recent Australian report pegs the overlap in coverage for physics and physical sciences at 95-100%.[172]

4.2. Classification codes:

Approaches using classifications are particularly helpful because the classification systems remain fairly constant over a long time whereas the alternatives, words and expressions, change relatively quickly.[173] In addition, classification numbers are assigned to publications more uniformly, irrespective of the terminology in use. One can even construct networks (maps) of co-classification fields and analyze

---

[170]Jan Vlachy, "Publication Output of World Physics," *Czechoslovak Journal of Physics B* 29 (1979): 475-80; Erno Bujdoso and Tibor Braun, "Publication Indicators of Relative Research Efforts in Physics Subfields," *Journal of the American Society for Information Science* 34 (1983): 150-5; and Radosvet Todorov, "Distribution of Physics Literature," *Scientometrics* 7 (1985): 195-209.

[171]Francis Narin and Mark P. Carpenter, "National Publication and Citation Comparisons," *Journal of the American Society for Information Science* 26 (1975): 80-93.

[172]National Board of Employment, Education and Training. *Quantitative Indicators of Australian Research* (Canberra: Australian Government Publishing Service, 1994), 248-9.

[173]A.F.J. Van Raan and R.J.W. Tijssen, "Numerical Methods for Information on Aspects of Science: Scientometric Analysis and Mapping," *Perspectives in Information Management* 2 (1990): 203-28.

the dynamics of given subfields.[174] The advantages of classification-based studies have been well stated by Todorov in his work on co-classification analysis and are listed here:

- All documents, irrespective of their source and language of publication, are used in the analysis.
- Documents relevant to a given (sub)field are not randomly selected.
- Although there could emerge some problems of classification changes, the dynamics of links could be studied.
- Classification codes are easy to access and to process.
- Content of codes is explicit and, therefore, the observed links are more open to discussion.
- Time lag is smaller as compared to citation appearance.
- Maps of national research in a given (sub)field or of journal publication profiles could be produced by taking an appropriate subset of documents.[175]

4.3. Data collection:

Data on the growth of the literature of physics were obtained from the printed version of *Physics Abstracts*. The data were obtained from printed sources for two

---

[174]A.F.J. Van Raan and H.P.F. Peters, "Dynamics of a Scientific Field Analyzed by Co-subfield Structures," *Scientometrics* 15 (1989): 607-20.

[175]R. Todorov, "Co-classification Analysis for Science Mapping: An Example from Superconductivity," in *Science and Technology Indicators* ed. A.F.J. Van Raan et al. (Leiden: DSWO Press, 1988), 261-70.

major reasons. First, monthly update field codes that would allow online retrieval are not available from any online vendor, including the Dialog system that was used for online retrieval. The database files were reorganized in 1990, and update codes for the database prior to 1987 were eliminated. Second, and more important, abstracted articles often are given several classification numbers even if they are predominantly published in one class in the printed indexes. There are often cross-references to relevant items in other classes. With the classification numbers obtained from a computer search it is not possible to differentiate between the primary class number in which the abstract was published and secondary ones that are cross-references and are given to enhance retrieval.

The number of abstracts published in each chapter of *Physics Abstracts* was collected from each semimonthly issue published between January 1977 and December 1987 (inclusively). The year 1977 is the first year the internationally accepted classification scheme went into effect. The year 1987 is the last year before any changes to the scheme were made. Between those years the classification scheme of *Physics Abstracts* remained stable. Therefore, any changes in the growth patterns in this study must be ascribed to the literature and not to changes in classification. Otherwise, any partial changes to the classification scheme would have to be explained in relation to other chapters that did not change as well as in relation to outliers, should they be present in the data.

The data for each *Physics Abstracts* chapter were pooled into monthly data and were entered into a separate Lotus 123® spreadsheet for each chapter. The advantage here is that this cumulation eliminates some noise but does not change the regression line R^2.

Three of the sixty-one chapters of *Physics Abstracts* listed above have been excluded from analysis because they did not provide sufficient data: Chapter 86 began in 1979 on energy research and the environment. Due to its later starting date and to its peripheral subject nature, it has been omitted from consideration. The coverage for chapters 14 (Properties of specific particles and resonances) and 27 (Properties of specific nuclei listed by mass ranges) has been very sparse: together they include less than a dozen items per year and are therefore not worthy of inclusion. These chapters, when first introduced, may have been active areas of research but nowadays they represent a project largely completed and no longer current.

At the end, the final data pool consisted of fifty-nine files, one for each *Physics Abstracts* chapter and containing 11 years of monthly data (making up fifty-eight) and one for the total monthly output. It is these files of monthly data that were used as the base for analysis. The number of items counted reaches some 1.3 million abstracts in the aggregate. The time series charts for the chapters of *Physics Abstracts* are shown in Appendix A.

4.4. Representation:

Given that each chapter exhibited a multitude of spikes of differing magnitudes, it became necessary to adopt a systematic and statistically consistent technique for determining outliers. The personal computer version (SPSS-PC®, version

4) of the SPSS® program was chosen for this purpose. All the data were put into one large Lotus 123® spreadsheet with months by rows and chapter numbers by columns. The spreadsheet consisted of 132 rows for 132 months (for eleven years) and sixty-two columns (years, months, cumulative monthly data and fifty-eight chapters).

A regression analysis was conducted with SPSS-PC® on each chapter's data to obtain descriptive statistics, histogram frequencies, regression parameters, residuals, outliers and the autocorrelation function. The data obtained from the SPSS-PC® analysis were exported back to the Lotus 123® spreadsheet program in order to fit an envelope of plus and minus three standard deviations around the regression line and to graphically illustrate the outliers. Given the unsatisfactory graphic capabilities of the SPSS-PC® program at hand, it was preferable to use the spreadsheet program. These spreadsheets, in fact, provide the data for all the figures in this thesis.

The reason for choosing a threshold of 3.0 standard deviations is that 3.0 standard deviations account for 99.74% of the area under a normal curve.[176] The chance of finding any point on either extreme of 3.0 standard deviations is only 0.26%. In other words, any point falling on either extreme of a 3.0 standard deviation envelope is considered to be truly exceptional and worthy of being judged an outlier.[177]

---

[176]Hubert M. Blalock, *Social Statistics*, rev. 2d ed. (New York: McGraw-Hill, 1979), 602.

[177]The criterion for inclusion here is a z score equivalent to 1/2n, where n = no. of cases. With 132 cases in the data, 1/264 = 0.0037878. To get the area under the standard normal curve, subtract that from 1.0 and divide by 2 (to be able to read the statistical tables). Then, (1 - 0.0037878) / 2 = 0.4981, which is equivalent to 2.9

4.5. Criteria for analysis:

The list of outliers by chapter is given below in Table 2 (page 117). Each outlier carries a number designation. The listing also indicates the month and year for the occurrence of the outlier as well as the source publication(s) (be it a journal, be it a conference proceeding) that contributed significantly. The only threshold imposed a priori was that a publication source would be listed as being significant for an outlier so long as it accounted for more than ten percent of the outlier's contents. In other words, only those sources that contributed more than ten percent to an outlier's volume are listed in Table 2 (page 117).

As the table shows, there are altogether eighty-one outliers obtained from the analysis. However, not all outliers are meaningful. If a given *Physics Abstracts* chapter did not show any growth between 1977 and 1987, there was no point in analyzing an outlier further once it became known that the principal portion of its contents were due to one or two conference proceedings. For example, such is the case for Chapter 67 (Quantum Fluids and Solids) where the three outliers are composed of conference articles by 81% to 89%. If a chapter did show growth accompanied by one or two outliers, but those outliers were due to a conference proceeding or a special journal issue, again, the chapter was omitted from consideration. For example, Chapter 91 (Solid Earth Geophysics) presents such a

---

standard deviations. Thus, the 3.0 standard deviations criterion is more than adequate to establish outliers.

case: it showed significant growth, but its single outlier is composed of one special journal issue and a conference proceeding.

If a given chapter exhibited a number of outliers but these were interspersed regularly in time, they too were omitted from analysis so long as the chapter did not reveal any other outliers that were exceptional. Chapters 29 (Experimental Methods for Elementary Particles), Chapter 67 (Quantum Fluids and Solids) and, Chapter 75 (Magnetic Properties and Materials) belong to this category. Chapter 29 contains 5 outliers, each separated by two years, because the outliers were due to the proceedings of the biennial Particle Accelerator conferences (Appendix A, Figure A17). Chapter 67 contains three outliers due to the proceedings of the International Conference on Low Temperature Physics held every three years (Appendix A, Figure A38). Chapter 75 also contains three outliers due to the proceedings of International Conference on Magnetism held every three years (Appendix A, Figure A44).

An additional reason for removal from analysis was the lack of outliers despite interesting looking dynamics. In other words, if a given *Physics Abstracts* chapter exhibited major ups and downs in its data with occasional spikes, but none of them were outliers as such (i.e. pierced the 3.0 standard deviation line), then that chapter was eliminated from consideration. If the moving average (as determined by its autocorrelation function for the time series) reflected a major up and/or down pattern but the data did not contain any outliers, then that chapter was eliminated. For example, Chapter 65 (Thermal Properties of Condensed Matter) does not reveal any outliers, but has nevertheless interesting dynamics between 1979 and 1983.

It is possible that a new set of standards within a 2.5 or even 2.0 standard deviation envelope would have produced more outliers and consequently a different list of *Physics Abstracts* chapters for further analysis. Such a test was run for 2.0, 2.5, 3.0, 4.0 and 5.0 standard deviation lines, and the number and the characteristics of the outliers were noted down. Table B2 and B3 of Appendix B list a comparison of the results. The conclusion was that while more outliers presented themselves for consideration, they did not add to the number of chapters that would have been analyzed under the 3.0 standard deviation rule. A larger list of outliers only meant more noise but not a larger number of case studies to adopt. The number of cases presented by the 2.0 and 2.5 standard deviation envelopes were qualitatively not any more significant. Those obtained from 4.0 and 5.0 standard deviation envelopes kept some of the more important outliers (such as Chapter 2 (Mathematical Physics)) but also eliminated others that were finally chosen for further analysis. Thus, the 3.0 standard deviation norm was found satisfactory on additional counts, and it was decided therefore to adhere to it.

Overall, then, given the volume of data obtained from *Physics Abstracts*, the following criteria were adopted to choose data appropriate for detailed analysis:

1. The chapter must show dramatic or sudden growth patterns.
2. The regression statistics must be significant.
3. The growth portion must include data that pierce the upper regression envelope, that is exhibit outliers at higher than 3.0 standard deviations.

4. The moving average must reflect a jump in growth. Due to the volume and noise in the data, it is not always possible to recognize significant patterns at first sight.[178]

By these criteria only four chapters were retained for analysis:

- Chapter 2: Mathematical methods in physics;
- Chapter 36: Studies of special atoms and molecules;
- Chapter 73: Electronic structure and electrical properties of surfaces, interfaces, and thin films;
- Chapter 74: Superconductivity.

Each one includes significant growth over 11 years, includes outliers, and its moving average reveals sudden jumps in its growth patterns.

4.6 Data preparation:

All the abstracts contained within an outlier were searched on *INSPEC* online (Dialog file 2), and their abstracts numbers, titles and sources of publication were identified and downloaded. Each record was in turn identified on *Science Citation Index* online (Dialog file 434) and downloaded in format 4 (F4). Format 4 contains the full record of each abstract with tagged fields that facilitate further analysis. A typical record downloaded from the *Science Citation Index* with format 4 is shown below:

---

[178]A smoothed curve, the periodicity of which is determined according to the largest ACF (autocorrelation function) factor for the raw data, eliminates much of the shorter-term detail and at the same time dramatically reflects the suddenness of the jump, if present.

```
FN- Scisearch(R)_1974-1994/Mar W4
AN- 07519661¦
GA- D9600¦
TI- EVALUATION OF INTEGRALS INVOLVING POWERS OF (1-X2) AND 2
    ASSOCIATED LEGENDRE FUNCTIONS OR GEGENBAUER POLYNOMIALS¦
LA- ENGLISH¦
AU- RASHID MA¦
CS- AHMADU BELLO UNIV,DEPT MATH/ZARIA//NIGERIA/¦
GL- NIGERIA¦
JN- JOURNAL OF PHYSICS A-MATHEMATICAL AND GENERAL, 1986,
    V19, N13, P2505-2512¦
PY- 1986¦
DT- ARTICLE¦
NR- 4¦
SF- SciSearch; CC PHYS--Current Contents, Physical, Chemical
    & Earth Sciences¦
SC- PHYSICS¦
RF- 86-0109 001   (TIME-VARYING SYSTEMS VIA GENERALIZED
    ORTHOGONAL POLYNOMIALS; SHIFTED TSCHEBYSCHEFF SERIES;
    LINEAR DYNAMIC-SYSTEMS)
CR- BAILEY WN, 1964, GENERALISED HYPERGEO
    GRADSHTEYN IS, 1965, TABLES INTEGRALS SER
    LAURSEN ML, 1981, V14, P1065, J PHYS A
    ULLAH N, 1984, V25, P872, J MATH PHYS¦¦
```

The field tag designators for both *INSPEC* and *Science Citation Index* online are given in Appendix C.

There are, however, two complications that arise in data preparation here. One is that some records are incomplete: there is a field missing. The other is that the record is not available from *Science Citation Index* because it belongs to a conference proceeding that was never indexed by it.

In the first case, the most important omissions were the lack of an address field (CS-) or a "Citations Received" field (CR-). In the case of a missing address a "CS - NOT GIVEN" line was added to the record. In the case of missing citations a "CR - NONE GIVEN" line was added. For one, the programs to analyze the data depend on the presence of these fields in the records; otherwise they crash. As well, it would be useful to obtain statistics on the number of instances, in the total data, where fields were missing, and therefore useful information for further analysis was missing. As it turns out, the lack of the address field was a significant factor in one of the case studies presented in Chapter 6. The abstract number of the record from *Physics Abstracts* was also added. The reason for this was quality control: to ensure that all the records obtained from *Physics Abstracts* were eventually included in the files to be analyzed. A typical record then had the following look:

```
FN- Scisearch(R)_1974-1994/Mar W4
AN- 07480276¦
GA- D6679¦
TI- IMPROVING THE COMPUTATION ACCURACY BY APPLYING THE FAST
    FOURIER TRANSFORMATION ALGORITHM¦
LA- RUSSIAN¦
AU- SADYKOV IK¦
CS- NOT GIVEN¦
JN- IZVESTIYA VYSSHIKH UCHEBNYKH ZAVEDENII AVIATSIONAYA
    TEKHNIKA, 1986, N2, P58-62¦
PY- 1986¦
DT- ARTICLE¦
NR- 0¦
SF- CC ENGI--Current Contents, Engineering, Technology &
    Applied Sciences¦
CR- NONE GIVEN¦¦
```

There were less than a hundred records with no citations among the more than ten thousand records directly analyzed for this thesis (<1%). On the other hand, the lack of addresses was more prevalent, especially in the case of Russian and Japanese publications. The majority of Russian records did not carry an address field. In the case of Japanese publications, Japanese letters journals carry neither an address field nor a *Citations Received* field.

In the case of conference papers missing from the *Science Citation Index* database the massaging of data was more extensive. A typical full record (format 4) from *Physics Abstracts* has the following arrangement:

```
FN- DIALOG(R)File   2:INSPEC¦
CZ- (c) 1994 Institution of Electrical Engineers. All rts.
    reserv.¦
AZ- 02489431¦
AZ- <INSPEC> A85086321¦
TI- Stability of superconducting composites in a magnetic
    field¦
AU- Gray, K.E.¦
CS- MST Div., Argonne Nat. Lab., IL, USA¦
AU- <EDITOR> Collan, H.; Berglund, P.; Krusius, M.¦
SP- Finnish Min. Educ.; Helsinki Univ. Technol.; OY LM
    Ericsson; Fincoil-Teollisuus OY; Huure OY;
    Instrumentarium OY; Kone OY; et al¦
CP- UK¦
PG- 616-19¦
PY- 1984¦
CT- Proceedings of the Tenth International Cryogenic
    Engineering Conference¦
CL- Helsinki, Finland¦
CY- 31 July-3 Aug. 1984¦
PU- Butterworth Guildford, Surrey, UK¦
PG- xx+844¦
BN- 0 408 01257 9¦
DT- Conference Paper (PA)¦
LA- English¦
TC- Theoretical (T)¦
RF- 6¦
AB- The recent calculation of the stability criterion for
    ... [shortened for brevity] ...
    nevertheless important to the agreement.¦
DE- composite superconductors; critical current density
    (superconductivity); stability¦
ID- one dimensional model; critical current density;
    superconducting composites; magnetic field; stability
    criterion; flux flow; multifilamentary conductors;
    thermal conductivity; windings¦
CC- A7460J (Critical currents)¦¦
```

First, all the records for a given conference from *Physics Abstracts* were put into a separate file. The author (AU-) and address (CS-) fields were capitalized (all

Science Citation Index records are in capital letters). Author first names, available in full in *Physics Abstracts*, were initialized, up to three initials (as per *Science Citation Index*). Other reformatting included ensuring that the address started with the name of an institution and was followed by a departmental name, adding a slash (/) between the name of an institution and its department, and the addition of a semicolon between addresses. A *Citations Received* (CR-) field was also added and the references at the end of the conference papers (obtained locally or by interlibrary loan) were typed in *Science Citation Index* format. Fields not needed for identification nor analysis were removed in order to reduce the size of the files and to speed up analysis. These included the CP-, PG-, BN-, DT- as well as the AB-, DE- and ID- fields. For the sake of consistency with all other records, the abstract number, under a new field IN-, was also added to the record. Once ready for analysis, the record above finally looked as follows:

```
IN- A85086321
AN- 0000¦
FN- DIALOG(R)File   2:INSPEC¦
AZ- 02489431¦
AZ- <INSPEC> A85086321¦
TI- Stability of superconducting composites in a magnetic
    field¦
AU- GRAY KE¦
CS- ARGONNE NAT LAB/ MST DIV¦
PG- 616-19¦
PY- 1984¦
CT- Proceedings of the Tenth International Cryogenic
    Engineering Conference¦
CL- Helsinki, Finland¦
CY- 31 July-3 Aug. 1984¦
RF- 5¦
CR- GRAY KE, 1983, V13, P405, J PHYS F
    BROOM RF, 1960, V11, P292, BR J APPL PHYS
    SCOTT CA, 1982, V22, P577, CRYOGENICS
    MARTINELLI AP, 1972, PROC APPL SUPERCOND CONF
    MADDOCK BJ, 1969, V9, P261, CRYOGENICS¦¦
```

4.7. Analysis of outliers:

All the authors, corporate sources and citations that were contained in the records were identified with the help of two programs specially written for the purposes of this thesis.[179] The first (FILTER.EXE) identifies all the occurrences of a given author, corporate source or citation and makes a single-column list of them. In the case of citations, if the name in a citation matched that in the author field of the same record (self-citation), it was omitted. The reason is that while a self-citation may be important to its author, a work cannot be judged to represent a significant impact unless it is highly cited by others. The second program (REPORT.EXE) summarizes and concatenates the data obtained from FILTER.EXE, and presents the results in comma-delimited files. These can then imported into the Lotus 123® spreadsheet program into separate columns by name and frequency, and the results can be reranked by frequency.

All the data making up an outlier (sometimes a given outlier may span several months) were then cumulated and reranked by the number of citations received by authors. All those authors with less than 3 citations to their names were dropped from consideration. One or two citations to any work can be regarded as accidental or spurious whereas three, within a short period of time, must be examined with more attention. The advantage for partial cumulations came from the fact that monthly tables gave sparse and inconsistent results whereas cumulative tables consistently

---

[179]The two programs, called FILTER.EXE AND REPORT.EXE, were written by Larry McGoldrick of Concordia University in TurboPascal especially for this analysis.

listed most cited authors. The same procedures were followed to rank the most productive authors and the most productive corporate sources.

Once the most consistently and highly cited authors were identified, searches were conducted in the *Science Citation Index* online for their most cited papers. For this purpose, the command "E CR= ..." in Dialog was used. It lists all the cited works of an author in chronological order, along with the number of citations obtained. Statistics for all those papers cited more than 100 times among the top cited authors were retained and listed on a table. This represents an extraordinary threshold. Garfield reports that of all papers indexed in the *Science Citation Index* between 1945 and 1988 less than half a percent obtained more that 100 citations and less than 0.01% obtained more than 500 citations in their lifetimes.[180] A more recent but unpublished survey by Henry Small found that only 1% of all items in the ISI database between 1981 and 1994 had been cited 100 or more times.[181]

The result of all this was to obtain lists of highly cited authors and corporate sources with the purpose of identifying influential works as well as works that further

---

[180]Eugene Garfield, "The Most-Cited Papers of All Time, SCI 1945-1988. Part 1A. The SCI Top 100 - Will the Lowry Method Ever be Obliterated?" in *Essays of an Information Scientist* 13: 45-56.

[181]Henry Small, letter to author, 5 September 1995.
To be precise, the data sample in question is composed of *all published items covered by ISI in 1981 and cumulated citations to them over the period 1981-1994 (14 years). There were a total of 783,339 1981 items. ... 47% of items [were] uncited after 14 years, but 19% had been cited 10 or more times, and 1% had been cited 100 or more times after 14 years.*

contributed to knowledge epidemics, that is those that sustained a specialty area and promoted its growth well after the first influential paper had been published.

4.8. Limitations:

Since the data are obtained from the physics literature only, it may be difficult to generalize the conclusions of this study to all fields of science. Two of the primary impediments would be the uniqueness of the classification used for physics and the characteristics of the physics field itself. The primary communication vehicle for researchers in physics is the preprint; the journal article acts as an archival depository. Other fields where the journal publication is not so predominant (such as geology or engineering) or where the patterns of referencing are different (such as mathematics) may show different growth characteristics. However, a priori, there is no reason why collective excitement could not exist in any science or for that matter in any research endeavour.

The classification system used in this study is the same adopted by *Physics Abstracts*. It is a unique classification in the sense that it is valid only for physics and no other field. Although it has been adopted by the three other major physics indexes and has remained consistent for many years, there is no certainty that a different classification system based on different premises will give the same results. There are few indexes representing scientific fields with a classification scheme as detailed as that of *Physics Abstracts*. For example, the classification scheme used for *Chemical Abstracts* is relatively small. Each class comprises many subspecialties so as to make the technique unusable. On the other hand, the one field that deserves to be followed along with physics, the literature of molecular biology and biochemistry, where a large number of modern information epidemics presumably have occurred, is not organized along any classification scheme. A study following the growth of specialties in biochemistry or molecular biology would per force have to use title words or keywords, which would bring in a set of other serious constraints and would lose the type of simplicity demonstrated by this study. On the other hand, mathematics is one other potentially fruitful area for applying the methodology developed here. Although its research patterns are very different, and many of its specialties are still dominated by influential works that were published over a century ago, the amount of detail provided in its classification scheme renders *Mathematical Reviews* easily subject to the same analysis conducted in this thesis.

Another limitation with classification schemes in general is that they are academically oriented, that is they reflect the practice of science by researchers. On the other hand, a large and increasing quantity of recent research is interdisciplinary and problem-oriented, irrespective of academic disciplines. A strict adherence to classifications may miss the mark on some important discoveries. On the other hand, given that a complete survey of the 1977-1987 *Physics Abstracts* was undertaken, no matter whether a work is interdisciplinary or not, it will have to be classified somewhere, and if it gives rise to an information epidemic, sooner or later the swelling of output will reflect this.

This study has restricted itself to publication indicators alone and is not using sociological methodology. Most scientometric research needs to be accompanied by expert advice and interviews with workers in given fields. However, that would take the study in a different direction. The purpose here is to test for the prevalence of information epidemics as an exploratory study. If the test is successful and the hypotheses of this thesis are supported, future studies can follow certain cases in more depth. There are other input/output factors such as the analysis of research funding, the number of scientists active in a given field at a given time, the number of graduate students working in those same fields and the content analysis of influential works that could also be included in the examination of information epidemics. However, the focus in this thesis is the information epidemic as an output indicator and the mass of work being published as a result of scientific activity.

Finally, this research is unidirectional in the sense that it seeks to identify outliers and their prevalence. It does not test for the reverse effect of identifying seminal work by other means or from other authors and then testing to see whether they were followed by information or knowledge epidemics. That is certainly a valid test, but it is not the focus of this study, and will have to await further work in the future.

CHAPTER 5

DATA ANALYSIS: PHYSICS ABSTRACTS

5.1. The Growth of *Physics Abstracts*

Figure 1 illustrates the number of abstracts published per year from its inception until the end of 1994. The Abstracts was first published as part of *Science Abstracts* between 1898 to 1902 (vol.1-5). In 1903 it was split into two parts: *Section A: Physics* and *Section B: Electrical Engineering*. In 1941 it was renamed *Physics Abstracts* but retained as part of *Science Abstracts*, and it continues to be published under this title today.

Physics Abstracts has gone through three major phases of growth. From its start in 1903 until World War II the number of items in *Physics Abstracts* was largely stable and showed little growth, at an average rate of 2.2% per annum. Between 1948 and 1970 it grew much faster at a rate of 13.8% p.a. Since 1970 the growth has slowed down to 3.4% p.a.. For the period of time that is the focus of this study, that is 1977-1987, the literature has grown at a compounded 4.1% p.a. (Figure 2).

When compared to the previous 80 years, the last 20 years have displayed rather quiet and stable growth (Figure 3). In fact, this steadiness in yearly growth rates and numbers confers an added dimension of stability to the results obtained from this study. There were two major decreases in output, one in 1905, due to the huge rise the previous year, and one in 1940 (presumably due to World War II), when the number of items abstracted dropped by 31.4%. On the other hand, there were five

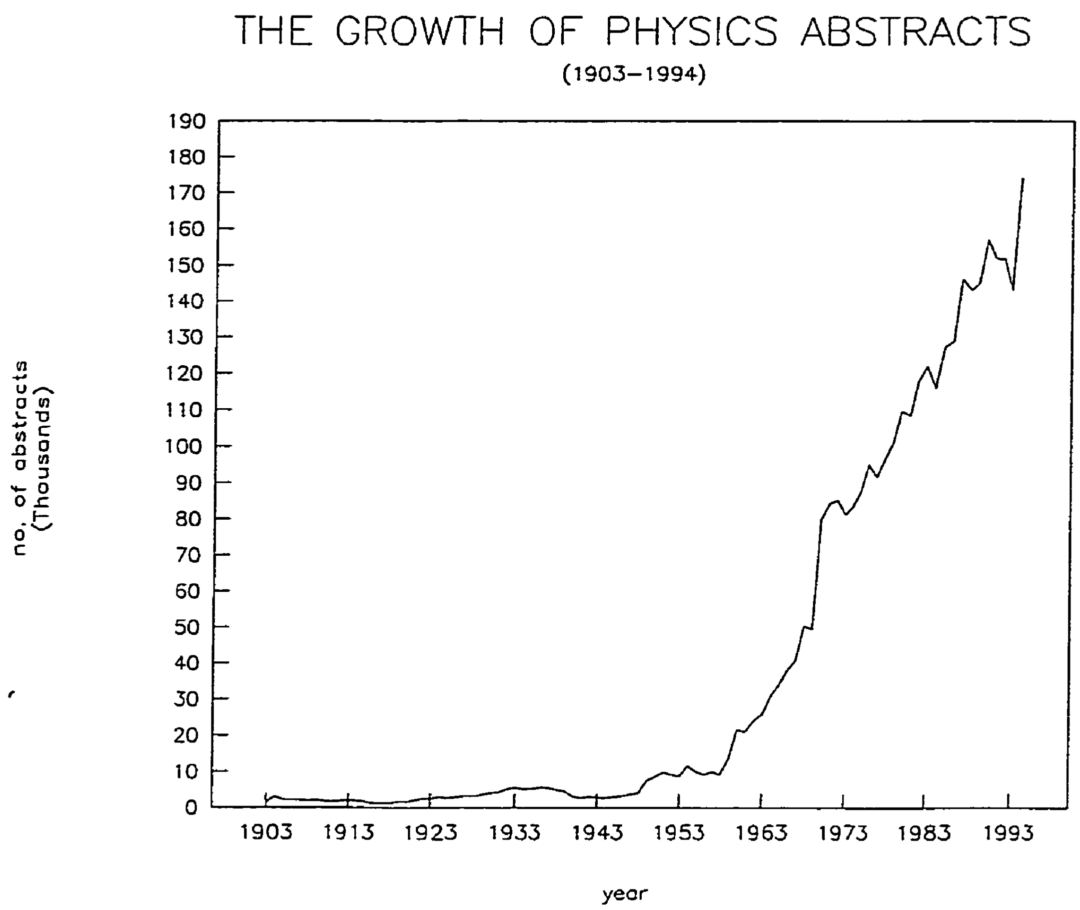


Figure 1. Yearly Growth of *Physics Abstracts* from its inception.

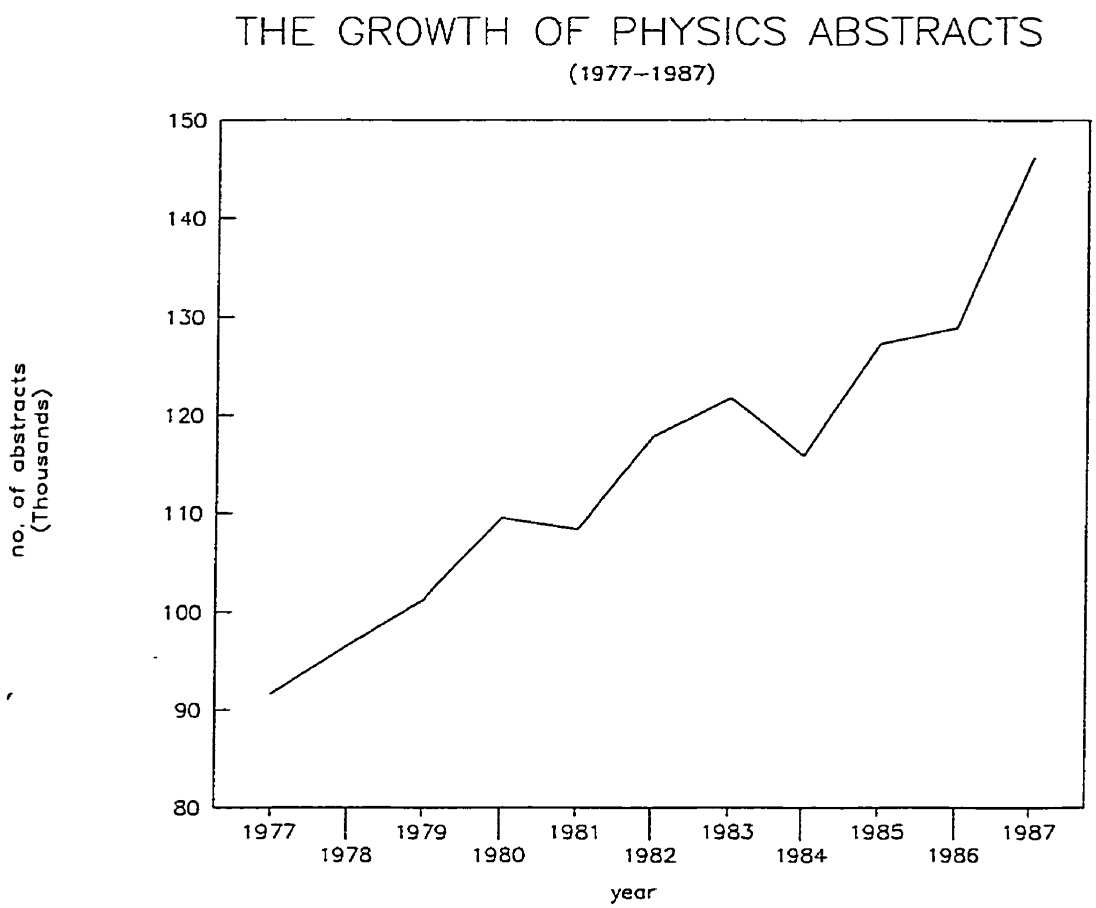


Figure 2. The Growth of *Physics Abstracts* for 1977-1987.

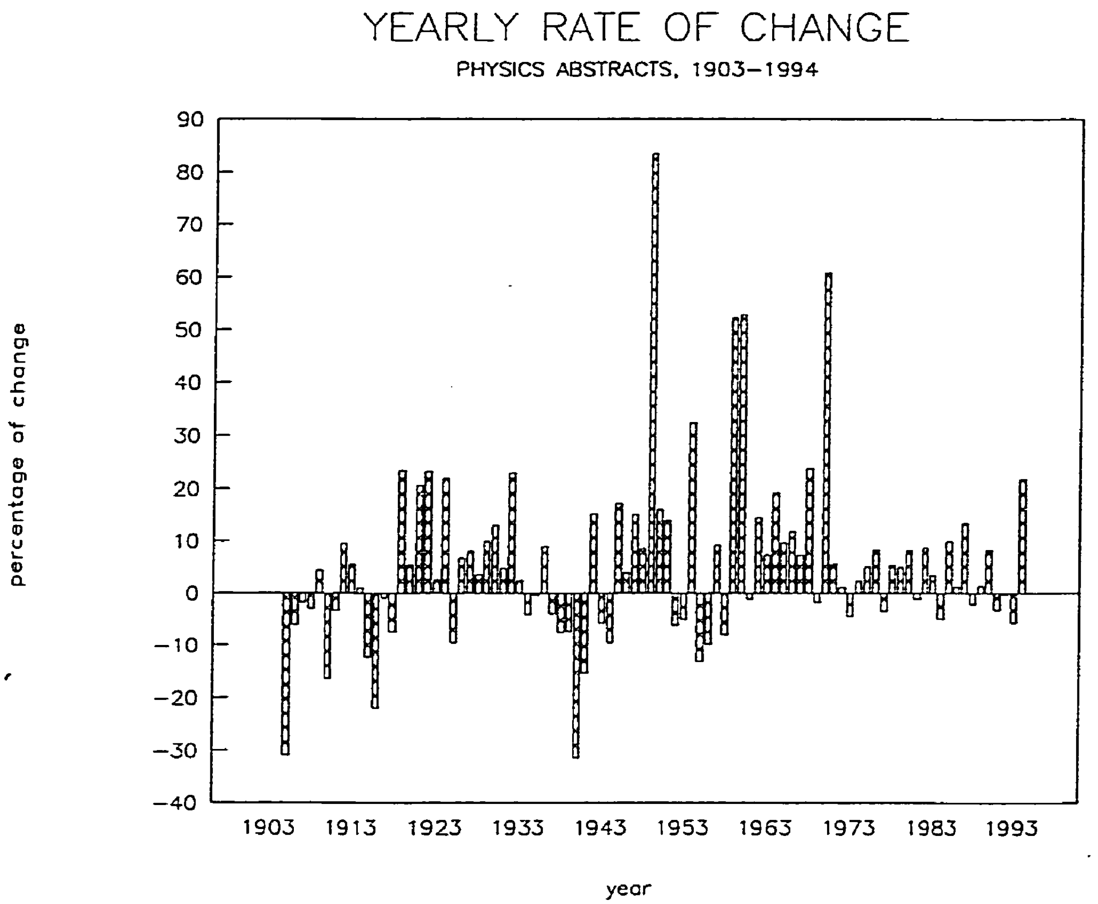


Figure 3. Year by year percentage change of *Physics Abstracts*.

periods of major increases: in 1949 at 83.5%, in 1954 at 32.4%, in 1959 and 1960 each at 52% and in 1970 at 60.9%. All these increases were accompanied by classification changes. The increase for 1903 and the following drop for 1904 are anomalies due to the large number of journals adopted for indexing at the inception of *Physics Abstracts*.

This stability in growth rates is also reflected in the comparison of subject allocations for 1977 versus 1987. As the chart in Figure 4 shows, there was hardly any change in the proportionate coverage of the various topics in *Physics Abstracts* between the first (1977) and the last (1987) year of study. The graph compares the number of abstracts for 1977 versus 1987 by the divisions that represent major fields of study in physics. These are the same divisions listed above in Table 1 (pages 80-81).

Overall, while there has been an increase in the total numbers of abstracts published per year, the proportion of volume accorded from one field to another has not changed much. With two exceptions, the changes amount to less than 1% difference. The exceptions are in Condensed Matter I (Division 6) which went from 9.1% in 1977 to 11.8% in 1987 while Condensed Matter II (Division 7) fell from 15.5% in 1977 to 14.8% in 1987, for an overall increase for Condensed Matter of 1.8%. Geophysics (Division 10) decreased from 15.7% in 1977 to 13.4% in 1987, a reduction of 2.3%. On the average, the difference comes to less than 2000 articles over 58 chapters per year or 34 abstracts per chapter per year. Thus, growth cannot be ascribed to the development of one field alone.

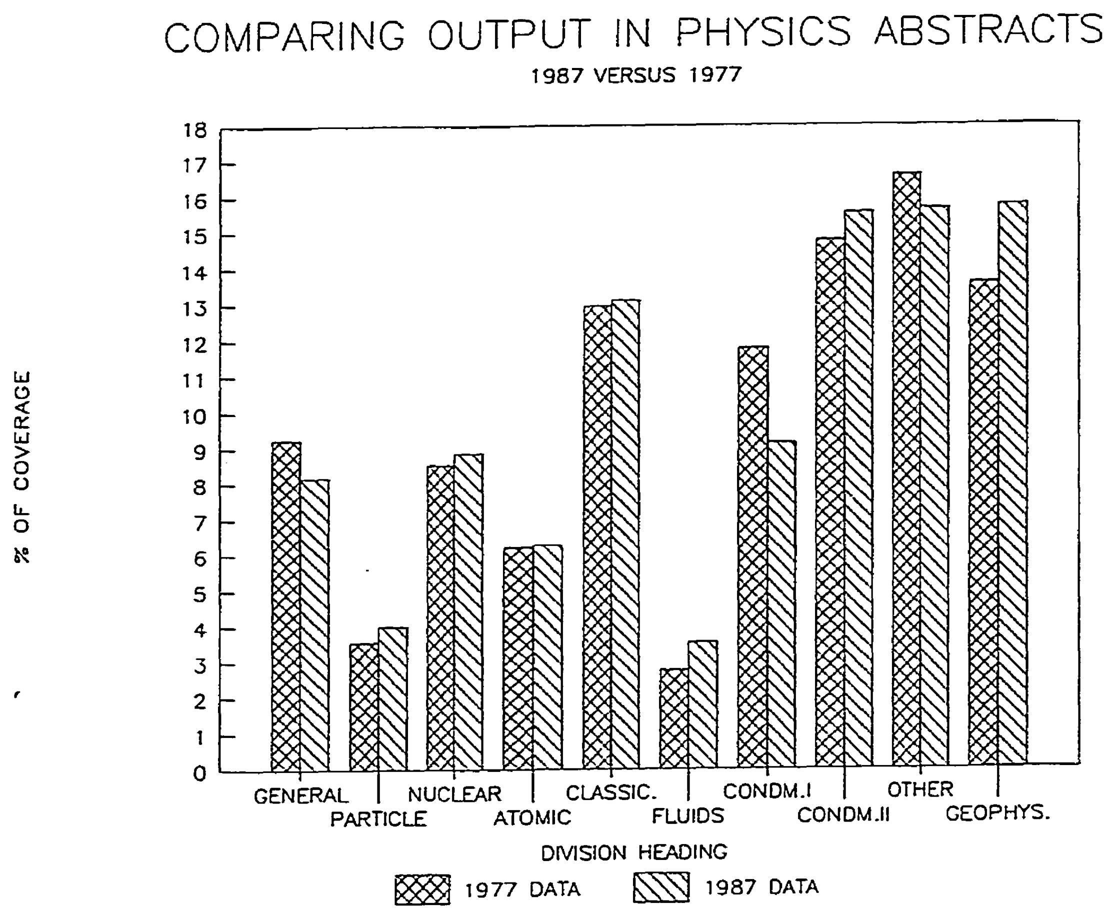


Figure 4. Coverage of major divisions of *Physics Abstracts* for 1977 and 1987.

Finally, Figure 5 illustrates the number of articles carried each month by *Physics Abstracts* between 1977 and 1987. The growth during this period is linear with a coefficient of determination of R^2=0.49.[182] There are several spikes but no outliers in the data. Most months of above average output are followed by months of below average output. The average number of items abstracted each month rose from 7600 in 1977 to about 12,000 in 1987. Overall, the literature of physics exhibited a uniform rate of growth during that time.

5.2. General characteristics of the chapters and outliers:

The raw data indicate that although the total *Physics Abstracts* output grew from 91,677 in 1977 to 146,131 in 1987 (up 59%), not all chapters showed growth. Based on the values of b_1 (the trend) in the regression statistics (see Table B1 in Appendix B, page B2) the trends can be classified into six groups: up, flat, and down, depending whether the statistics are significant or not significant. It is understood that even a flat output every month still represents growth in the field, except that that growth is simply linear. Thus, "up" represents an acceleration in the growth pattern of a chapter, but not necessarily exponential growth, "flat" means the growth has remained linear, and "down" means that over the eleven years the growth of the chapter has displayed a decelerating pattern. Since there can be no negative publication output, the worst that can happen is that a field slows down or simply stops growing. It cannot go down.

---

[182]The regression equation is $y= 7454.87 + 32.03b_1$.

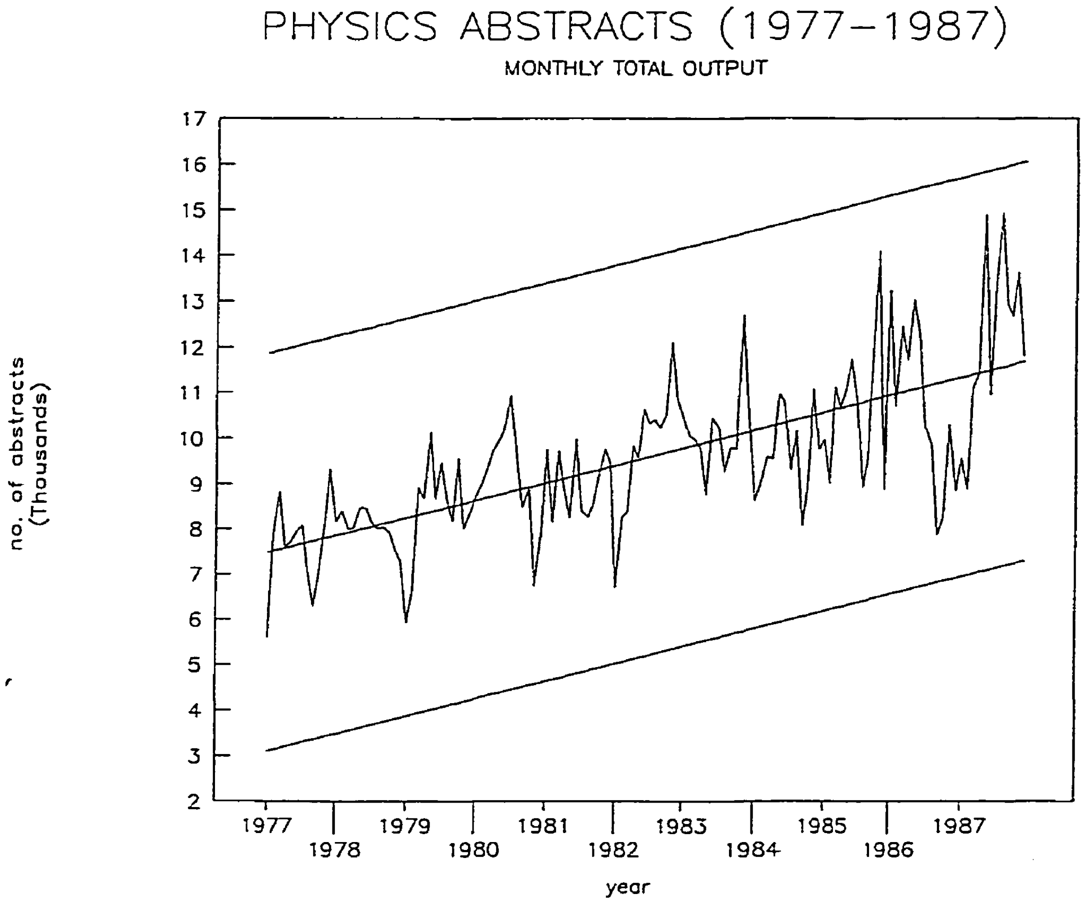


Figure 5. The number of abstracts published in *Physics Abstracts* by month, 1977-1987.

1. **Up**: Up were those chapters where b_1 was positive and more than 0.1 and significant ($p < 0.01$). Those chapters where the value of b_1 was positive but between zero and 0.1 did not have significant statistics and as such could not be considered to have accelerated over the 1977-1987 period. This can also be verified by visual inspection. Thus, forty of the fifty-eight chapters (69%) met those criteria. They cannot be further differentiated with respect to outliers: several had them, several others did not. All four case studies (i.e. *Physics Abstracts* Chapters 2, 36, 73 and 74) analyzed in Chapter 6 below belong to this category and all at one point or another exhibit acceleration and exponential growth.
2. **Flat**: Those chapters where either b_1 was positive and above 0.1 but not significant ($p > 0.01$), or b_1 was between 0 and 0.1 (whether it was significant or not) or b_1 was negative but not significant were categorized as flat or showing no trend. Seventeen of the fifty-eight chapters (29%) showed no trend. They cannot be further differentiated in terms of having outliers at all or with respect to having more than one outlier.
3. **Down**: Those chapters where b_1 was negative and $p < 0.01$ were categorized as showing a deceleration. Only one of the fifty-eight chapters (2%) showed a downward trend and that was Chapter 35 (Properties of atoms and molecules; instruments and techniques).

No correlation could be found between trend and the presence of outliers in a chapter or their number. Some chapters that accelerated strongly had no outliers, such as Chapter 42 ($b_1 = 1.89$), Chapter 61 ($b_1 = 2.18$) and Chapter 81 ($b_1 = 1.86$). Chapter 35 that showed a downward trend ($b_1 = -0.1$) did not exhibit any outliers either. On the other hand, some remaining at linear growth had several outliers. Chapter 52 (b_1 = 0.4), Chapter 66 (b_1 = 0.39), Chapter 67 (b_1 = 0.02) and Chapter 75 (b_1 = 0.51) all had three outliers each. Chapter 77 (b_1 = 0.17) had four. Chapter 29 (b_1 = 0.79) had five outliers occurring two years apart.

It is thus interesting that chapters with significant as well as insignificant growth statistics had several outliers present. Those with three or more outliers equally spaced apart were for the most part linear or showed a weakly accelerating trend. In other words, chapters with occasional instances of increased activity nevertheless remained linear over eleven years (1977-1987). Given that all of these outliers are due to conference proceedings, it is telling that a field of science with regular active conferences would remain linear over the course of a decade. In other words, even though there were a number of conferences where a large number of papers were presented on seemingly new discoveries or new ideas, the total number of publications extant increased at a linear pace but the field overall did not grow and showed no acceleration. It seems this can be used as another confirmation of the idea of physics as a mature discipline in the 1980s. This theme will be revisited in the discussion in Chapter 7.

As to outliers, three types of *outlier behaviour* were observed. They are listed in Table 2 below (page 117).

1. **Random:** The vast majority of the chapters showed single outliers or a random number (if more than one), coinciding with occasional conferences and the increased number of papers they attracted. Of the fifty eight chapters in the study sample,

twenty two (38%) had only one outlier. Except for Chapter 2 (Mathematical Methods in Physics) and Chapter 73 (Electronic structure and properties of surfaces), all can be attributed to conference proceedings.

Fifteen chapters (26%) had two outliers, anything from a few months to eleven years apart. The two extremes include, for example, Chapter 25 (Nuclear reactions and scattering) where the two outliers occur in April 1977 and November 1987 and Chapter 94 (Astronomy and space physics) where the two outliers belong to the October/November 1984 and February/March 1985 periods and are entirely due to the proceedings of the huge International Cosmic Ray Conference held in Bangalore, India in 1983.

Three chapters also exhibited negative outliers, that is, points piercing the ± 3 standard deviation envelope on the negative side. These were Chapter 33 (Molecular spectra and interactions) for January 1987, Chapter 64 (Equations of state and phase equilibria) for October 1986 and Chapter 65 (Thermal properties of condensed matter) for April 1987. These months belong to the time period between September 1986 and April 1987 when the total monthly output of *Physics Abstracts* itself fell to unusual lows. The reasons for these declines are unknown at this time.

It is difficult to attach any importance to specialties that carry one or two outliers that arise as a result of conference papers getting indexed. It is not inconceivable that a given field of study could be subject to an occasional moment of excitement and that a conference could be held to examine the matter in some depth. If the excitement were to lead to a string of new discoveries one would expect an acceleration and, conceivably, more conferences and more outliers. The fact that those fields reflect one or two outliers shows that the excitements were short lived. As well, there are regular conferences held in the normal course of the life of a professional association one of whose duties is to promote communication among its members. Therefore, the occasional spikes that intersperse the data must simply be seen as noise accompanying a growth pattern.

2. **Regular:** These are characteristic of relatively slow moving specialties, where the only outliers are associated with regularly held biannual or triennial conferences. These are Chapter 29 (Experimental methods and instrumentation for elementary particles and nuclear physics) and the Particle Accelerator Conference held every two years, Chapter 67 (Quantum fluids and solids) and the International Conference on Low Temperature Physics held every three years, Chapter 75 (Magnetic properties of materials) and the International Conference on Magnetism held every three years, and Chapter 77 (Dielectric properties of materials) and the International Meeting on Ferroelectricity and the European Meeting on Ferroelectricity held alternately every two years. In all these outliers the proportion of conference proceedings account for 48% to 89% of their contents. Especially Chapter 67 represents an oddity because three regularly held conferences with about 150 papers each appear in the midst of an almost perfectly linear trend ($b_1 = 0.02$) of about thirty items indexed per month.

There are two common denominators to all these regularly held conferences. One is that they all have been sponsored by major professional associations and the second is that the proceedings of almost all have been published in journals. Given that *Physics Abstracts* indexes all those journals, it is inevitable that the large numbers of papers indexed would produce outliers. Except for the Particle Accelerator Conference that is part of Chapter 29, all of the other conferences have been sponsored by, among others, the *International Union for Pure and Applied Physics* (IUPAP). Depending on the location where the conferences were held, one or more additional local or national professional associations co-sponsored the conferences. The *Particle Accelerator Conferences* were sponsored by the *Institute of Electrical and Electronics Engineers* (IEEE) since their inception, and starting with the 1987 conference the title of the conference was changed to the *IEEE Particle Accelerator Conference.*

It is not difficult to speculate why such conferences will be held regularly over two or three decades. For one, given that they are all international conferences, they bring together a large number of scientists, giving them the opportunity to exchange views in person. Second, one of the fundamental raisons d'être of an association is to facilitate communication both among its members and with the outside world. IUPAP, being an international and non-profit association, is active in supporting the learned and educational activities of its members. Third, conferences are money makers. Associations largely sponsor conferences because they will bring significant revenues that can be used to further their agendas nationally and/or internationally. Thus, the regular outliers seen in the various chapters of *Physics Abstracts* are nothing more than the reflection of regular business by various physics associations that sponsor conferences. The fact that during the period of study (1977-1987) they occurred in the midst of otherwise linear growth patterns shows that they fulfilled the regular communication function for physicists but did not add significantly more to the discipline of physics per se.

The one odd case among these conferences is that of the Particle Accelerator Conference in Chapter 29. During the examination of the abstracts it was noted that a large number of them carry two abstract numbers: one for *INSPEC, Part A: Physics Abstracts* but also a second one for *INSPEC, Part B: Electrical and Electronics Abstracts*. Given that the problem of particle accelerators is largely an engineering problem, why should so many abstracts find their way into Physics Abstracts? In addition, is the linear growth seen in Chapter 29 also present in its counterpart, the *Electrical and Electronics Abstracts*?

An online search was conducted for the number of items abstracted for Section B74.10: Accelerators. Section B of INSPEC is the electronic database version of the *Electrical and Electronics Abstracts*. The search was widened to include the number of items abstracted yearly between 1970 and 1994, that is practically the entire span of *Physics Abstracts*. The results are shown below in Figure 6. The jagged curve is the yearly number of items abstracted, and the smooth curve is the six year centered average of the raw data. The six year average was chosen arbitrarily, only to reflect the accelerating growth pattern starting at the end of the 1970s and continuing into the 1980s. The data show that in fact the topic of *particle accelerators* has been an active one over the last twenty-five years. Even between 1977 and 1987 the average

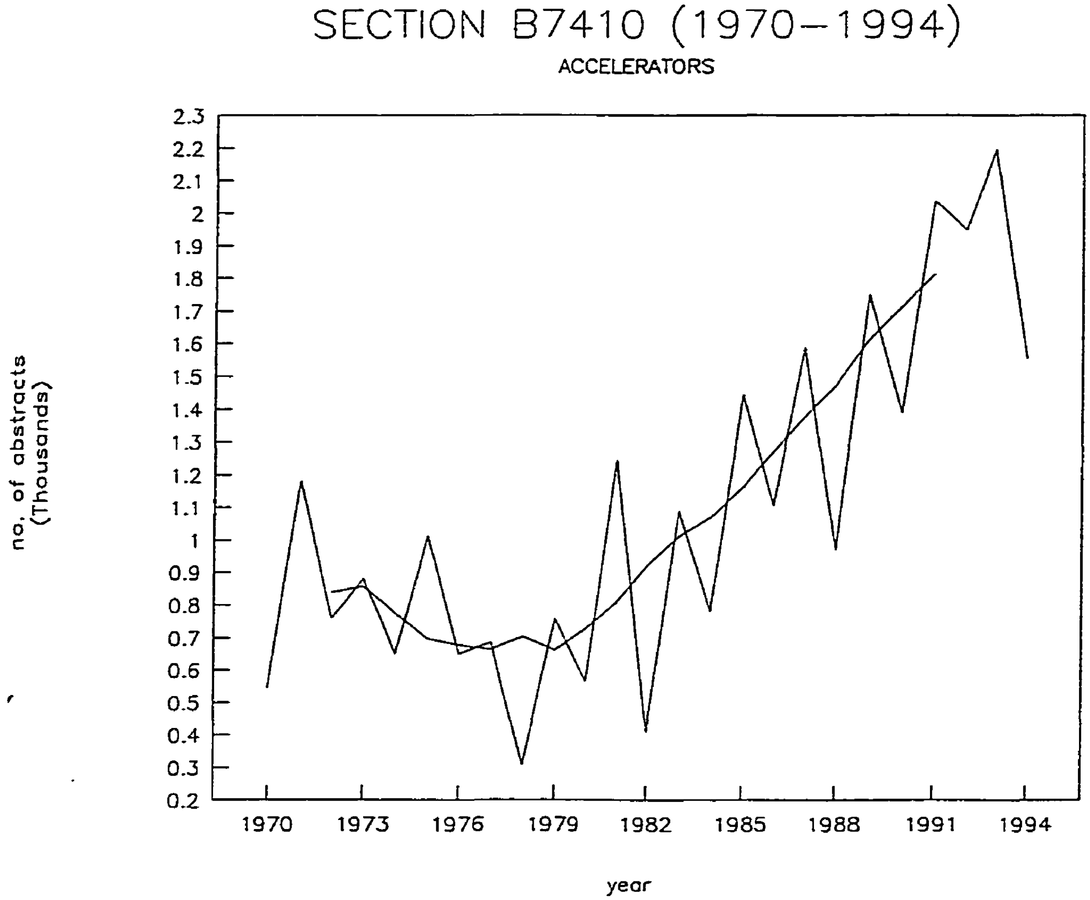


Figure 6. The growth of items abstracted for Section B74.10 (Accelerators) of *Physics Abstracts*, 1970-1994. The jagged curve is the raw data and the smooth curve represents the six-point centered moving average.

numbers abstracted doubled from 685 in 1977 to 1374 in 1987, and the growth curve followed a steep upward trend.

These results indicate that *particle accelerators* has not been a dormant topic but one which has seen significant growth. The fact that Chapter 29 of *Physics Abstracts* displays a flat curve interspersed with the triennial conference outliers means that some of the papers held during the conferences of an active research area (Section B74.10) were also relevant to physics and were thus abstracted for it. It is understandable that a major engineering problem will attract some fundamental work, and it is that fundamental work that was being abstracted by *Physics Abstracts*, even if the field it belonged to was not showing much growth.

3. **Significant**: These are chapters that exhibit major outliers and fulfil the criteria described in Section 4.5 of the chapter on methodology. They all contain outliers at more than three standard deviations, the entire chapter exhibits significant growth, and their moving averages reflect significant acceleration patterns. These are Chapter 2 (Mathematical methods in physics), Chapter 36 (Studies of special atoms and molecules), Chapter 73 (Electronic structure and electrical properties of surfaces, interfaces, and thin films), and Chapter 74 (Superconductivity). These chapters are analyzed in detail in the next chapter.

The fact that only four chapters so far have fulfilled all the criteria for information epidemics indicates that, in physics at least, epidemics are not widely occurring phenomena. They remain unusual events. This means either that the criteria developed in this work are too strenuous and that epidemics have characteristics that are not captured with the methodology of this thesis or else that physics has become a *mature* field where novelties that make news headlines are not necessarily discoveries that change the course of a given specialty permanently.

5.3. General comments on outliers:

Table 2 below summarizes the outliers obtained from analysis with SPSS-PC®. Each line indicates the chapter number and the number of outliers that belong to the chapter. It is accompanied by the month in which it occurs, the number of abstracts giving rise to it, the percentage of abstracts that are due to a given (or a number of given) conference proceedings, and the title of the proceedings that form the largest portion of the outlier. For example, in Chapter 4 there are two outliers, one occurring in June 1985 and the other in July 1987. The first contains 199 abstracts, the second contains 203 abstracts. The portion of conference papers in them is 59.8% and 61.6% respectively. Thus, 40.2% of the first outlier and 38.4% of the second are due to journal articles and other types of publications. Given that the number of publication in formats other than journal articles and conference papers is less than 5%, it can be safely assumed that these portions are largely due to conference papers. If there are no conference proceedings to account for a given outlier and it is due to journal articles, it is indicated as such. In a few cases where the outlier pierces the negative regression line, it is indicated as a "NEGATIVE outlier". Each outlier follows the numbering given in the graphs of Appendix A and follows a chronological order.

Table 2--List of Outliers (*Physics Abstracts*, 1977-1987)

Ch. no.	Outlier no.	Year/ Month	No. abs.	Portion of outlier	CONTENTS
2	1	8707	282		articles in journals
3	1	8211	252	10.7	International Conference on Mathematical Problems in Theoretical Physics, 6th, 1981. International Conference on Nonpotential Interactions & Their Lie-Admissible Treatment, 1982.
4	1	8506	199	59.8	Marcel Grossman Meeting on General Relativity, 3rd, 1982.
4	2	8707	203	61.6	Marcel Grossman Meeting on General Relativity, 4th, 1985.
5	1	8311	218	16.1	Dynamical Systems & Chaos: Proc. of the Sitges Conference, 1982.
5	2	8705	253	11.4	Statistical Physics and Dynamical Systems: Rigorous Results, 1984.
6	1	8511	132	40.9	Annual Frequency Control Symposium, 38th, 1984. International Conference on Precision Measurement and Fundamental Constants, 1981.
11	1	8211	226	20.8	International Conference on Mathematical Problems in Theoretical Physics, 6th, 1981. *Physical Review D*, vol.25, no.12 and vol.26, no.2, 1982.
21	1	8710	264	59.1	International Conference on Hyperfine Interactions, 7th, 1986. International Nuclear Physics Conference, 1986.
23	1	7704	84	45.2	International Conference on Nuclei Far from Stability, 3rd, 1976.
23	2	8705	80	47.5	International Symposium on Weak and Electromagnetic Interactions in Nuclei, 1986.
24	1	8107	90	53.3	Topical Conference on Giant Multiple Resonances, 1979.
25	1	7704	322	67.7	International Conference on Interactions of Neutrons with Nuclei, 1976.
25	2	8711	314	31.2	International Nuclear Physics Conference, 1986.

Table 2--Continued

Ch. no.	Outlier no.	Year/ Month	No. abs.	Portion of outlier	CONTENTS
28	1	8306	816	72.4	1982 Nuclear Science Symposium.
					ASTM-EURATOM Symposium on Reactor Dosimetry, 4th, 1982.
					Canadian Nuclear Society, 3rd Annual Conference, 1982.
					Conference on Fast, Thermal and Fusion Reactor Experiments, 1982.
					Fusion Technology Symposium, 12th, 1982.
					IMACS World Conference on System Simulation and Scientific Computation, 10th, 1982.
					International Symposium on Actinide Recovery from Waste and Low Grade Sources, 1981.
					LMFBR Safety Topical Meeting, 1982.
					Neutron and Its Applications, 1982.
28	2	8412	937	87.4	Annual Meeting of Nuclear Technology, 1984.
					Annual Meeting of the American Nuclear Society, 1984.
					Annual Symposium on Safeguards and Nuclear Materials Management, 6th, 1984.
					International Conference on Radioactive Waste Management, 1983.
					Symposium on Fusion Engineering, 10th, 1983.
					Topical Meeting on Fusion Reactor Materials, 3rd, 1983.
29	1	7712	430	67.4	1977 Particle Accelerator Conference
29	2	7909	444	73.4	1979 Particle Accelerator Conference
29	3	8110	474	84.2	1981 Particle Accelerator Conference
					Fifth Tandem Conference, 1980.
					International Conference on Experimentation at LEP, 1980.
					International Conference on Polarization Phenomena in Nuclear Physics, 5th, 1980.
29	4	8311	555	82.7	1983 Particle Accelerator Conference
					International Workshop on Mercuric Iodide Nucleic Reaction Detectors, 5th, 1982.
					Yamada Conference VI on Neutron Scattering in Condensed Matter, 1982.
29	5	8605	557	60.7	1985 Particle Accelerator Conference
					International Conference on Electrostatic Accelerator Technology and Associated Boosters, 1985.
					Symposium on X and Gamma Ray Sources and Applications, 6th, 1985.

Table 2--Continued

Ch. no.	Outlier no.	Year/ Month	No. abs.	Portion of outlier	CONTENTS
32	1	8606	142	30.3	International Conference on the Physics of Electronic and Atomic Collisions, 1985.
33	1	8701	125		NEGATIVE outlier
33	2	8708	471	21.2	International Conference on Raman Spectroscopy, 10th, 1986. International Conference on Resonance Ionization Spectroscopy, 3rd, 1986.
34	1	8402	404	79.2	International Conference on the Physics of Electronic and Atomic Collisions, 13th, 1983.
34	2	8607	311	71.7	International Conference on the Physics of Electronic and Atomic Collisions, 14th, 1985.
36	1	8511	72	38.9	International Meeting on Small Particles and Inorganic Clusters, 3rd, 1984.
36	2	8702	84	33.3	International Symposium on Metal Clusters, 1986.
36	3	8705	78	23.1	International Conference on Muon Spin Rotation, Relaxation and Resonance, 1986.
43	1	8601	293	36.5	IEEE 1984 Ultrasonics Symposium.
44	1	7812	163	63.8	International Heat Transfer Conference, 6th, 1978.
46	1	7905	342	9.6	International Conference on Experimental Stress Analysis, 6th, 1978.
47	1	8705	592	10.0	International Symposium on Finite Element Methods in Flow Problems, 6th, 1986. IUTAM Symposium on Fluid Mechanics in the Spirit of G.I. Taylor, 1986. Spatio-Temporal Coherence and Chaos in Physical Systems, 1986.
51	1	8102	56	51.8	International Conference on Gas Discharges and Their Applications, 6th, 1980.
51	2	8211	63	38.1	International Conference on Gas Discharges and Their Applications, 7th, 1982.

Table 2--Continued

Ch. no.	Outlier no.	Year/ Month	No. abs.	Portion of outlier	CONTENTS
52	1	7809	495	51.9	International Conference on Phenomena in Ionized Gases, 13th, 1977.
52	2	8005	559	59.2	International Conference on Phenomena in Ionized Gases, 14th, 1979.
52	3	8505	529	62.2	1984 IEEE International Conference on Plasma Science, 1984.
					International Conference on Plasma Surface Interactions in Controlled Fusion Devices, 6th, 1984.
					International Symposium on Heating in Toroidal Plasma, 4th, 1984.
					Symposium on Plasma Double Layers and Related Topics, 2nd, 1984.
62	1	8206	150	40.7	International Conference on Internal Friction and Ultrasonic Attenuation in Solids, 7th, 1981.
					International Conference on Phonon Physics, 1981.
63	1	7810	132	59.8	International Conference on Lattice Dynamics, 1977.
63	2	8206	154	56.5	International Conference on Phonon Physics, 1981.
64	1	8311	216		articles in journals
64	2	8610	81		NEGATIVE outlier
65	1	8704	14		NEGATIVE outlier
66	1	8012	152	53.3	Europhysics Topical Conference: Lattice Defects in Ionic Crystals, 3rd, 1979.
					Fast Ion Transport in Solids, Electrodes and Electrolytes, 1979.
66	2	8203	161	54.7	International Conference on Fast Ionic Transport in Solids, 1981.
					International Meeting on Solid Electrolytes, Solid State Ionics and Galvanic Cells, 3rd, 1980.
66	3	8606	198	47.0	International Conference on Solid State Ionics, 5th, 1985.
67	1	7907	160	81.3	International Conference on Low Temperature Physics, 15th, 1978.
67	2	8205	139	89.2	International Conference on Low Temperature Physics, 16th, 1981.
67	3	8509	140	85.9	International Conference on Low Temperature Physics, 17th, 1984.
68	1	8511	377	7.7	International Conference on Solid Films and Surfaces, 1984.
71	1	8510	328	22.3	International Conference on Valence Fluctuations, 4th, 1984.

Table 2--Continued

Ch. no.	Outlier no.	Year/ Month	No. abs.	Portion of outlier	CONTENTS
72	1	8601	359	26.7	International Conference on the Physics and Chemistry of Low-Dimensional Synthetic Metals (ICSM '84), 1984.
73	1	8511	297	13.5	International Conference on Solid Films and Surfaces, 1984. National Symposium of the American Vacuum Society, 31st, 1984.
	2	8705	278	16.9	International Conference on Solid State Devices and Materials, 18th, 1986. International Conference on Superlattices, Microstructures and Microdevices, 2nd, 1986. International Winter School on Two-Dimensional Systems: Physics and New Devices, 1986.
74	1	8509	284	61.6	International Conference on Low Temperature Physics, 17th, 1984. International Cryogenic Engineering Conference, 10th, 1984. International Cryogenic Materials Conference, 5th, 1984.
74	2	8709	273		articles in journals
74	2	8710	353		articles in journals
74	2	8711	277		articles in journals
75	1	8007	493	58.4	International Conference on Magnetic Fluids, 2nd, 1980. International Conference on Magnetism, 1979.
75	2	8306	516	69.4	International Conference on Magnetism, 1982. Impact of Polarized Neutrons on Solid State Chemistry and Physics, 1982.
75	3	8606	545	65.7	Annual Meeting of the Magnetics Society of Japan, 8th, 1984. Conference on Electronic Structure and Properties of Rare Earth and Actinide Intermetallics, 1984. International Conference on Magnetism, 1985
76	1	8005	280	50.4	International Conference on Mossbauer Spectroscopy, 1979. Joint Intermag-MMM Conference, 1989.
76	2	8505	257	39.7	Congrès Ampère on Magnetic Resonance and Related Phenomena, 22nd, 1984.

Table 2--Continued

Ch. no.	Outlier no.	Year/ Month	No. abs.	Portion of outlier	CONTENTS
77	1	8008	160	48.1	European Meeting on Ferroelectricity, 4th, 1979.
77	2	8205	165	56.4	International Meeting on Ferroelectricity, 5th, 1981.
77	3	8408	176	42.6	European Meeting on Ferroelectricity, 5th, 1983.
77	4	8609	231	73.2	International Meeting on Ferroelectricity, 6th, 1985. International Symposium on Electrets, 5th, 1985.
78	1	8206	527	38.3	International Conference on Amorphous and Liquid Semiconductors, 9th, 1981. International Meeting on Ferroelectricity, 5th, 1981. International Conference on Luminescence, 1981. International Conference on Phonon Physics, 1981.
82	1	7909	386	13.7	Conference on the Applications of Small Accelerators in Research and Industry, 1978. Informal Conference on Photochemistry, 12th, 1976. International Topical Meeting on Muon Spin Rotation, 1st, 1978.
82	2	8708	411	22.6	Colloquium Spectroscopicum Internationale XXIV, 1985 International Conference on Particle Induced X-ray Emission and Its Analytical Applications, 4th, 1986. International Conference on Resonance Ionization Spectroscopy, 3rd, 1986. International Conference on the Application of Accelerators in Research and Industry, 1986.
87	1	8708	982	23.2	Annual Conference of IEEE Engineering in Biology and Medicine Society, 8th, 1986. International Conference on Solid State Dosimetry, 8th, 1986.
91	1	8211	437	10.5	Papers on MAGSAT: *Geophysical Research Letters*, vol.9, no.4, April 1982. Papers on the Eruption of Soufriere Volcano, St. Vincent, 1979: *Science*, vol.216, no.4550, 1982. Symposium on Properties of Materials at High Pressures and High Temperatures, 1981.

Table 2--Continued

Ch. no.	Outlier no.	Year/ Month	No. abs.	Portion of outlier	CONTENTS
92	1	8511	517	7.5	International Atmospheric Electricity Conference, 7th, 1984. *Journal of Geophysical Research*, vol.90, no.C4, 20 July 1985.
93	1	7711	183	40.4	Annual Symposium on Machine Processing of Remotely Sensed Data, 4th, 1977. International Congress on Electronics, 24th, 1977.
93	2	8711	249	32.5	*Journal of Atmospheric and Oceanic Technology*, vol.2, no.1,2,3,4 and vol.3, no.1,2, 1985.
94	1	8410	321	64.8	International Cosmic Ray Conference, 18th, 1983. NATO Advanced Study Institute on Composition and Origin of Cosmic Rays, 1982.
94	1	8411	373	69.2	European Incoherent Scatter Radar: Papers from the EISCAT Workshop, 1983. International Cosmic Ray Conference, 18th, 1983.
94	2	8502	304	68.1	International Cosmic Ray Conference, 18th, 1983.
94	2	8503	225	59.6	International Cosmic Ray Conference, 18th, 1983.
95	1	8303	170	31.2	IAU Colloquium on Instrumentation for Astronomy with Large Optical Telescopes, no.67, 1981. Oberwolfach Conference on Mathematical Methods in Celestial Mechanics, 7th, 1981.
96	1	8708	405	57.3	20th ESLAB Symposium on the Exploration of Halley's Comet, 1986.

Each line in the table above also indicates which issue in *Physics Abstracts* provided the data, what percentage of items in the outlier is accounted for by a significant publication (almost all being conferences), and the name of the conference(s). Some conferences were big enough that they gave rise to outliers in more than one chapter of *Physics Abstracts*. For example, the *International Nuclear Physics Conference* of 1986 produced outliers in both Chapter 21 and Chapter 25. The *International Low Temperature Physics Conferences* produced outliers in Chapter 67 and Chapter 74. One conference, the 18th *International Cosmic Ray Conference*, generated several outliers in Chapter 94 for October-November 1994 and February-March 1985 because of its massive volume (779 abstracts).

This also shows that a given conference may not necessarily be limited to one specialty in physics but may cover several specialties and can produce growth in several of them at the same time. Some general conferences may be organized around special themes that carry a large number of papers and produce outliers in certain chapters. For example, the 17th *International Conference on Low Temperature Physics* of 1984 gave rise to outliers in Chapter 67 (Quantum fluids and solids) as well as Chapter 74 (Superconductivity). The conference papers accounted for 140 of the 163 abstracts (85.9%) in the outlier no.3 of Chapter 67 - all in a chapter that normally carried about 40 abstracts per month and where the trend was flat between 1977 and 1987. In the same vein, the superconductivity papers from the same conference contributed another 175 abstracts to Chapter 74 (61.6% of the papers in that outlier), a chapter that normally carried about 100 abstracts a month, and gave rise to outlier no.1 there. Finally, Figure 7 illustrates the number of outliers obtained from the data over the 11 year period (1977-1987) that is the focus of this study. It is based on results obtained from regression with the use of SPSS-PC•. It clearly indicates that the number of outliers is increasing with time. This is likely due to an increasing number of conferences being held. More conferences attract more attendance, and more presentations provide more spikes and outliers.[183] Overall, this reflects increasing volatility with time, that is, increasing activity accompanied by increasing volume. It also indicates that, where physics is concerned, the growth of the literature may be due more to conference papers than journal articles as such (the portion of all other forms of literature is too small to be significant here). Physics seems to be growing more thanks to an increasing number of conferences being organized and an increasing number of papers being presented. The journal literature is increasingly being viewed as the archival repository for knowledge in physics, most communication for a number of years now has been taking place via preprints or electronic mail, and conference papers have acted more as an outgrowth of the effort to enhance personal communication among physicists. For example, the electronic print archives at the Los Alamos National Laboratory have grown in only three years to attract some 20,000 users from over sixty countries, processing over 30,000 messages per day[184]. A number of physicists have ceased relying on the printed

---

[183]H.H. Barschall and W. Haeberli, "Conference Proceedings in Physics," *College and Research Libraries* 53 (1992): 563-6.

[184]Paul Ginsparg, "First Steps Towards Electronic Research Communication," *Computers in Physics* 8 (1994): 390-6.

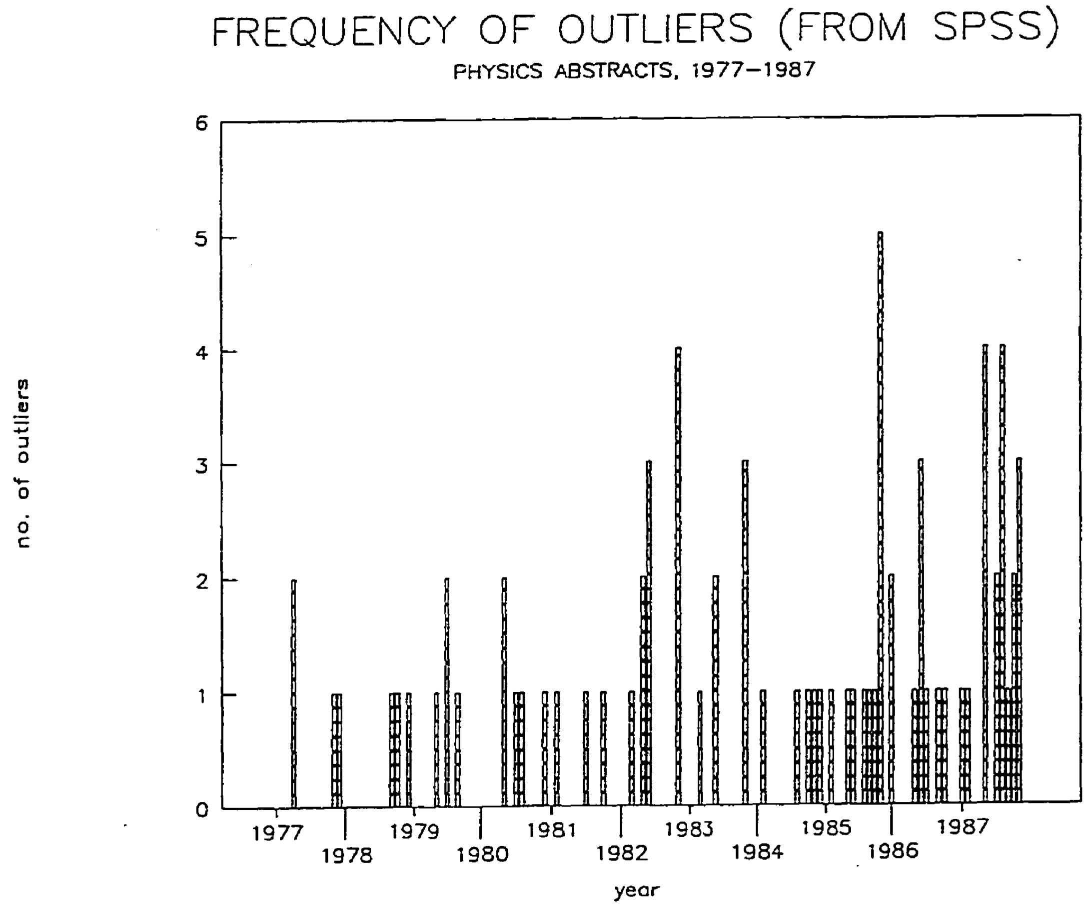


Figure 7. The monthly number of outliers in *Physics Abstracts* chapters for 1977-1987.

journal. They now read the electronic preprint archive exclusively and refer to a paper's e-print index number in their publications. Similarly, many institutions have now discontinued mailing paper preprints.[185] In other words, the increasing concentration of outliers into recent years (see Figure 7 above) and the fact that the vast majority of them are due to conference proceedings, indicate that the communication system in physics is undergoing a fundamental change. The repercussions of these developments will be discussed in Section 7.4 (page 218) below.

---

[185]Ibid., 393.

CHAPTER 6

DATA ANALYSIS: SELECTED CHAPTERS

The characteristics of the case studies presented in this chapter fulfill the four inclusion criteria set out in Section 4.5 (pages 87-89). Again, they are:

1. The chapter must show dramatic or sudden growth patterns.
2. The regression statistics must be significant.
3. The chapter must include an outlier that pierces the upper regression envelope at +3.0 standard deviation.
4. The moving average must reflect a jump in growth.

Only four chapters meet these criteria: Chapter 2: Mathematical Methods in Physics, Chapter 36: Studies of Special Atoms and Molecules, Chapter 73: Electronic Structure and Electric Properties of Surfaces, Interfaces and Thin Films, and Chapter 74: Superconductivity. Each of these is analyzed separately in the following pages according to its growth characteristics and the contents of its outlier(s).

6.1. Mathematical Methods in Physics (Chapter 2)

6.1.1. <u>Growth</u>

This chapter is entitled "Mathematical Methods in Physics". There is a major outlier in 1987 that dominates the chart (see Figure 8). During most of the period

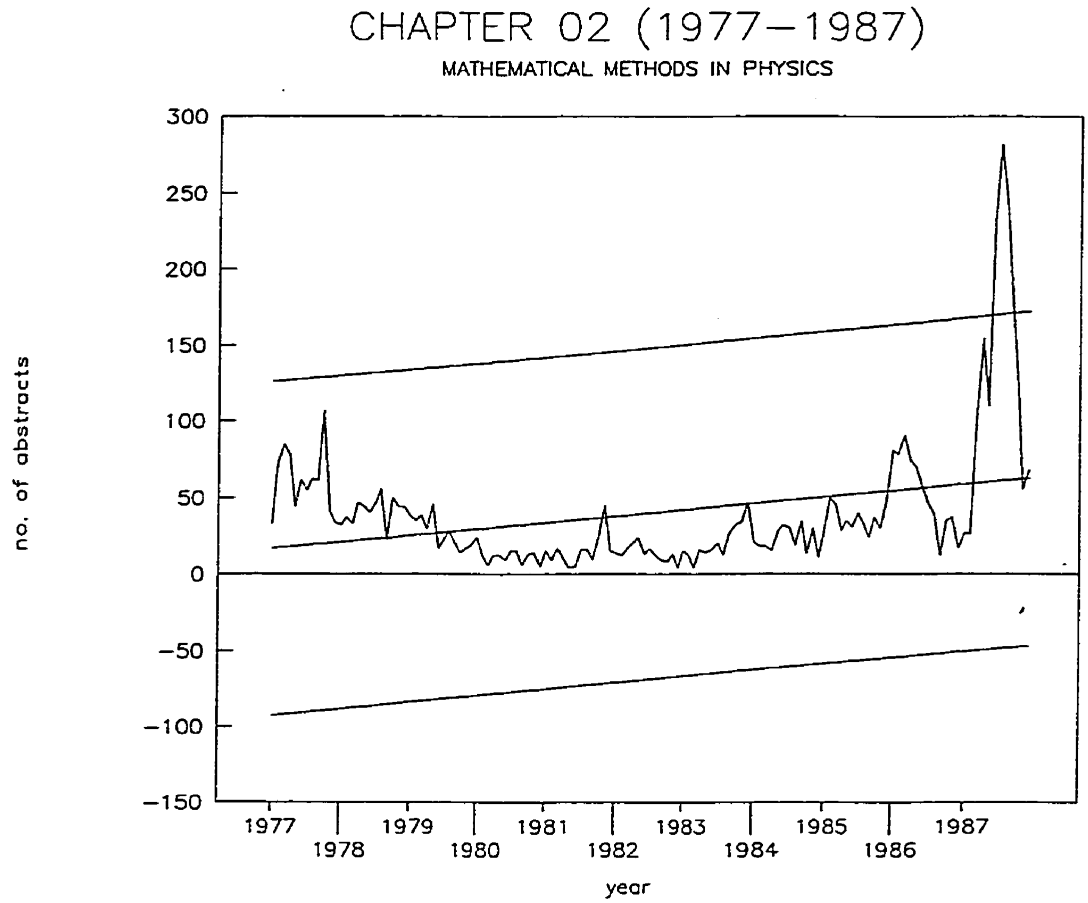


Figure 8. Monthly number of abstracts in Chapter 2 for 1977-1987.

covered here, that is 1977 to 1987, the field was largely quiescent. Following some activity in the late 1970s output kept dropping, and there was very little publication until late 1984.

From late 1985 to early 1986 the number of publications increased up to 91 for March 1986 and subsequently declined. In early 1987 a major flurry of activity resulted in the inclusion of up to 282 abstracts for the month of July 1987. The activity then declined to 68 for December 1987. The regression and the ± 3 standard deviation lines reveal a significant outlier that covers much of 1987. Given that it stands out dramatically from previous activity and that it rose and declined within the span of one year, and not just one or two issues of *Physics Abstracts*, it must be regarded as the *information epidemic* that is being sought here.

The predominant outlier of 1987 in this chapter fulfils criteria 1 and 3. The value for the slope is a 0.35 and statistically significant ($p < 0.001$). This fulfils criterion no.2. Given the size and volume of the outlier, it is obvious that criterion 4 is met by the very fact that there is a major jump in growth in 1987. Consequently, there is no need to plot a graph of the moving average. Thus, all four criteria for analysis are met. In addition, given that the major outlier arises and falls back within one year on an otherwise flat graph, there is no point in doing any nonlinear regression.

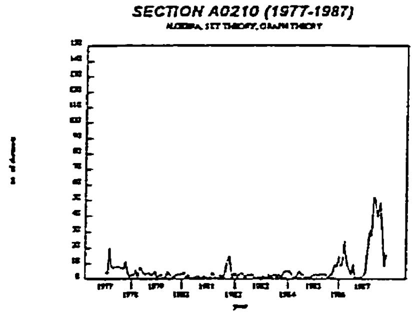
Chapter 2 is subdivided into eight sections. They are:

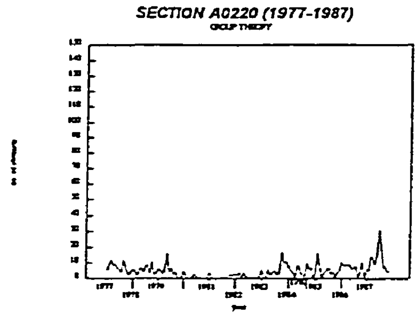
A02.10: Algebra, set theory, and graph theory
A02.20: Group theory
A02.30: Function theory, analysis
A02.40: Geometry, differential geometry, and topology

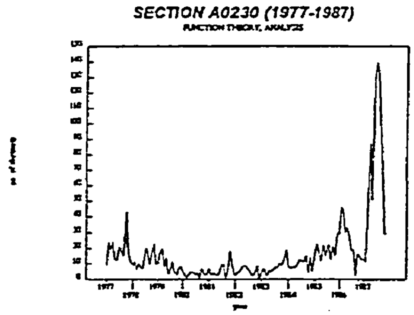
A02.50: Probability theory, stochastic processes, and statistics
A02.60: Numerical approximation and analysis
A02.70: Computational techniques
A02.90: Other topics in mathematical methods in physics

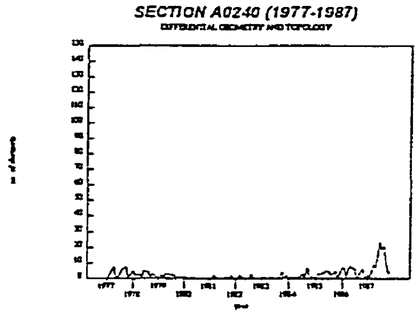
Figure 9 illustrates the activity in each of the sections. The chart for Section A02.90 has been omitted because there was very little included in it, only thirty-nine abstracts over eleven years. The figure reveals that while overall activity increased in almost all the sections in 1987, by far the most significant portion was in Section A02.30. Therefore, the remainder of the analysis for this chapter will be conducted on Section A02.30 (see Figure 10).

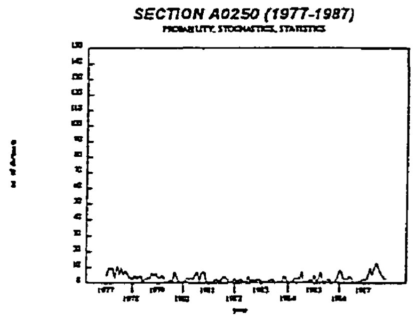
Section A02.30 is entitled "Function Theory, Analysis". Its growth patterns parallel those of the overall Chapter 2. The most interesting part in it is the epidemic that occurred in 1987.

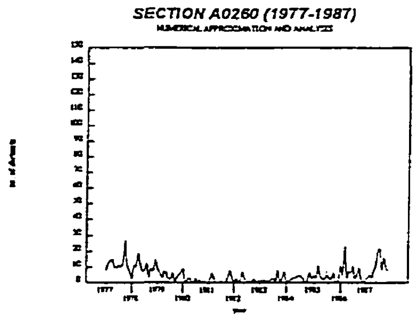












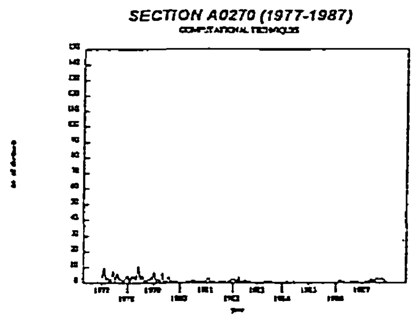


Figure 9. Comparison of the growth patterns for the sections comprising Chapter A02.

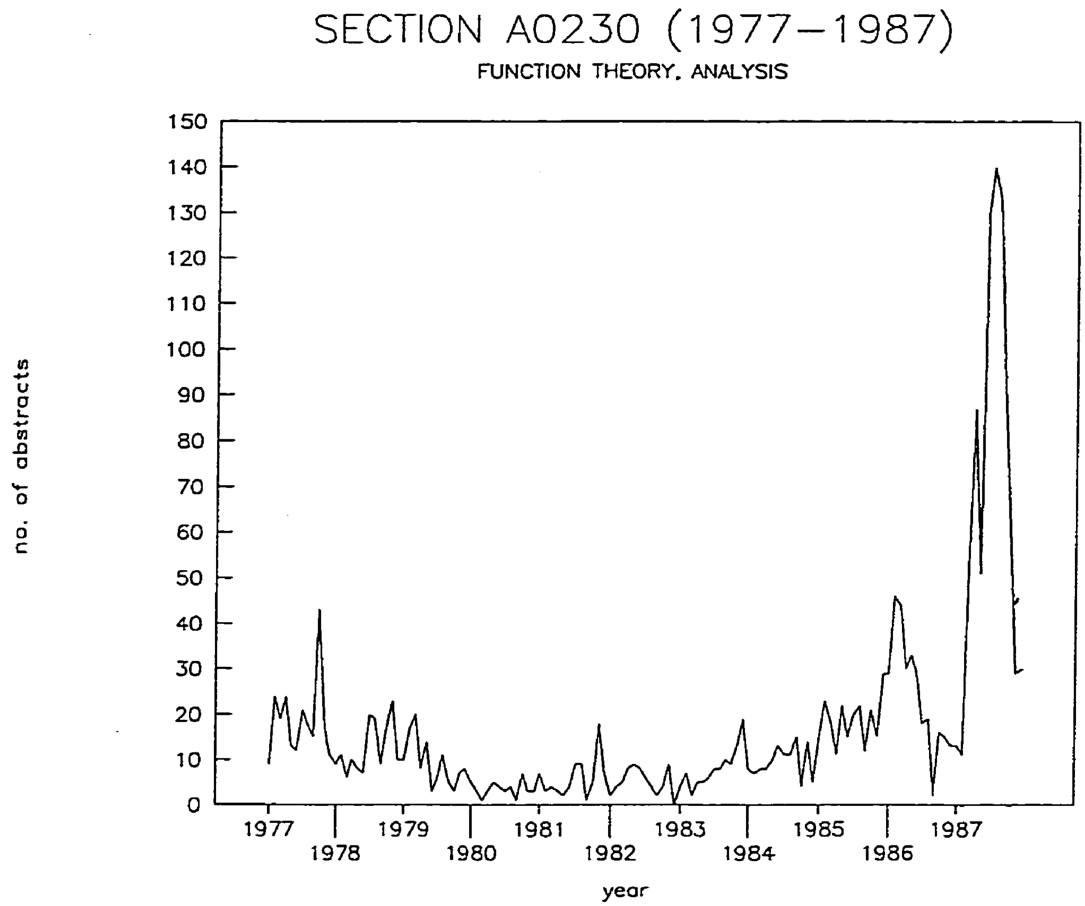


Figure 10. Monthly numbers of abstracts for Section A02.30 for 1977-1987.

6.1.2. Contents of the outlier

A cursory examination indicates that the volume of abstracts in this section is almost completely accounted for by journal articles. If one is expecting to find an influential article giving rise to an epidemic, it is to be expected that it will be found, by definition, in the reference lists of the works making up that epidemic. Thus, if a given field is under the strong influence of a given work, this influence should be reflected in the citation patterns of the papers making up the outlier.

Most of the journals indexed by *Physics Abstracts* are also indexed by the *Science Citation Index*, and therefore the records of most of the abstracts for each chapter and section are available from the *Science Citation Index* database. The full records of all the abstracts for 1987 were downloaded from SCI, and the list of references of all the papers were examined.

In the few cases where a journal is not abstracted by SCI, a photocopy of the bibliography available either in Montreal or at CISTI (Canada Institute for Scientific and Technical Information) in Ottawa was obtained. This was done for all journals or conference proceedings that contributed three or more abstracts to the contents of the epidemic. Thus, for Section A02.30, 98.3% of all the 829 records of the 1987 *Physics Abstracts* have been included in the analysis. The remaining 14 abstracts (1.7%) are from non-English or non-French journals not available at McGill University or Université de Montréal. Given their small number, they can be safely neglected.

These references (a total of 61) were entered manually into an ASCII file. For the sake of consistency as well as for the purposes of the two programs used in the analysis (FILTER.EXE AND REPORT.EXE) the data were prepared using Format 4 of Dialog. The file of records for all the abstracts for 1987 were subjected to analysis by the two programs mentioned above. The results are shown in the tables below.

Table 3 below gives the author names as listed in rank order by the number of citations obtained.

Table 3
Top Cited Authors for 1987 in Rank Order

ABLOWITZ MJ	46
BREZIS H	38
HALE JK	33
KATO T	33
LIONS JL	32
MAWHIN J	32
GELFAND IM	31
ZAKHAROV VE	30
AMANN H	29
FOKAS AS	29
ARNOLD VI	26
COPPEL WA	24
GOHBERG IC	23
KRASNOSELSKII MA	22
ERDELYI A	21
SAMARSKII AA	21
ABRAMOWITZ M	20
GUCKENHEIMER J	20
HIRSCH MW	19
KANTOROVICH LV	19
DUNFORD N	18
HARDY GH	18
HARTMAN P	18
HELGASON S	18
KREIN MG	18
OLVER FWJ	18
RABINOWITZ PH	18
YOSHIDA H	18

There is no single name or group of names that dominates the list above. The most popular cited author is Ablowitz at 46 times. From a list of 829 articles containing well over 5500 citations, 46 (less than 1%) cannot be considered dominant. In addition, there is no group or focus that dominates the citations. Of the top seven authors who were cited more than thirty times, none has a single work or a small group of closely related works that are cited consistently. Ablowitz received 46 citations to 13 different works published over fourteen years. Brezis received thirty-eight citations to twenty-one works published over sixteen years. Hale has thirty-three citations to twenty-five works published over twenty-four years. Kato also received thirty-three citations, to fourteen works published over twenty-eight years, the first published in 1953. Lions obtained thirty citations to eleven works published over thirteen years. Mawhin received thirty-two citations to twenty-two different works published over seventeen years. Finally, Gelfand received thirty-one citations to nineteen works published between 1939 and 1980. Obviously, authors are citing these researchers for their cumulative work and their influence on their field as a whole rather than reacting to an influential work or a attractive idea. Some names, such as Hale, Gelfand, Arnold, Guckenheimer, Kantorovich and Yoshida are well known names in the study of nonlinear dynamics, but they cannot be said to be a dominant group (in terms of citations) in this cohort.

Table 4 below gives the names of corporate sources as listed in rank order by the number of contributions made.

Table 4
Top Contributing Corporate Sources for 1987 in Rank Order

BROWN UNIV,DIV APPL MATH	9
MV LOMONOSOV STATE UNIV	9
UNIV MONTREAL,CTR RECH MATH	7
CORNELL UNIV,DEPT MATH	6
TEL AVIV UNIV,SCH MATH SCI	5
UNIV MIAMI,DEPT MATH & COMP SCI	5
UNIV ROMA 2,DEPT MATH	5
VA STEKLOV MATH INST	5
ANDHRA UNIV,DEPT APPL MATH	4
CTR MATH & COMP SCI	4
ECOLE POLYTECH,CTR MATH APPL	4
ECOLE POLYTECH,CTR PHYS THEOR	4
FLINDERS UNIV S AUSTRALIA	4
INDIAN INST TECHNOL,DEPT MATH	4
INT CTR THEORET PHYS	4
MARATHWADA UNIV,DEPT MATH & STAT	4
N CAROLINA STATE UNIV,DEPT MATH	4
NO ILLINOIS UNIV,DEPT MATH SCI	4
PURDUE UNIV,DEPT MATH	4
UNIV CALIF DAVIS,DEPT MATH	4
UNIV DURHAM,DEPT MATH SCI	4
UNIV IDAHO,DEPT MATH & APPL STAT	4
UNIV IOANNINA,DEPT MATH	4
UNIV LONDON QUEEN MARY COLL,SCH MATH SCI	4
UNIV MANCHESTER,DEPT MATH	4
UNIV MICHIGAN,DEPT MATH	4
UNIV NICE,DEPT MATH	4
UNIV PARIS 06,ANAL NUMER LAB	4
UNIV PARIS 09 DAUPHINE,CEREMADE	4
UNIV PIERRE & MARIE CURIE	4
UNIV RHODE ISL,DEPT MATH	4
UNIV ROME LA SAPIENZA,DEPT MATEMAT	4
UNIV TENNESSEE,DEPT MATH	4
UNIV TORONTO,DEPT MATH	4
UNIV WARWICK,INST MATH	4

Once again, there is no institution that dominates the field. The top three contributors are Brown University - Division of Applied Mathematics, M.V. Lomonosov State University in Moscow, and the Centre de Recherche Mathématique at the Université de Montréal. In 97 cases (11.7%) no address was given. This is a large portion of the addresses that introduces a large uncertainty factor into the analysis. However, since there is no single author or group of authors that dominate(s) the epidemic, ultimately the matter is without consequence.

6.1.3. Summary

Work on function theory and analysis that comprises Section A02.30 of Chapter 02 reveals an outlier for 1987. However, analysis of citations and sources of authorship has not uncovered any dominant or influential work. Therefore, none of the hypotheses, as stated in Section 2.4, are supported.

6.1.4. Further Analysis

Despite the fact that the hypotheses are not supported and that there is no single influential work or a small group of influential works that dominate this outlier, there is nevertheless a dramatic surge in the number of articles abstracted in Section A02.30 for 1987. In addition, Figure 9 above shows that several other sections (Sections A02.10, A02.20, A02.40) within Chapter 02 with similar spikes in 1987. Thus, one wonders where this sudden surge is coming from and what is contributing to it. In other words, while there is smoke, where is the fire?

To answer the question, all the abstracts for 1987 were analyzed for descriptors. Physics Abstracts carries a controlled vocabulary, and descriptors in the records are assigned by experienced and dedicated indexers.[186] If there is a surge of records for 1987, the descriptors should reveal where the surge took place and which were the topics that benefited the most from the increase. Table 5 below gives the frequency of descriptors that occurred twenty or more times.

Table 5
Frequency of descriptors with twenty or more occurrences

functional analysis	243
differential equations	166
boundary-value problems	150
transforms	126
partial differential equations	106
nonlinear differential equations	104
functions	63
eigenvalues and eigenfunctions	47
integral equations	45
difference equations	41
nonlinear equations	41
matrix algebra	40
initial value problems	35
integration	34
topology	32
algebra	31
group theory	25
polynomials	24
linear differential equations	22
integro-differential equations	22
numerical methods	20

[186]Mrs. G.M. Wheeler, Editorial Director, INSPEC, letter to the author, 22 July 1993.

The 829 documents carry a total of 1879 descriptors of which 158 are unique. The table shows the twenty-one that account for 50% of the occurrences. Among those that occurred the most, six occur more than a hundred times. However, three of these six are closely associated, such that they can be grouped together. Differential equations, partial differential equations, and nonlinear differential equations can be grouped together and simply called "differential equations". Functional analysis and functions can also be grouped together as "functional analysis". Thus, the top seven descriptors that account for 958 occurrences (51%) can be grouped into: 1. differential equations (376 documents), 2. functional analysis (306 documents), 3. boundary-value problems (150 documents), and 4. transforms (126 documents).

Figure 11 below shows the occurrence of these descriptors month by month in Section A02.30. It is clear that the number of documents indexed by these descriptors increase and reach a maximum at the same time as the epidemic in Section A02.30. Given that the four altogether account for 51% of all the occurrences of descriptors, it is clear that they are in large part responsible for the epidemic in Section A02.30. As to the reason(s) for which the epidemic among such articles occurred in the first place, this must await future analysis. One possibility is the growth of work on nonlinear dynamics and chaos in the late 1970s and its spillover effect into work on nonlinear and partial differential equations.

The topic of Chapter A02 is "Mathematical Methods in Physics". The most appropriate index or database to start such a search would be *Mathematical*

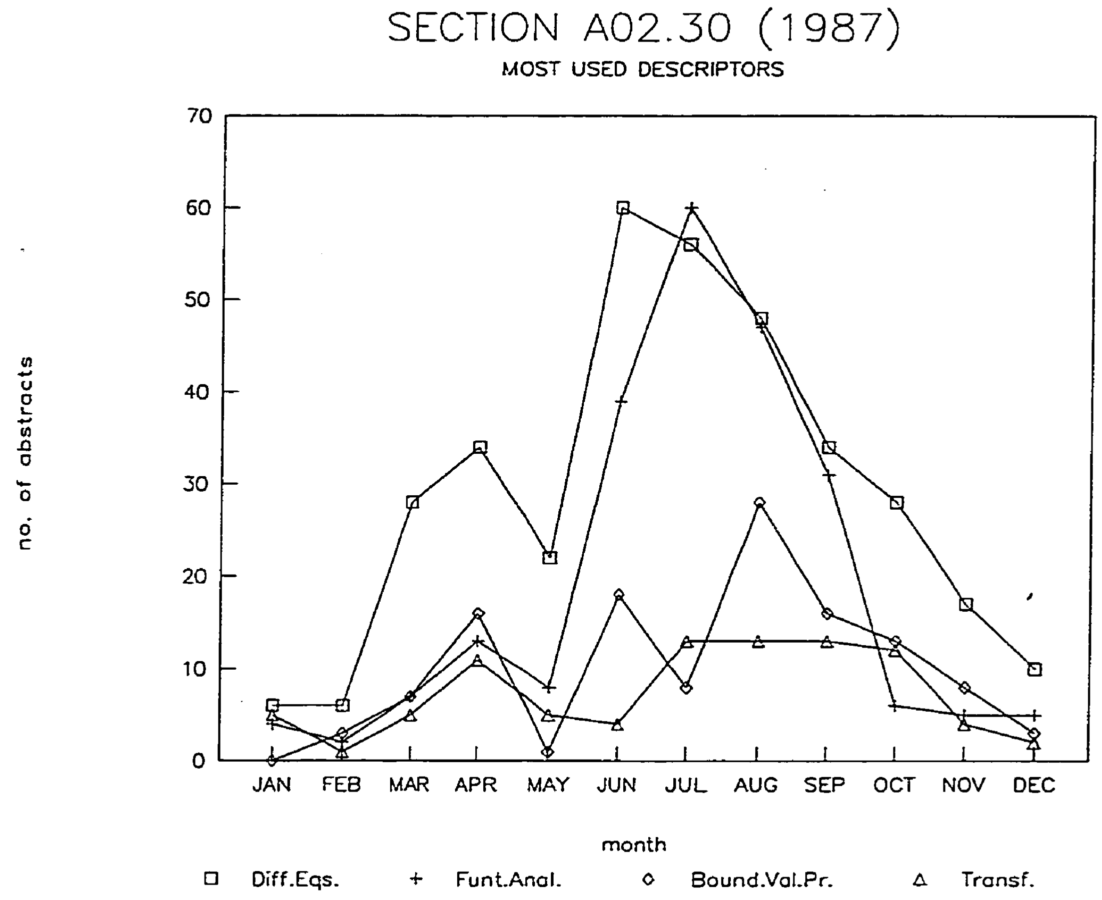


Figure 11. The four most used descriptors in Section A02.30 for 1987.

Reviews. The Reviews are also classified and provide a most suitable source for further work on information epidemics in mathematics.

6.2. Studies of Special Atoms and Molecules (Chapter 36)

6.2.1. Growth

This chapter is entitled *Studies of Special Atoms and Molecules*. It reveals three outliers in 1985 and 1987. From 1977 to late 1984 the activity in the field was rather flat. From 1985 on it shows a significant jump accompanied by several outliers. This is reflected in Figures 12 and 13. Figure 12 displays the monthly number of abstracts from 1977 to 1987 and marks the outliers. It also gives the regression line and the ± 3 standard deviation envelope. Figure 13 was prepared with the help of the shareware program NLREG (see above). It reflects the acceleration taking place in the studies on special atoms and molecules after 1984 accompanied by the exponential regression line and the $+3$ standard deviation envelope. This is also reflected in the 15 month moving average[187] in Figure 14 reflecting a smoothed curve but also the significant acceleration after 1984. Thus, the 4 criteria appropriate to data analysis (see Section 4.5, page 89 above) are met.

---

[187]The 15 month moving average is a curve where each point represents the arithmetic average of the preceding 15 months of activity. Its purpose is to smooth the original curve and reveal major tendencies in the data. The figure *15* here was obtained from the autocorrelation function following the SPSS analysis of the data of Chapter 36.

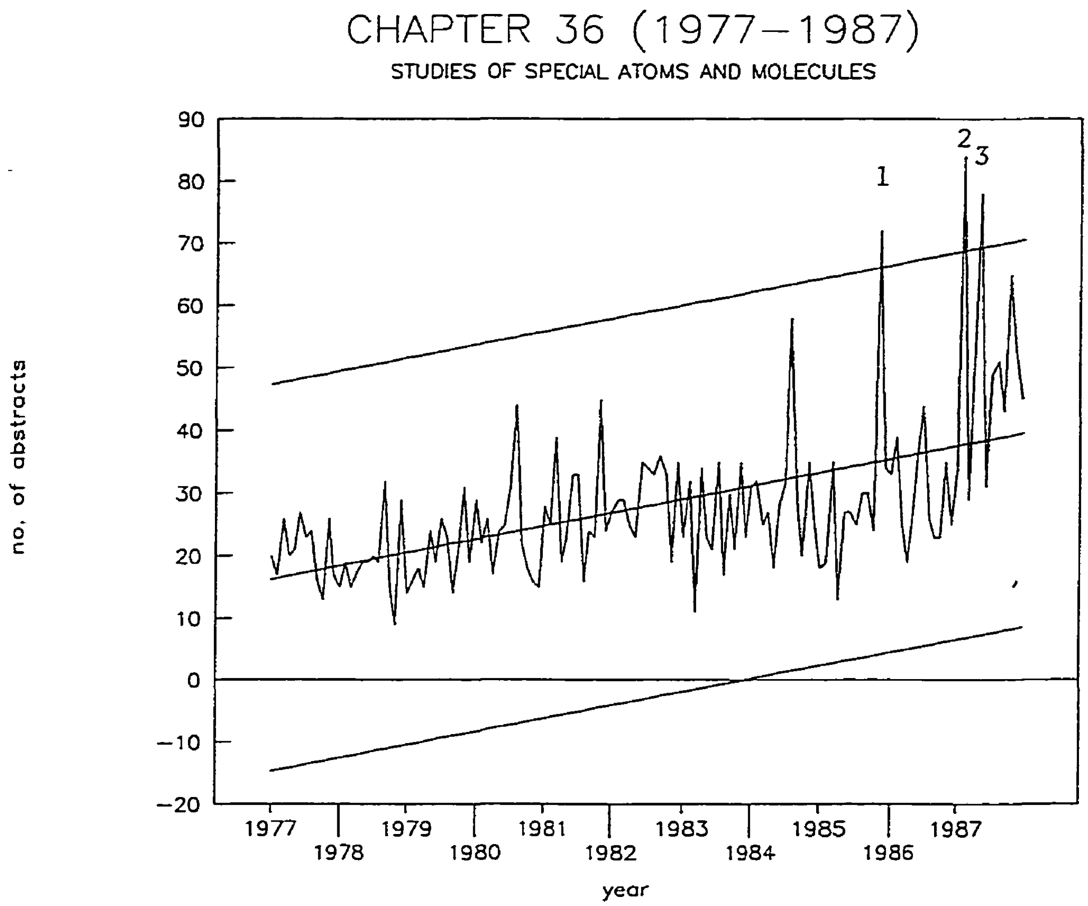


Figure 12. Monthly number of abstracts for Chapter 36 for 1977-1987.

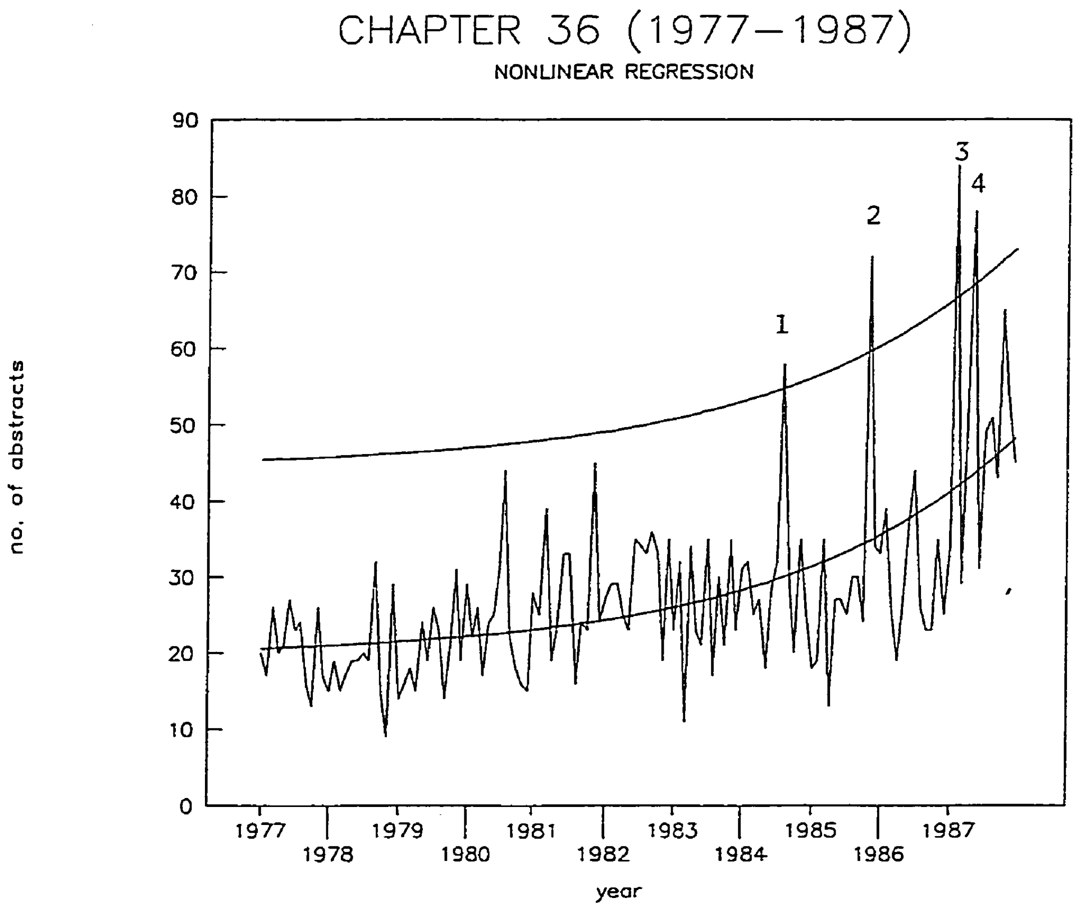


Figure 13. The growth of Chapter 36 with the nonlinear regression line and the accompanying +3.0 standard deviation envelope.

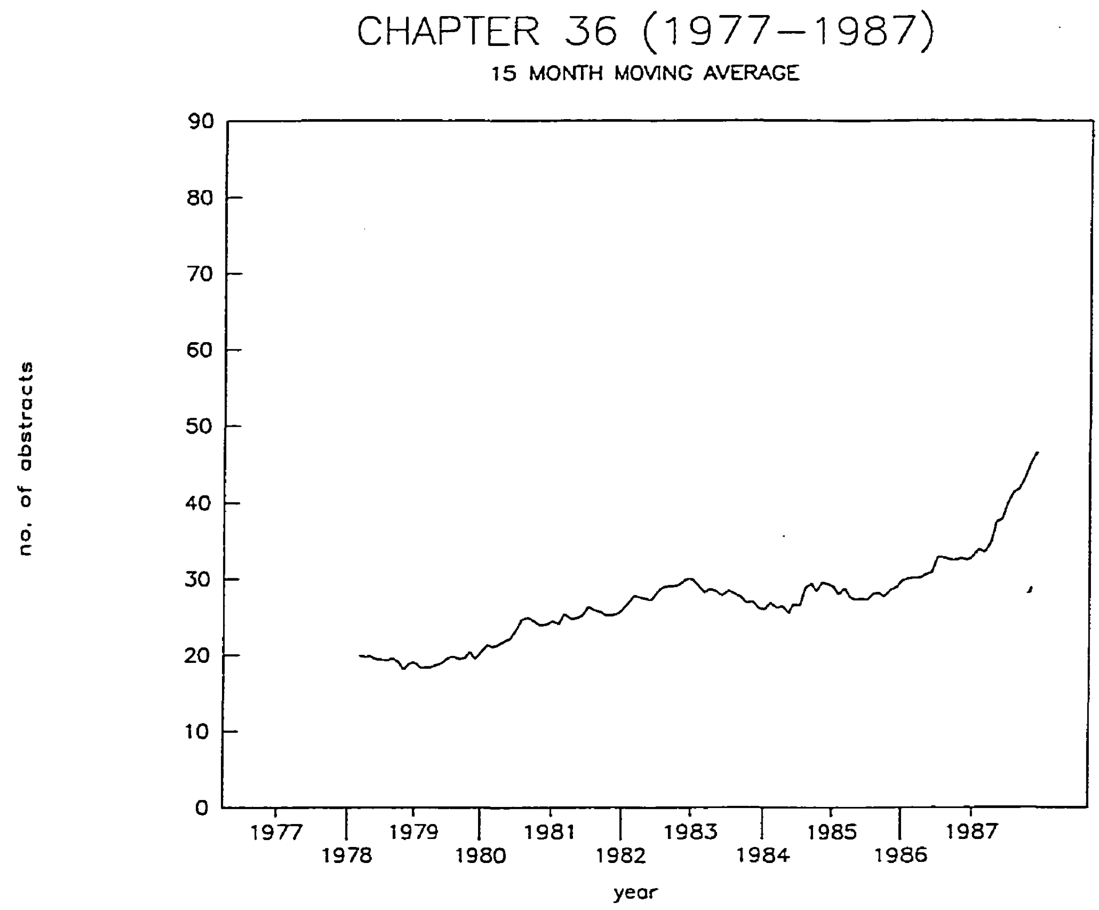


Figure 14. A fifteen month moving average of Chapter 36.

Chapter 36 consists of 3 sections:

A36.10: Exotic atoms and molecules
A36.20: Macromolecules and polymer molecules, and
A36.40: Atomic and molecular clusters.

Figure 15 reveals that the significant growth that took place between 1977 and 1987 can be largely attributed to Section A36.40. The other two sections reflect no acceleration at all. Therefore, the remainder of this analysis will be carried on Section A36.40 reflecting progress on cluster physics and chemistry in the 1980s.

Atomic and molecular clusters are small groupings of atoms usually numbering less than 100. Their properties are of interest because they have potential applications in photography, lubrication and catalysis. It is a field that is at the beginning of an exponential phase of growth, and continues to grow strongly to this day.

While only two of the three outliers in Chapter 36 come from the contributions of Section A36.40, there are two others within Section A36.40 that also cross beyond the +3 standard deviation threshold. These are shown in Figure 16, also prepared with the help of the shareware program NLREG. In this instance, the significance of the exponential regression line lies in the fact that the last spike in the graph (for October 1987) is no longer considered to be an outlier. The table below shows the considerable improvement in fit gained by the use of exponential regression over linear regression (obtained with SPSS). The R^2 values are considerably higher for both the whole chapter as well as the subsection under discussion, reflecting an improved fit to the data. The F statistics are also higher, reflecting an improved model over the linear one. Given the size of the outliers, the only difference comes

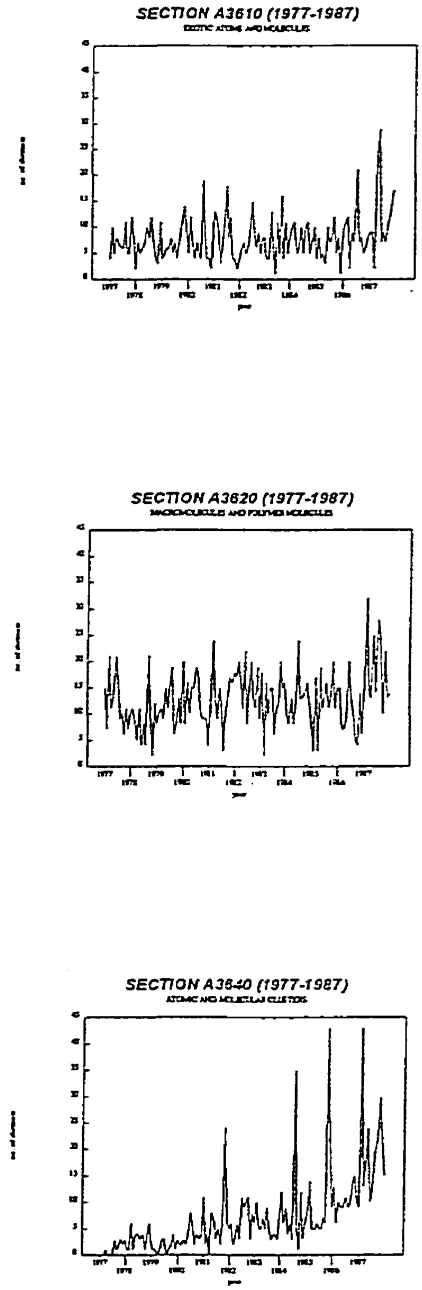
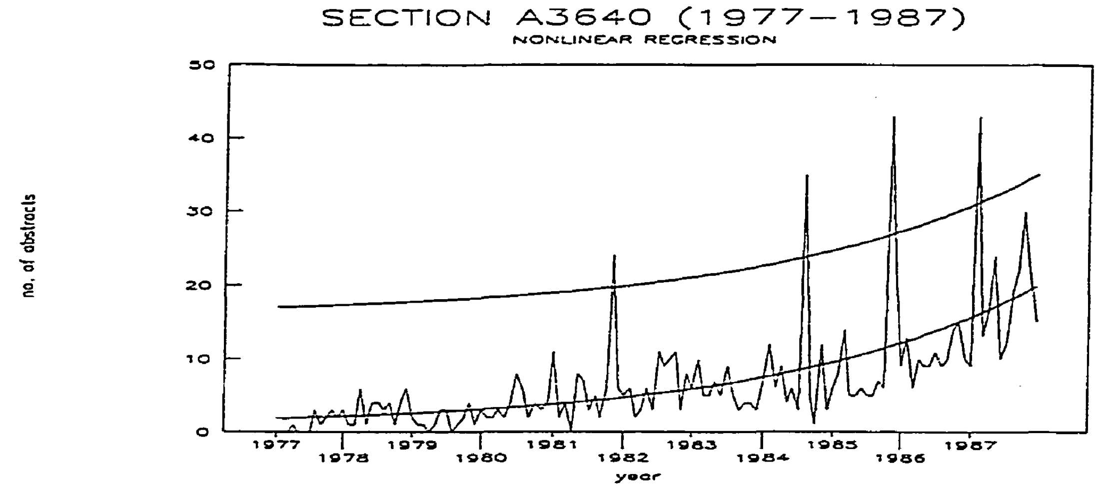
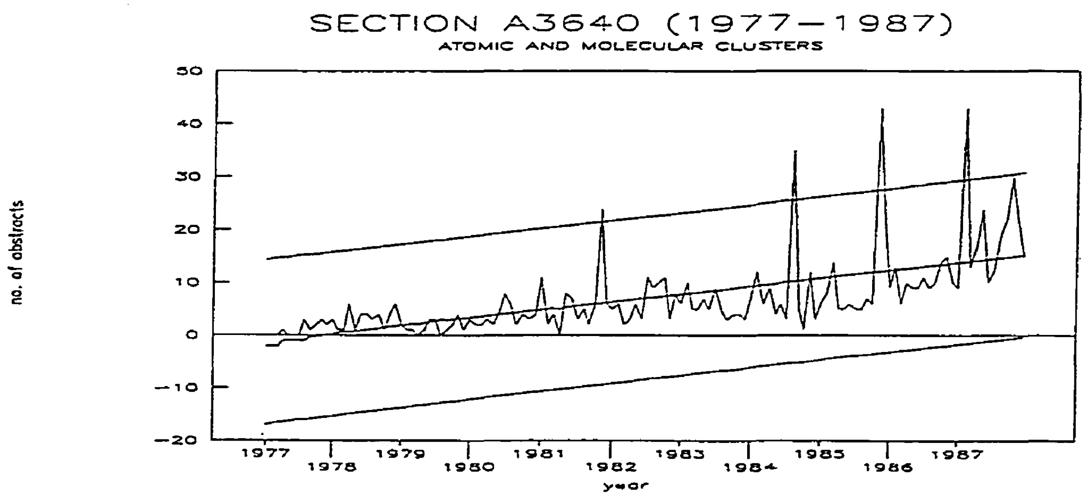


Figure 15. A comparison of the growth patterns of the three sections comprising Chapter 36.



a



b

Figure 16. The growth of Section A36.40: *a*, with the nonlinear regression line and the accompanying +3.0 standard deviation envelope; *b*, with linear regression and the accompanying ±3.0 standard deviation envelope. The contents of the outliers are listed in Table 7, page 151.

with the last spike (for October 1987) that is no longer considered significant once exponential regression is used.

Table 6
Comparison of the Statistics Between Linear and Nonlinear Regression

Parameters	Chapter A36	Section A36.40
R^2 (SPSS)	0.29861	0.38111
R^2 (NLREG)	0.3558	0.4267
F (SPSS)	55.34507	80.05496
F (NLREG)	71.80	96.76
Outliers (SPSS):	November 1985 February 1987 May 1987	November 1981 August 1984 November 1985 February 1987
Outliers (NLREG):	August 1984 November 1985 February 1987 May 1987	November 1981 August 1984 November 1985 February 1987

All of the outliers found belong to conference proceedings held on clusters: Lausanne in 1980, Konigstein in 1983, Berlin in 1984 and Heidelberg in 1986. In fact, the very first conference on clusters was held in Lyon in 1977 but it was not indexed by *Physics Abstracts*, possibly because Chapter 36 had just started in 1977 and because the conference was regarded to be more in the realm of chemistry and thus irrelevant to *Physics Abstracts*. The distinctive but insignificant spike is due to a special issue of the *Journal of Physical Chemistry* (v.91, no.11 for 21 May 1987) that was a Festschrift to a deceased colleague (Gilbert Stein).

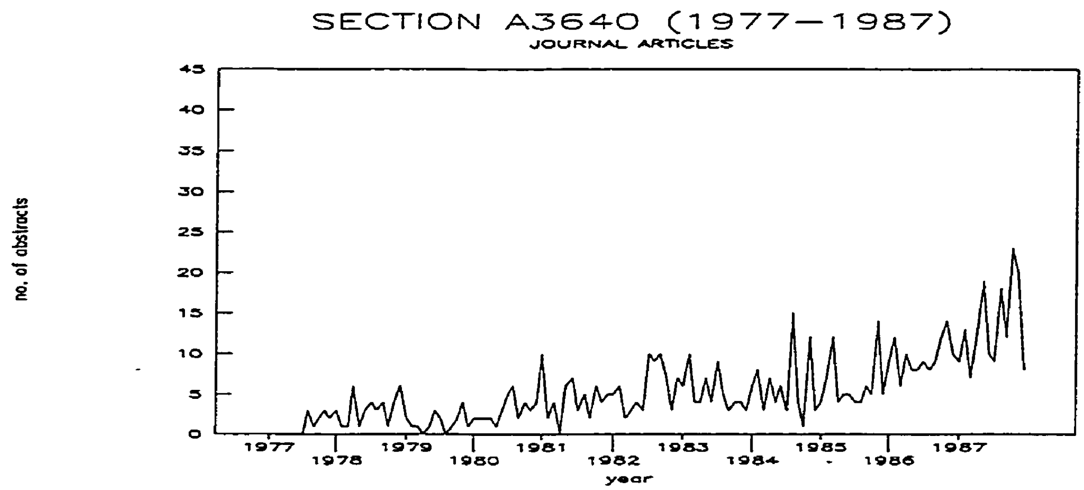
The details are given in Table 7.

Table 7
Data for the Outliers of Section A36.40

Outlier no.	PA issue date	Conference location and year held	No. of abstracts in PA	Conference papers abstracted
1	Nov.1981	Lausanne, 1980	24	19
2	Aug.1984	Konigstein, 1983	35	20
3	Nov.1985	Berlin, 1984	43	28
	Dec.1985		22	17
4	Feb.1987	Heidelberg, 1986	43	28

The numbers of papers and dates of conferences suggest that over time more conferences were held on this topic and more papers were given at each conference. Figures 17a and 17b contrast the growth of conference papers with the growth of journal papers. Figure 17b is a bar chart because papers from a conference form single entities in the time series whereas journal articles are abstracted continuously.

Overall, the growth of the field in large measure was sustained by a large and growing number of conferences. The last three years compared to the first three years in the data (1977-79 versus 1985-87) reveal a much larger number of conference papers indexed in Section A36.40 - not only from specialized conferences on clusters but also from the contributions of papers on clusters at other conferences. Thus, this



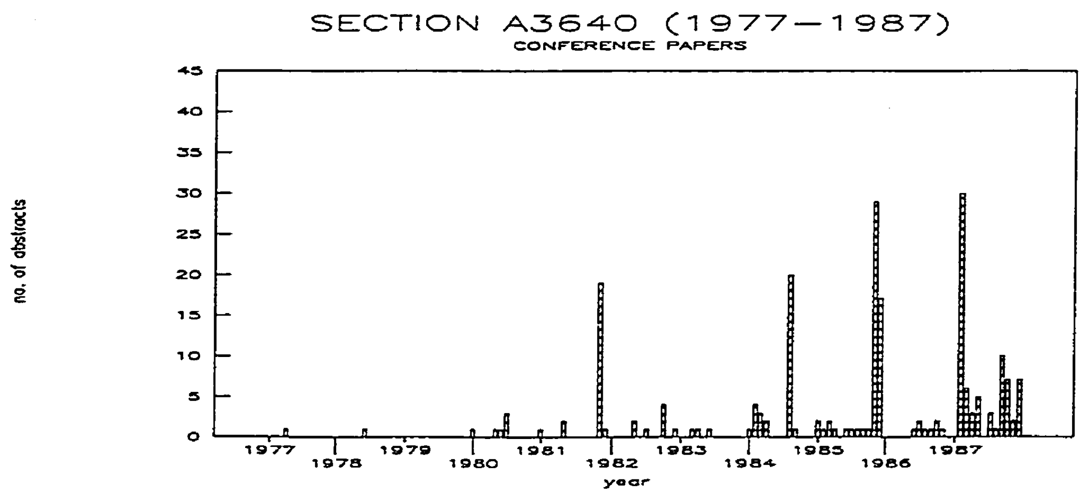


Figure 17. Comparing the production of journal articles with conference papers in the growth of Section A36.40.

evidence supports hypothesis 3b that as activity increases in an area of specialty so does the proportion of conference papers given in it.

6.2.2. Contents of outliers

If one is expecting to find an influential article as a precursor of an epidemic, it is to be expected that it will be found, by definition, in the reference lists of a large number of works making up that epidemic, including conference papers. In other words, if a given field is under the strong influence of a given work, this influence should be reflected in the citation patterns of conference papers as well. Since these proceedings were all published in journals, the records are available from the Science Citation Index database. The full records of all the abstracts were downloaded from SCI and the list of references of all the papers in the conferences were examined. Table 8 gives the author names as listed in rank order by the number of citations obtained.

There is no one name that dominates the table, but there are several active and productive authors that show up repeatedly. They have participated at most of the conferences and have also published several articles in the journal literature. Being so active, they also have been recognized by their peers and coworkers.

Do the authors in Table 8 represent any groupings by institutions or do they all work at different addresses? Table 9 lists the institutional origins of authors in rank order by the number of contributions to the conferences. There are clearly a number of addresses that stand out, most of which are of European origin. Only one of the

Table 8
Top Cited Authors by Conference in Rank Order[188]

LAUSANNE, 1980		KONIGSTEIN, 1983		BERLIN, 1984		HEIDELBERG, 1986	
INO S	11	HERRMANN A	14	MARTIN TP	21	MARTINS JL	15
BAETZOLD RC	10	HAGENA OF	11	SATTLER K	20	KNIGHT WD	13
HERRMANN A	9	SATTLER K	9	HERRMANN A	19	YAMADA I	12
JOHNSON KH	8	MOSKOVITS M	9	OZIN GA	14	ROHLFING EA	11
HOARE MR	7	LEUTWYLER S	8	SCHULZE W	13	WHETTEN RL	9
SOLLIARD C	7	AMIRAV A	7	ECHT O	12	KAPPES MM	9
SLATER JC	7	GOLE JL	7	MORSE MD	11	HERRMANN A	8
MASON MG	7	BECKMANN HO	7	MUHLBACH J	10	KOUTECKY J	7
GRANQUIST CG	7	BONDYBEY VE	6	MOSKOVITS M	10	PACCHIONI G	7
YANG CY	7	HABERLAND H	5	FARGES J	10	SHENG P	6
KIMOTO K	6	MOTT NF	5	BAETZOLD RC	10	PETERSON KI	6
LINDSAY DM	6			HABERLAND H	10	HENKES W	5
SALAHUB DR	6			MESSMER RP	9	GEUSIC ME	5
BUFFAT P	6			RILEY SJ	9	DELLEY B	5
CASTLEMAN AW	6			HOARE MR	9	STEPHAN K	5
GORDON MB	6			HUBER KP	9		
BECKER EW	6			EKARDT W	9		
VOSTRIKOV AA	5			RUPPIN R	8		
DESJONQUERES MC	5			STACE AJ	8		
COUCHMAN PR	5			COTTON FA	8		
MESSMER RP	5			MARKS LD	8		
FARGES J	5						
MUHLBACH J	5						
YOKOZEKI A	5						
YACAMAN MJ	5						
HAGENA OF	5						

[188]Only authors with more than four citations are shown.
LAUS = Lausanne, 1980; KONG = Konigstein, 1983; BERL = Berlin, 1984; HEID = Heidelberg, 1986.

Table 9
Top Contributing Corporate Sources by Conference in Rank Order[189]

	LAUS	KONG	BERL	HEID	TOTAL
ECOLE POLYTECH FED LAUSANNE	8	2	11	4	25
FREE UNIV BERLIN	0	5	12	1	18
MAX PLANCK GESELL,FRITZ HABER INST/ BERLIN	0	4	11	0	15
UNIV PARIS 11	6	1	7	0	14
KERNFORSCHUNGSZENTRUM, KARLSRUHE	4	1	2	2	9
UNIV CONSTANCE,FAK PHYS	2	1	6	0	9
MAX PLANCK INST FESTKORPERFORSCH/ STUTTGART	1	1	6	1	9
IBM CORP,ALMADEN RES CTR	0	1	2	3	6
KFA JULICH GMBH,INST FESTKORPERFORSCH	0	1	4	1	6
TECH UNIV DENMARK,APPL PHYS LAB	2	0	4	0	6
UNIV SAARLAND,FACHBEREICH PHYS/ SAARBRUCKEN	1	0	4	1	6
INST RECH CATALYSE/ VILLEURBANNE//FRANCE	4	0	1	0	5
STANFORD UNIV	2	0	2	1	5
UNIV FREIBURG,FAK PHYS	0	2	3	0	5
UNIV KARLSRUHE	2	2	1	0	5
PENN STATE UNIV, DEPT CHEM	0	2	1	1	4
UNIV BERN INST ANORGAN ANALYT & PHYS CHEM	1	1	1	1	4
UNIV HAMBURG,INST PHYS CHEM	2	0	1	1	4
UNIV STUTTGART,INST THEORET CHEM	1	1	2	0	4

[189]Only institutions with four or more contributions shown.

top 10 names is of American origin: IBM Corporation. Of the others, six are German, and one each are from France, Switzerland and Denmark. There are in total 145 addresses on Table 9 accounting for 342 contributions.[190] The top ten account for about one third (34.2%) of the contributions and about one sixth (15.9%) of them (23) account for 50% of the contributions. Of those, 16 are European addresses, 6 are American and one is Japanese. One must then conclude that cluster work between 1977 and 1987 is a field largely centered in Europe. In fact, it started and grew in Europe, all of its initial conferences were held in European cities, and all the proceedings were published in European journals.

Does the number of citations obtained by the authors increase as the years go by? Table 8 also indicates that this is in fact what happens. Lausanne was the location for the second *International Conference on Small Particles and Inorganic Clusters* (ISPIC), and Berlin for the third. Table 8 indicates that a larger number of authors at Berlin obtained more citations than those at Lausanne. Konigstein 1983 and Heidelberg 1986 were special conferences held by invitation only; they exhibit lower citation counts than Berlin, but the Heidelberg citations are higher than those at Konigstein. Thus, there is a marked difference between three years of activity.

---

[190]In reality, the combined number of papers from the four conferences is lower than 342. In the case of coauthorships from different institutions, each address was counted as a separate contribution.

6.2.3. Summary

Cluster work between 1977 and 1987 shows exponential growth marked by four conferences held respectively in Lausanne (1980), Konigstein (1983), Berlin (1984) and Heidelberg (1986). These four outliers are characterized by contributions mostly from Europeans, but none of them reflects any dominance from a single influential work. On the other hand, the dominance of a group of work on the fractional quantum Hall effect is unmistakable. With respect to the hypotheses in this thesis:

Hypothesis 2: is partially supported; there is no single influential work, but a group of works dominates.

Hypothesis 3a: is partially supported; there is no single influential work, but a group of works dominates.

Hypothesis 3b: is supported (see Figure 17b); as the field grows, so does the proportion of conference papers.

The evidence lends support to the epidemic hypothesis in so far as it reflects the early, exponential phase of growth.

6.3. Electronic Structure and Electrical Properties of Surfaces, Interfaces, and Thin Films (Chapter 73)

6.3.1. Growth

This chapter is entitled *Electronic Structure and Electrical Properties of Surfaces, Interfaces, and Thin Films*. The data reveal only one outlier from late 1985 that is due to a relatively small number of contributions (a total 13.5% of the abstracts) from two conferences (see Figure 18). The overall activity in the chapter shows no tendency until mid-1985 after which it accelerates significantly. This can also be discerned from Figure 19 and 20. Figure 19 shows the exponential regression line and the +3 standard deviation envelope. Figure 20 gives the 5-month moving average. This smoothed curve also shows the dramatic rise that took place since 1984. Overall, then, this chapter also meets the four criteria for data analysis.

Chapter 73 consists of six sub-chapters and 15 sections within them. Most of the subsections contain too little data (less than 100 abstracts per year) to be worthy of attention. However, the following five sections are large and active enough (i.e. they contain more than 100 abstracts per year) that they have been analyzed further below:

73.20:	Electronic surface states,
73.40L:	Semiconductor-to-semiconductor contacts, p-n junctions, and heterojunctions,
73.40N:	Metal-nonmetal contacts,
73.40Q:	Metal-insulator-semiconductor structures, and
73.60F	Semiconductor films.

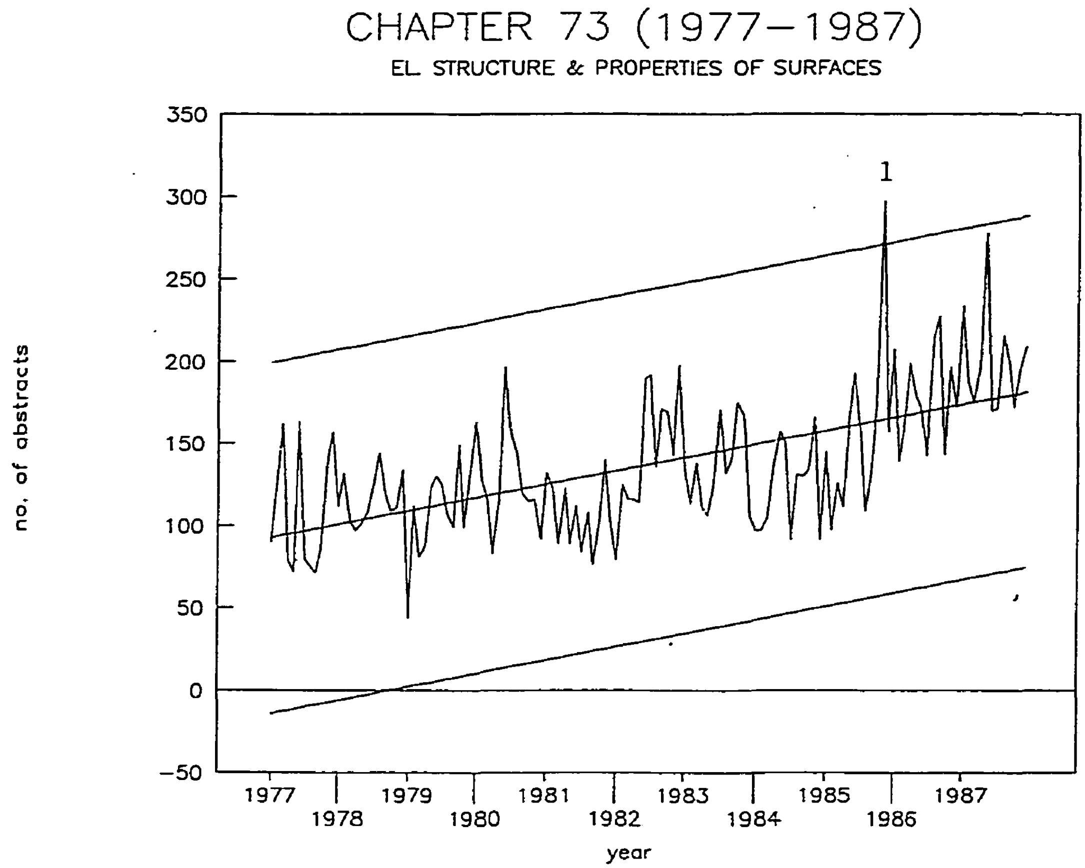


Figure 18. Monthly number of abstracts for Chapter 73 for 1977-1987.

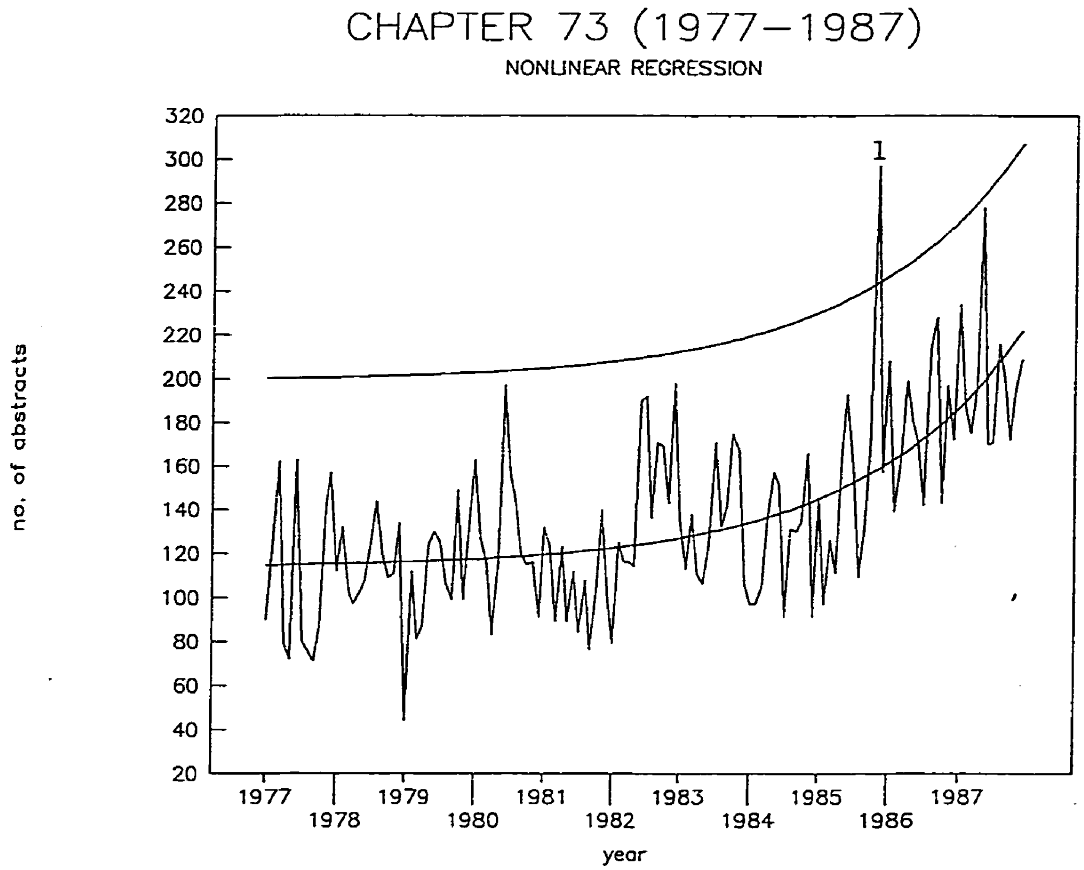


Figure 19. The growth of Chapter 73 with the nonlinear regression line and the accompanying +3.0 standard deviation envelope.

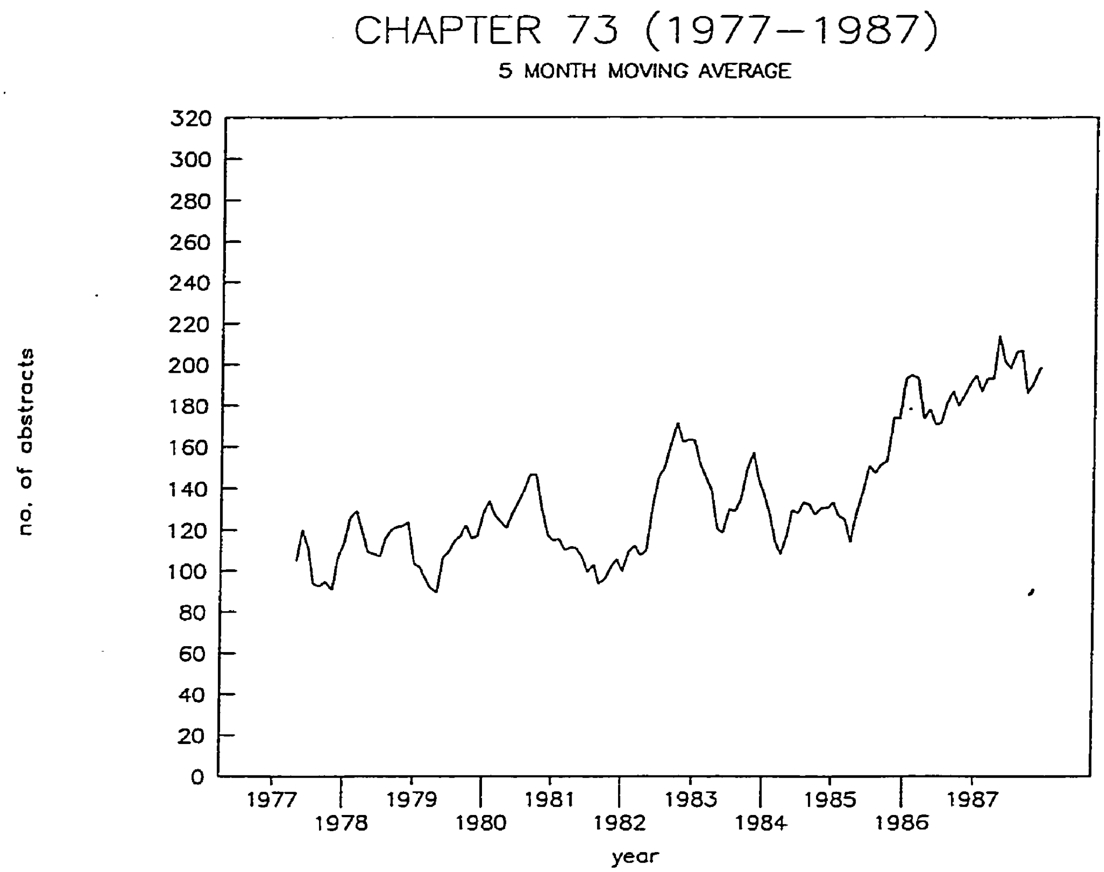


Figure 20. A five month moving average of Chapter 73.

The growth of these sections is illustrated in Figure 21 below. Of these only Section 73.40L shows a clear and significant acceleration pattern during 1977-1987. The others over the eleven year period are flat, with regression coefficients at zero. The rest of the analysis for this chapter will therefore be conducted on Section 73.40L.

Section 73.40L includes material on *semiconductor-to-semiconductor contacts, p-n junctions and heterojunctions, including superlattices and quantum dots, wells and wires.*[191] It was introduced into the classification in 1973. The topic is a subspecialty of solid state physics and pertains to semiconductor materials and the physical properties of the different layers of materials that make them up. Work on these materials affects the commercial applications and the performance of devices such as infrared detectors, microwave amplifiers, computer processors, laser diodes and high-efficiency solar cells. Heterojunctions refer to *the interfaces between different semiconducting materials.*[192] The section carries abstracts on work dealing with the difficulties and complexities of exploiting an increasing number of thinner layers of semiconducting materials sandwiched on the surface of a device and the physical properties exhibited by that device.

---

[191]*INSPEC Classification*, 1992, 50.

[192]Robert S. Bauer and G. Margaritondo, "Probing Semiconductor-Semiconductor Interfaces," *Physics Today* 40 (1987): 27-34.

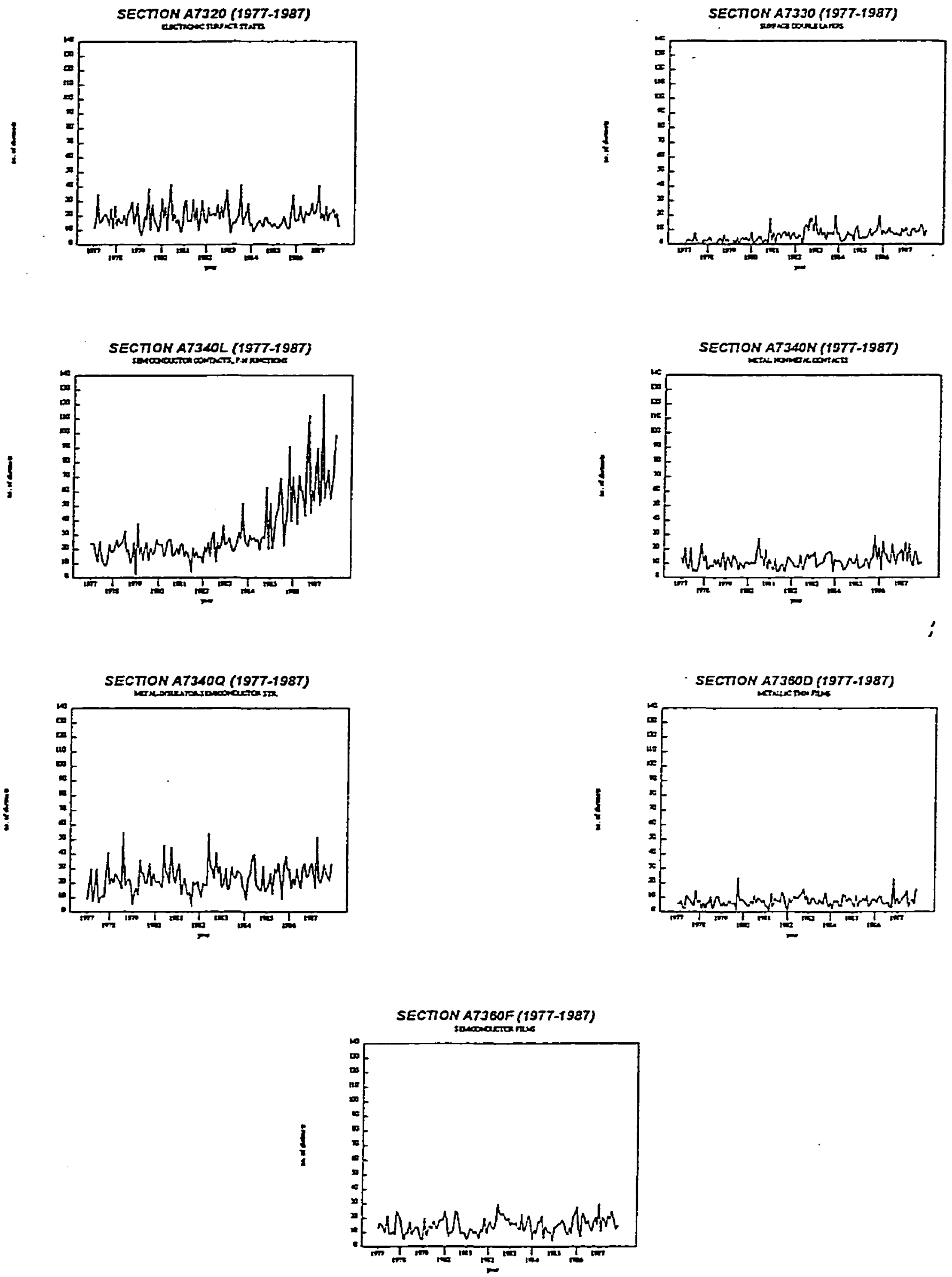
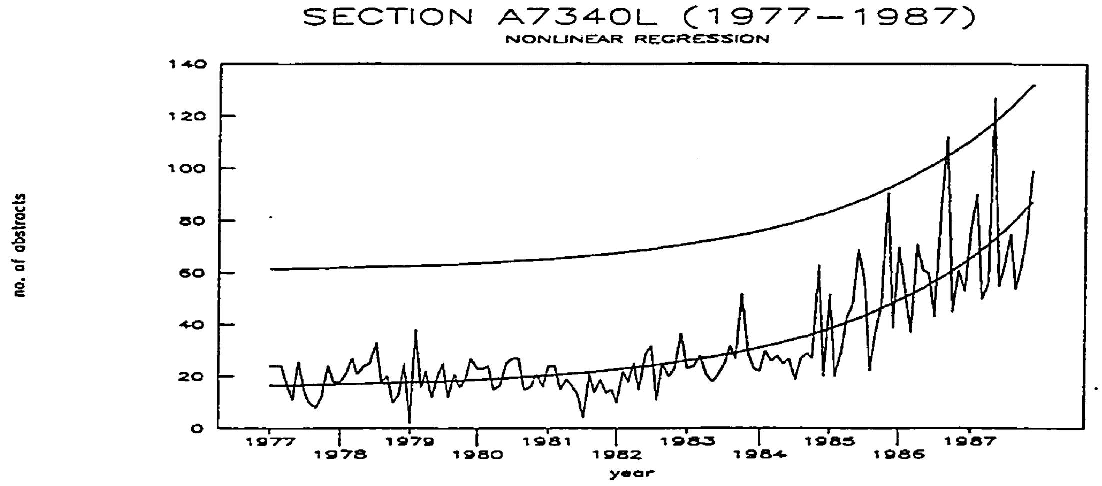


Figure 21. Comparison of the growth patterns for the sections comprising Chapter A73.



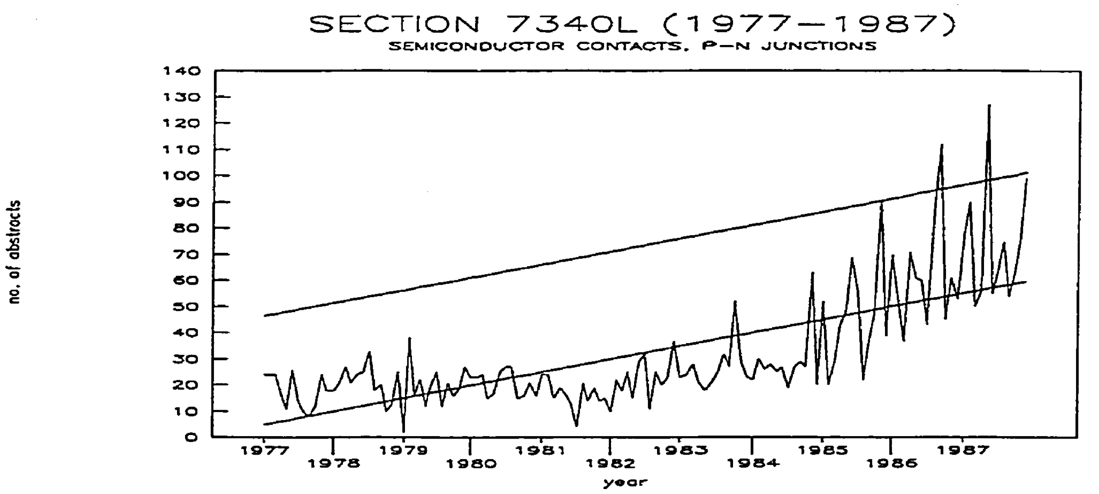


Figure 22. The growth of section A73.40L: *a*, with the nonlinear regression line and the accompanying +3.0 standard deviation envelope; *b*, with linear regression and the accompanying ±3.0 standard deviation envelope. The contents of the outliers are listed in Table 11, page 167.

Figure 22 above reveals an increasing number of outliers with an increasing volume of abstracts. The graph illustrates the data, an exponential regression line (prepared with the use of NLREG) and the +3 standard deviation envelope. There are two outliers that lie beyond the +3 standard deviation line, both of them largely due to conferences. Their contents are further analyzed below. It is also noteworthy that one of the outliers (for November 1985) identified as significant with linear regression is no longer considered so with the use of nonlinear regression.

Table 10 below once again displays the considerable improvement in fit gained by the use of nonlinear regression over linear regression. The R^2 values are considerably higher for both the whole of Chapter 73 as well as the subsection 73.40L under discussion, reflecting an improved fit to the data. The F statistics are also higher, reflecting an improved model over the linear one.

Table 10
A comparison of statistics between linear and nonlinear regression

Parameters	Chapter A73	Section A73.40L
R^2 (SPSS)	0.36499	0.50466
R^2 (NLREG)	0.4239	0.6748
F (SPSS)	74.72048	132.44653
F (NLREG)	95.64	269.77
Outliers (SPSS):	November 1985	November 1985 September 1986 May 1987
Outliers (NLREG):	November 1985	September 1986 May 1987

Both of the outliers belong to conference proceedings held on superlattices and heterostructures: San Francisco in 1984, Kyoto in 1985, and Goteborg and Mauterndorf in 1986. The November 1985 outlier that turns out to be non-significant after the nonlinear regression is accounted for by the indexing of all the four issue of the new journal *Superlattices and Microstructures*, volume 1, issue nos. 1-4 for 1985. It is the first time the journal was being covered by Physics Abstracts. The putative outlier therefore is a spurious result introduced by an infrequent procedure.

The details are given in Table 11 below.

Table 11
Data for the outliers of Section 73.40L

Outlier no.	PA issue date	Conference location and year held	No. of abstracts in PA	Conference papers abstracted
1	Aug.1986	San Francisco, 1984[193]	87	47
	Sep.1986	Kyoto, 1985[194]	112	47
2	May 1987	Goteborg, 1986[195]	127	11
		Mauterndorf, 1986[196]		17
		Tokyo, 1986[197]		12

In support of hypothesis 3b, Figures 23a and 23b contrast the growth of conference papers to the growth of journals. Again, the growth of the field was in large measure sustained by a large and growing number of conferences. As the field developed and the number of contributions increased, so did the number of conference submissions. The last three years compared to the first three years in the data (1977-79 versus 1985-87) reveal a much larger number of conference papers indexed in

---

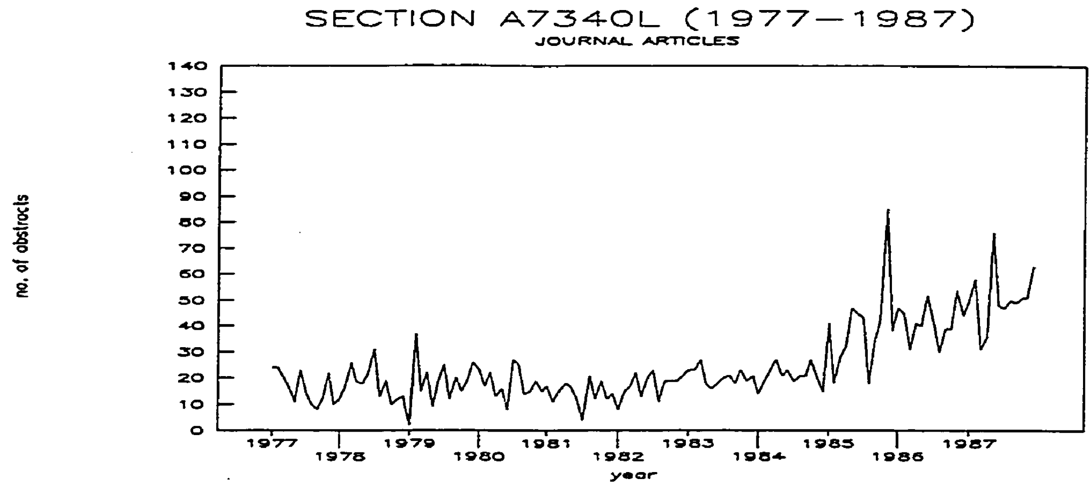
[193]*International Conference on the Physics of Semiconductors*, 17th, San Francisco, August 1984.

[194]*Yamada Conference XIII: Electronic Properties of Two-Dimensional Systems*, Kyoto, September 1985 (published in *Surface Science*, v.170, no.1-2, 1986).

[195]*International Conference on Superlattices, Microstructures and Microdevices*, 2nd, Goteborg, August 1986 (published in *Superlattices and Microstructures*, v.2, no.5, 1986).

[196]*International Winter School on Two-Dimensional Systems: Physics and New Devices*, Mauterndorf, Austria, February 1986.

[197]*International Conference on Solid State Devices and Materials*, 18th, Tokyo, August 1986.



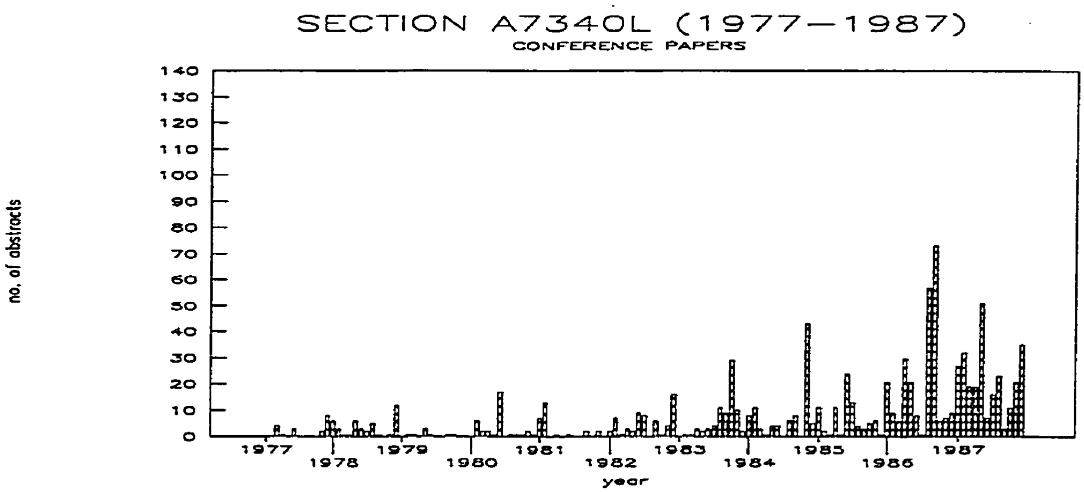


Figure 23. Comparing the production of journal articles with conference papers in the growth of Section A36.40.

Section 73.40L. A detailed look at other non-significant spikes with large conference papers also supports this contention, even if Table 11 does not necessarily reflect it clearly.

6.3.2. Contents of outliers

Similar to the results of Chapter 36 above, there is no single author that dominates the list. However, there are several highly productive and highly cited authors that show up consistently. As Tables 12 and 13 below show, they belong to a few institutions that dominate the field.

Table 12
Top Cited Authors by Conference in Rank Order[198]

SAN FRANCISCO, 1984		KYOTO, 1985		MAUTERNDORF, 1986	
STORMER HL	12	ANDO T	22	KROEMER H	9
TSUI DC	12	DASSARMA S	14	ANDO T	7
ANDO T	9	STORMER HL	11	BASTARD G	6
MILLER RC	9	EISENSTEIN JP	10	LAUGHLIN RB	6
BASTARD G	7	LAUGHLIN RB	10	TERSOFF J	6
CHANG LL	7	BROIDO DA	9	DINGLE R	5
MENDEZ EE	7	BANGERT E	7	MILLER RC	5
DINGLE R	5	EBERT G	7	SHAH J	5
WEISBUCH C	5	EKENBERG U	7	WANG WI	5
		ALTARELLI M	6		
		GULDNER Y	6		
		MILLER RC	6		
		VONKLITZING K	6		
		GIRVIN SM	5		
		GREENE RL	5		
		HORST M	5		

[198]Only authors with more than four citations are shown. None of the authors at the Goteborg conference of 1986 obtained more than four citations.

Table 13
Top Contributing Corporate Sources by Conference in Rank Order[199]

	SF	KY	GB	M2	TOTAL
AT&T BELL LABS, MURRAY HILL, NJ	8	11	0	0	19
IBM WATSON RES CENTER, YORKTOWN HEIGHTS	8	3	2	1	14
MAX PLANCK INST FESTKORPERFORSCH	1	3	1	4	9
BELL COMM RES INC, MURRAY HILL	8	0	0	0	8
FORSCHINST DEUT BUNDESPOST, DARMSTADT	0	5	0	2	7
THOMSON CSF,CENT RECH LAB	1	4	2	0	7
UNIV OXFORD,CLARENDON LAB	1	4	0	2	7
MIT, FRANCIS BITTER NAT MAGNET LAB	4	2	0	0	6
UNIV TOKYO, DEPT APPL PHYS	1	4	0	0	5
BROWN UNIV,DEPT PHYS	2	2	0	0	4
CTR NATL ETUD TELECOMMUN, PARIS	0	2	2	0	4
INST NATL SCI APPL LYON,PHYS SOLIDES LAB	0	2	2	0	4
PRINCETON UNIV,DEPT ELECT ENGN	0	4	0	0	4
UNIV PARIS 07	1	2	0	1	4
UNIV TOKYO, INST SOLID ST PHYS	3	1	0	0	4
UNIV TOKYO, RES INST IND SCI	2	2	0	0	4

[199]Only institutions with four or more contributions are shown.

SF=San Francisco, 1984; KY=Kyoto, 1985; GT=Goteborg, 1986; M2=Mauterndorf, 1986.

Of the top ten corporate sources, five are American, two are German, and one each comes from France, Great Britain and Japan. The top positions are dominated by corporate research centers: AT&T in the first place, IBM second, Bell Communication Labs fourth, and Thomson CSF of France in fifth place. There are a total ninety four addresses that made 223 contributions to the four conferences at issue here.[200] Of those, the top ten account for 39% of the contributions. Seventeen account for 50% of the contributions. Of these, six are American addresses, four are French, three are Japanese, two are German, and one each comes from Great Britain and Japan. In other words, the work is shared equally by sources of American and European origin.

It is difficult with these data to state unequivocally whether the number of conference papers increased with time. However, a cursory analysis of several other spikes that were not significant revealed that over the eleven years from 1977 to 1987 an increasing number of conferences were held and a larger number of papers were given.

Table 12 does not reveal whether the number of citations continued to grow with the passage of time. It is made out of four conferences bunched over the last two years in the data set. Previous years' data, on the other hand, did not produce outliers and therefore did not contribute any useful information. However, the analysis of

---

[200]As above, the combined number of papers from the four conferences is lower than 223. In the case of coauthorships from different institutions, each address was counted as a separate contribution.

some non-significant spikes showed that while the total number of conference papers went up (see also Figure 23b), the number of citations obtained by the authors were as high in 1983 or 1984 as they were in 1986 or 1987. Thus, the evidence here does not allow one to state that the number of citations obtained by its authors grows concomitantly with their field of specialty.

6.3.3. Summary

Progress on heterostructures and superlattices shows exponential growth between 1977 and 1987, marked by an increasing number of conference papers originating from an increasing number of institutions. The two outliers are due to two conferences each: San Francisco and Kyoto making up the outlier for August-September 1986 and the Goteborg and Mauterndorf conferences making up the outlier for May 1987. With respect to the hypotheses:

Hypothesis 2: is partially supported; there is no single influential work, but a group of work dominates.

Hypothesis 3a: is partially supported; there is no single influential work, but a group of works dominates.

Hypothesis 3b: is supported (see Figure 23b); as the field grows, so does the proportion of conference papers.

Overall, the evidence lends support to the epidemic hypothesis in so far as it reflects the early, exponential phase of growth.

6.4. Superconductivity (Chapter 74)

6.4.1. Growth

Superconductivity is a phenomenon where metals at temperatures close to absolute zero lose all their resistance to electrical current and become perfect conductors. They have potential applications in many areas, including high voltage electrical transmission, computer hardware, and medical imaging technology. Due to the importance of finding metals that are perfect conductors, superconductivity has been one of the consistent and major research specialties of the century since its discovery in 1911. While progress was made in finding materials that become superconductive at increasingly higher temperatures, the highest temperature (the critical temperature) achieved until 1986 remained around 23° Kelvin.

In late 1986 a paper by Bednorz and Muller reported superconductivity at 30° Kelvin with a new family of materials that was based on a mixture that included copper oxides.[201] Due to the unexpected nature of the materials and the crossing of the 23° barrier, the paper opened a new approach to the study of superconductivity. The main reasons for the popularity are that *the materials are easy to fabricate, they have high critical temperatures, they represent a challenge to theoreticians, and they are of considerable technological importance.*[202] Figure 24, borrowed from Muller

---

[201]Bednorz and Muller, *Possible High Tc Superconductivity.*

[202]K. Alex Muller and J. George Bednorz, "The Discovery of a Class of High-Temperature Superconductors," *Science* 237 (1987): 1133-9.

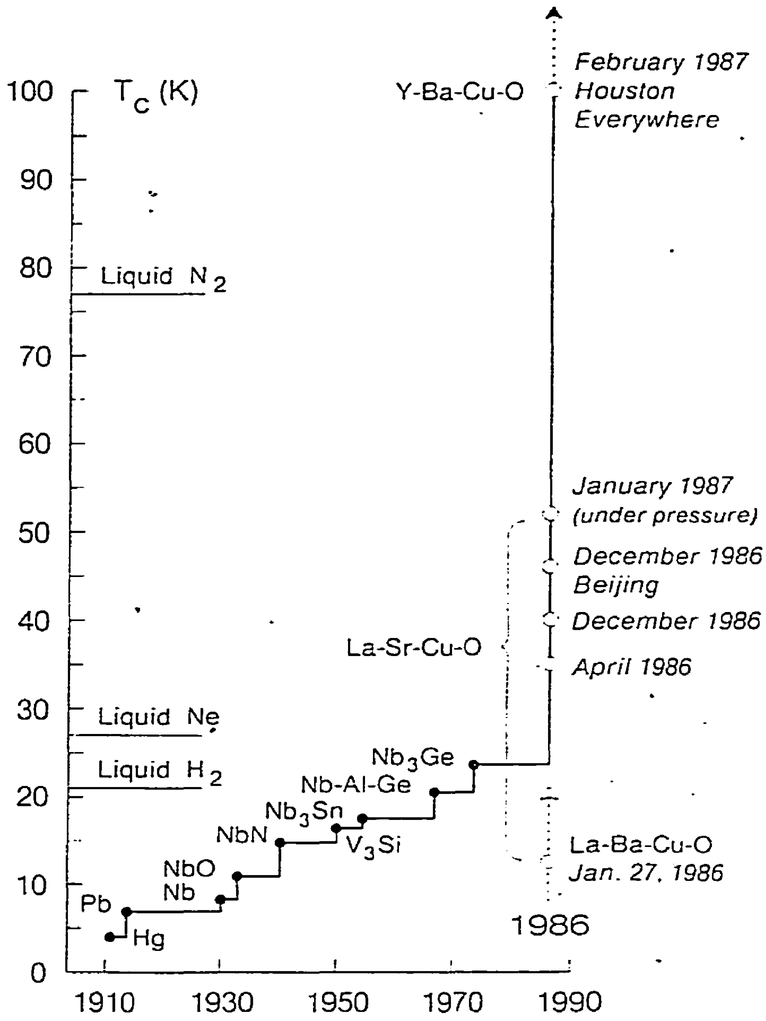


Fig. 1. Evolution of the superconductive transition temperature subsequent to the discovery of the phenomenon.

Figure 24. Evolution of the superconductive transition temperature subsequent to the discovery of High Temperature Superconductivity (from: Muller and Bednorz, 1987).

and Bednorz shows the history of attaining higher temperatures until 1987.[203] Achieving higher temperatures is critical because the cooling process is both cumbersome and rather expensive. The less cooling the materials need, the greater the potential for achieving perfect conductance at lower cost.

The exceptional interest in high-temperature superconductivity (HTSC) has been reflected in its literature. The short-term growth of articles indexed in Chapter 74 for 1977-1987 is depicted in Figure 25. As the figure shows, the time series for the chapter displays several spikes with growing intensity. Of the four major spikes, the first three are mainly due to the proceedings of the *International Conference on Low Temperature Physics* no.15-17, for 1978, 1981, and 1984. The outlier of 1987 is due to the dramatic growth of the superconductivity literature in reaction to the Bednorz and Muller paper and is mostly accounted for by the journal literature. The last two outliers (1985 and 1987) are important and will be analyzed below.

Chapter 74 is divided into the following sections:

A74.10: Occurrence, critical temperature
A74.20: Theory
A74.30: General properties
A74.40: Fluctuations and critical effects
A74.50: Proximity effects, tunnelling phenomena, and Josephson effect
A74.55: Type-I superconductivity
A74.60: Type-II superconductivity
A74.65: Insulator-superconductor transition
A74.70: Superconducting materials
A74.75: Superconducting films
A74.90: Other topics in superconductivity

---

[203]Ibid., 1134.

Figure 25. Monthly number of abstracts for Chapter 74 for 1977-1987.

Figure 26 shows the comparison of the sections comprising Chapter A74. Especially in 1987, when the critical growth took place, all sections have exhibited dramatic jumps. Spikes (or growth points) can be seen in several sections over the years, but all (that is in 1985 for Section A74.30 and 1979, 1982, and 1985 for Section A74.50) other than those in 1987 are due to conference proceedings. It is also revealing to see that some of the sections, such as A74.40 or A74.55 hardly have any data to display and that in all likelihood they reflect an interest or preoccupation of years gone by. This in a way demonstrates that it is easier to add to a classification schedule than to remove from it and that probably most current classifications support preoccupations of the past. Under these circumstances, it is not at all unusual to see that not all of the sections are showing dramatic jumps for 1987.

As to the moving average, given that the ACF=1, that the growth for 1987 is dramatic and that the outliers are clearly significant, there is no need to include a graph for it.

6.4.2. Contents of the outliers

As mentioned above, the first outlier (in 1985) is largely accounted for by the proceedings of the 17th *International Conference on Low Temperature Physics* in 1984. The second outlier from 1987 is accounted for largely by journal articles.

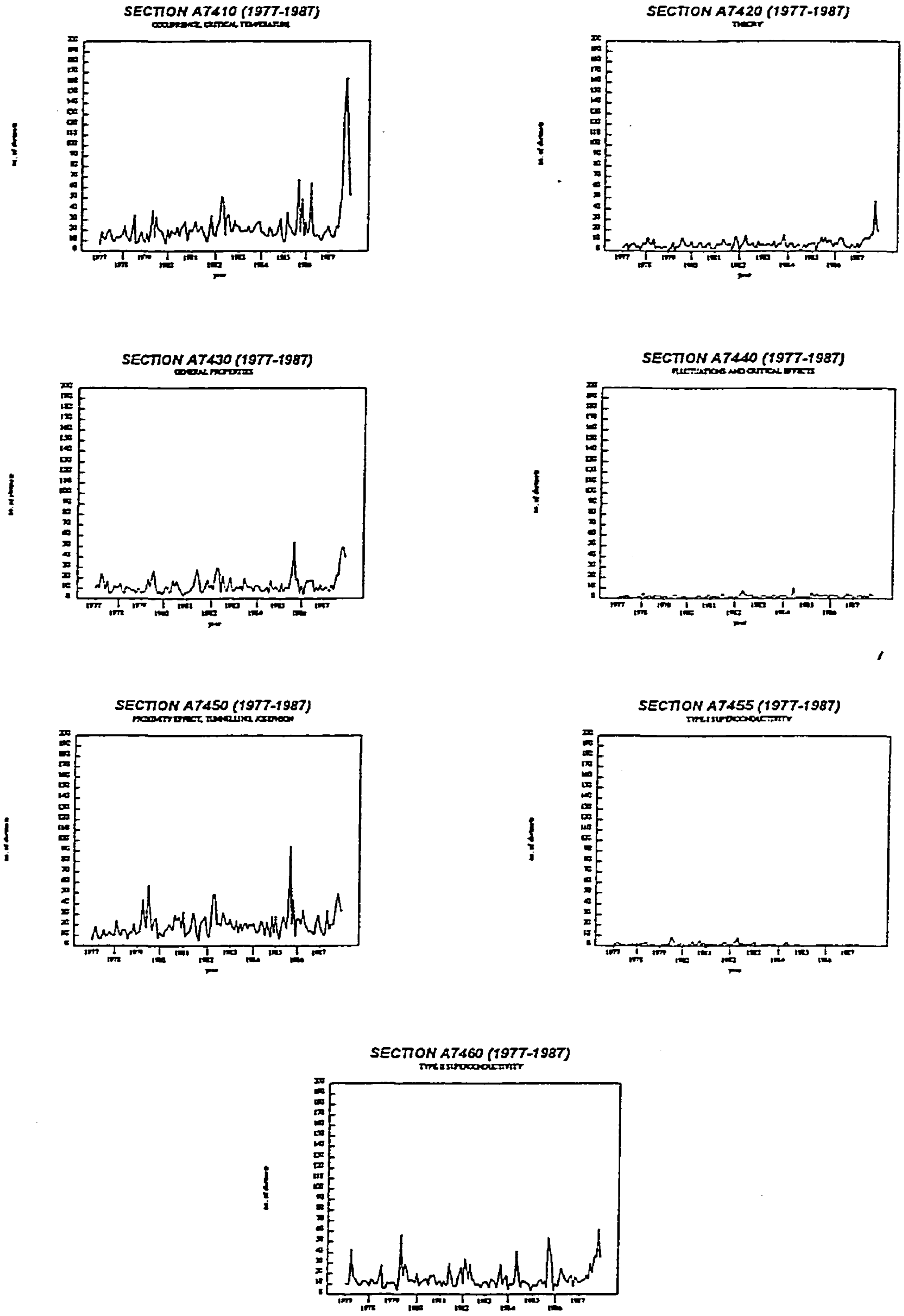


Figure 26. Comparison of the growth patterns for the sections comprising Chapter A74.

Given that there is already a seminal and influential article after which the dramatic growth of 1987 took place, outlier 2 supports hypothesis 2 by definition. The question then becomes: a) is there an influential work in outlier 1 and b) is the analysis appropriate for supporting the hypothesis? In other words, given that hypothesis 2 is already supported by outlier 2, does the analysis attempted here bring out results that give credence to this support?

The first outlier (hereafter named "outlier 741") consists of papers from four groups of works totalling 270 abstracts:

a) Proceedings of the *International Conference on Low Temperature Physics*, 17th, Karlsruhe, 1984 (175 abstracts);

b) Proceedings of the *International Cryogenic Materials Conference*, 5th, Colorado Springs, CO, 1984 (18 abstracts);

c) Proceedings of the *International Cryogenic Engineering Conference*, 10th, Helsinki, 1984 (12 abstracts); and

d) Miscellaneous journal articles (65 abstracts).

Of the 284 abstracts covered by *Physics Abstracts* for September 1985, 270 were analyzed. The remaining 14 (5%) are not covered by the *Science Citation Index*, are not available locally and are in foreign languages (Russian, Chinese or Japanese).

Since the papers from the three conference proceedings listed above are not covered by journals, they are not indexed by the *Science Citation Index*. To conduct the analysis, the proceedings were obtained by interlibrary loan, and the relevant references were typed into the *Physics Abstracts* records. The records were modified to fit the SCI format (such as capitalization of corporate sources and author names as well as the citation field) and analyzed with the FILTER.EXE and REPORT.EXE programs.

Table 14 below gives a ranked list of authors cited in outlier 741. There are a total of 712 authors in the outlier.

Table 14
Top Cited Authors for Outlier 741 by Conference in Rank Order[204]

ANDERSON PW	17
STEWART GR	17
BRANDT EH	13
MAEKAWA S	13
OTT HR	13
CLEM JR	12
FISCHER O	12
KRAMER EJ	12
WERTHAMER NR	12
MAPLE MB	11
VARMA CM	11
WOLF EL	11
DYNES RC	10
GRAY KE	10
LARKIN AI	10
STEGLICH F	10

[204]Only authors with more than ten citations are shown.

Table 15 below gives the corporate sources of the references in outlier 741. The list is dominated by American and Japanese institutions and the Soviet and Ukrainian Academies of Sciences. The table shows all institutions that contributed to five papers or more in the sample (eighteen in all). There are a total 175 institutions in the total sample.

Table 15
Top Contributing Corporate Sources for Outlier 741 by Conference in Rank Order[205]

UNIV CALIF LOS ALAMOS NATL LAB	12
ACAD SCI USSR	11
KYUSHU UNIV, FAC ENGN	11
ACAD SCI UKSSR	10
UNIV CALIF SAN DIEGO	10
ARGONNE NATL LAB	8
CNRS, CTR RECH TRES BASSES TEMP	8
IOWA STATE UNIV SCI & TECHNOL	8
MIT, CAMBRIDGE, MA, USA	8
TOHOKU UNIV, SENDAI, JAPAN	8
AT&T BELL LABS	7
UNIV SALERNO	7
IBM CORP, THOMAS J WATSON RES CTR	6
POLISH ACAD SCI, INST PHYS	6
KERNFORSCHUNGSZENTR KARLSRUHE	5
PHYS-TECHN BUNDESANSTALT, BERLIN	5
UNIV GENEVE, DEPT PHYS MAT COND	5
UNIV KARLSRUHE	5

The second outlier is the one following the Bednorz and Muller paper. Due to the burst of activity in superconductivity the outlier spans a period of three months:

[205] Only institutions with more than four contributions shown.

September to November 1987 with a total of 903 abstracts among them. There is no single conference proceeding that dominates the outlier, although there are several papers from conferences included.

As Table 16 below shows, the activity in this outlier is more intense, and the number of citations received by Bednorz and Muller clearly indicates that they stand well above the level of other papers in superconductivity. They are accompanied by three other authors who are also highly cited: Cava, Chu and Wu. The three published the first works confirming Bednorz and Muller's results as well as introducing new materials that exhibited superconductivity at yet higher temperatures. There are 2719 authors in outlier 742.

Table 16
Top Cited Authors for Outlier 742 by Month in Rank Order[206]

SEPTEMBER, 1987		OCTOBER, 1987		NOVEMBER, 1987	
BEDNORZ JG	103	BEDNORZ JG	126	BEDNORZ JG	70
CAVA RJ	85	CAVA RJ	117	CAVA RJ	68
CHU CW	72	WU MK	92	WU MK	65
UCHIDA S	42	CHU CW	69	CHU CW	48
WU MK	40	ANDERSON PW	50	ANDERSON PW	47
MATTHEIS LF	37	UCHIDA S	47	HOR PH	34
TAKAGI H	35	MATTHEISS LF	35	UCHIDA S	27
ANDERSON PW	31	HOR PH	31	MATTHEISS LF	23
KISHIO K	17	TAKAGI H	27	HIRSCH JE	20
NGUYEN N	16	TARASCON JM	25	MICHEL C	19
TARASCON JM	16	BARDEEN J	21	SUENAGA M	18
		BATLOGG B	21	EKIN JW	15
		HIRSCH JE	21	VARMA CM	15
		JORGENSEN JD	19		
		STEWART GR	19		
		WEBER W	19		
		GANGULY P	17		
		MIYAKE K	16		
		FISK Z	15		
		GRANT PM	15		
		LEE PA	15		
		MCMILLAN WL	15		

[206]Only authors with fifteen citations or more are shown.

The following table gives the list of the top five articles by the highly cited authors.

Table 17
List of five articles by the most cited authors in superconductivity

Authors	Journal	1987 citations
BEDNORZ JG & MULLER KA	ZEITSCHRIFT FUR PHYSIK B-CONDENSED MATTER, 1986, V64, N2, P189-193	890
CAVA RJ et al.	PHYSICAL REVIEW LETTERS, 1987, V58, N4, P408-410	400
CAVA RJ et al.	PHYSICAL REVIEW LETTERS, 1987, V58, N16, P1676-1679	389
CHU CW et al.	PHYSICAL REVIEW LETTERS, 1987, V58, N4, P405-407	400
WU MK et al.	PHYSICAL REVIEW LETTERS, 1987, V58, N9, P908-910	756

It is interesting to note that the work by M.K. Wu was also undertaken under the leadership of Paul (C.W.) Chu of the University of Houston. This list demonstrates that what was surprising in superconductivity was not that someone necessarily attained a higher critical temperature. The surprise was that the success occurred with an unsuspected family of materials. The fact that within a few months of the announcement by Bednorz and Muller several workers were able to produce materials that were superconductive at even higher temperatures and that those articles themselves are highly cited demonstrates that the field was ripe for a change and that Bednorz & Muller only provided the initial spark.

Table 18
Top Contributing Corporate Sources for Outlier 742 by Conference in Rank Order[207]

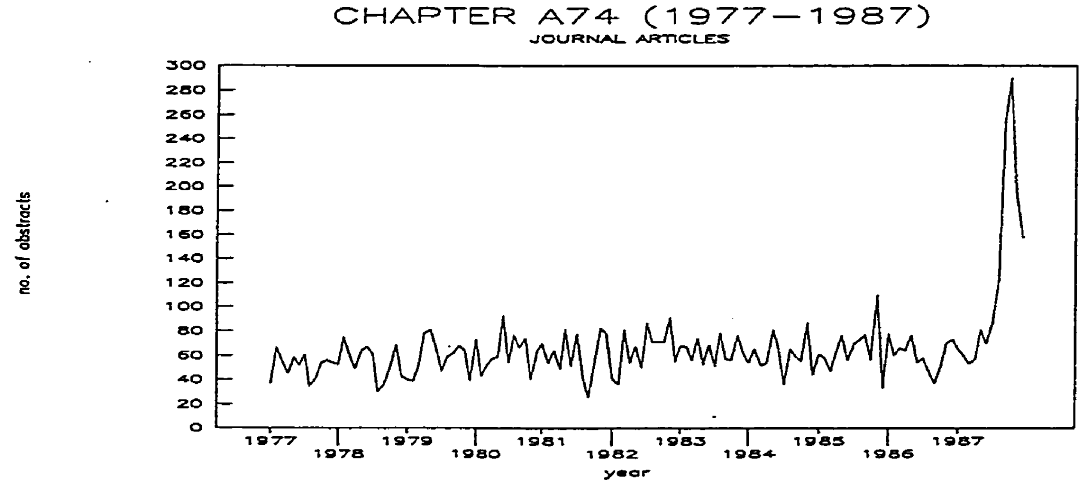
	SEPT	OCT	NOV	TOT
IBM CORP	12	18	9	39
UNIV CALIF BERKELEY	19	9	9	37
ACAD SCI UKSSR	4	5	22	31
TOHOKU UNIV	19	7	4	30
MIT	7	6	14	27
UNIV TOKYO	13	9	5	27
AT&T BELL LABS	10	7	8	25
ARGONNE NATL LAB	6	8	5	19
BELL COMMUN RES	6	7	5	18
CNRS	8	5	5	18
ACAD SCI USSR	7	2	8	17
CORNELL UNIV	6	4	6	16
INDIAN INST SCI	4	9	3	16
TATA INST FUNDAMENTAL RES	5	6	5	16
UNIV TSUKUBA	12	1	3	16
KERNFORSCHUNGSZENTRUM KARLSRUHE	3	3	9	15
UNIV CALIF SAN DIEGO	3	9	2	14
UNIV WISCONSIN	0	8	6	14
BHABHA ATOM RES CTR	4	7	2	13
CHINESE ACAD SCI, INST PHYS	6	3	4	13
UNIV CALIF LOS ALAMOS NATL LAB	2	8	3	13
NBS	2	3	7	12
OSAKA UNIV, FAC ENGN SCI	6	2	4	12
ROLAND EOTVOS UNIV	3	3	6	12
IOWA STATE UNIV SCI & TECHNOL	5	2	4	11
NIPPON TELEGRAPH & TEL PUBL CORP	5	5	1	11
ELECTROTECH LAB, JAPAN	5	1	4	10
KYOTO UNIV	4	5	1	10
STANFORD UNIV	4	3	3	10
UNIV SCI & TECHNOL CHINA	0	8	2	10

[207]Only institutions with ten or more contributions shown.

Table 18 above gives the corporate sources for outlier 742. The top ranks are still dominated by American and Japanese institutions. Compared to the previous outliers, the activity has increased considerably: there is a larger volume of works coming out from the top institutions. In addition, more institutions participate in work on superconductivity. The total sample contains 306 unique names of which 30 contributed to 10 or more articles.

All this lends support to the contention of this thesis that an information epidemic does indeed involve an influential work that gives spark to a field and attracts new workers to produce more. This increased attention is reflected in the dramatic growth of its literature.

Nevertheless, the dramatic growth needs to be qualified. The remarkable thing about Figure 27 below is that prior to the Bednorz and Muller article, superconductivity was a rather static field after all. When one separates the journal articles from conference papers, one sees that the apparent dynamism of the field was rather due to conference participation than journal publication. Given that journals are peer reviewed and sustain an archival function for the physics literature, the archive did not show any growth until the Bednorz and Muller work. There may have been a certain undercurrent that is reflected in an increase in conference papers during this period of time, but its interpretation will require more extensive analysis such as content analysis of the articles and a co-citation study of the field. From the titles of the papers in the sample it is clear enough that the major themes were "Josephson



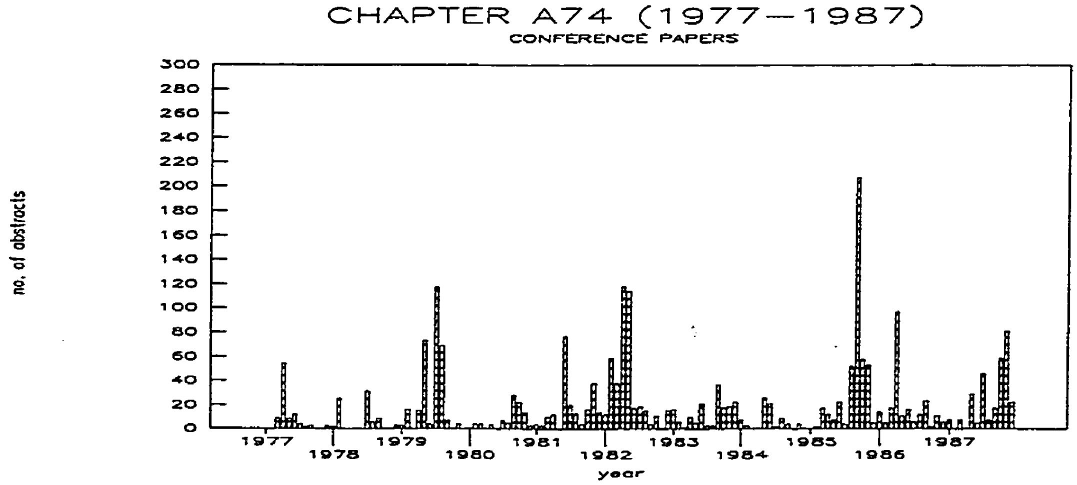


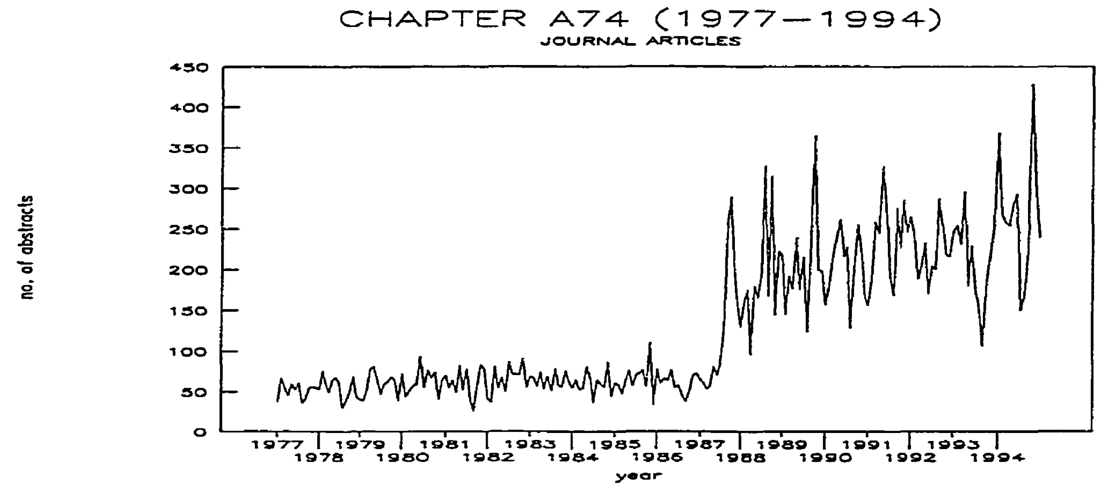
Figure 27. Comparing the production of journal articles with conference papers in the growth of Chapter 74.

junctions", "A15 superconductors" and "the Anderson model", P.W. Anderson that is consistently at the top of the citation rankings.

6.4.3. Summary

In hindsight, the major outlier of 1987 was not spurious and was sustained in subsequent years. An analysis of the literature indexed until the end of 1994 shows that superconductivity indeed exhibits a knowledge epidemic (not shown). It follows the Contagion Model described above in Section 2.1.2 (pages 25-27). The whole explosion took place following the publication of the Bednorz and Muller paper. Between 1987 and 1990 the publication rate remained high: the average number of journal articles is about three times the rate prior to 1987. Conference publication kept increasing, so much so that what was significant prior to 1987 in a regression analysis is no longer significant when compared to the outliers of 1989 and 1990 (Figure 28). Thus, high temperature superconductivity opened a new domain of research that is reflected in a new base of publications.

In this light, hypotheses 2, 3a and 3b of this thesis are supported.



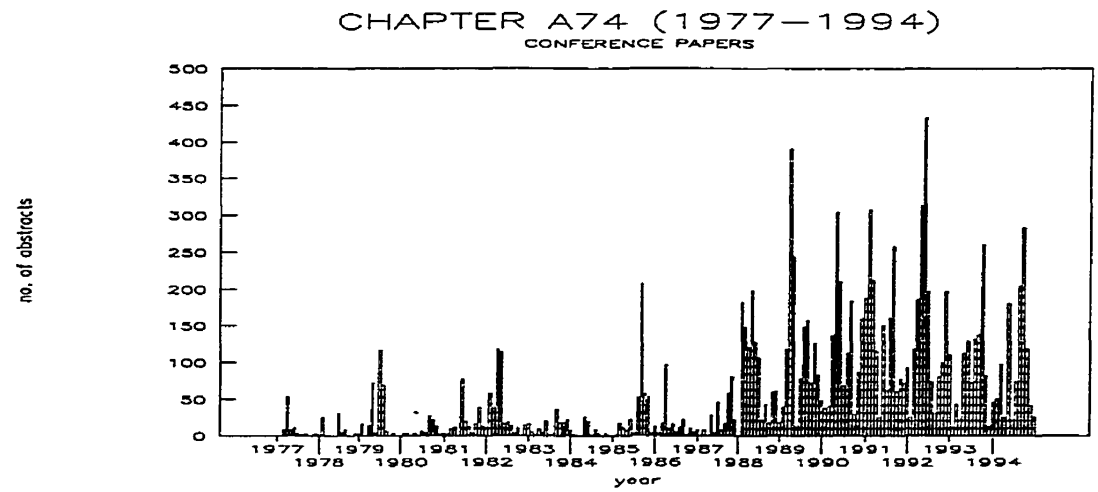


Figure 28. Comparing the production of journal articles with conference papers in the growth of Chapter A74 between 1977 and 1994.

CHAPTER 7

DISCUSSION

This study examined the prevalence of information epidemics in the physics literature between 1977 and 1987. The primary interest was to find out whether outliers observed on time series charts of literatures are due to information epidemics, whether these epidemics are widespread occurrences in physics, whether literatures showing such rapid growth arise mainly due to the influence of an important work and, if so, what characterizes these literatures. The work started with two research questions on page 2:

1. Is it possible that the yearly models [of growth] force an "extreme smoothing" on the observations and that the occurrence of spikes is more prevalent than hitherto suspected; this especially in view of the discontinuities and bursts of creativity that are prevalent in science?

2. Is it possible that some of the spikes are caused by influential papers that attract a large number of contributors to publish intensively in a given specialty area for a given time and cause a dramatic increase in the growth pattern of that area?

There was an additional question posed on page 25: Are the observed spikes meaningful in terms of the growth of knowledge and the evolution of ideas?

Based on these questions, a number of hypotheses were formulated and a methodology was developed to follow growth patterns and spikes, isolate outliers and analyze their contents for highly influential works in order to find support for the hypotheses. This chapter will discuss the findings in detail.

The survey of the *Physics Abstracts* chapters between 1977 and 1987 offers partial support for the hypotheses. Given the above-mentioned objectives, a number of major findings emerged. First, information epidemics exist, but they are not widespread. Second, some information epidemics are caused as well as carried forward by groups of influential works. Third, increased activity in a given field is accompanied by an increase in conference papers. Fourth, the journal literature is sufficient to obtain an accurate picture of the direction of literature growth in a given field.

In consideration of the findings, this discussion will first re-examine the hypotheses in light of the results obtained and then address four areas of interest: the implications for growth patterns, the prevalence of epidemics, the changing patterns of communication and a model for the dynamics of fast growing literatures.

7.1. Re-examination of the Hypotheses

Subsequent to the research questions posed at the outset, the following hypotheses were formulated in Section 2.4 on pages 63-64 above. It is now time to answer to each based on the results obtained from this work.

Hypothesis 1: **Sudden growth patterns termed information epidemics are widespread in the growth of the physics literature.** This hypothesis is not supported. The survey between 1977 and 1987 yielded eighty-two outliers. Of these, only two are outright information epidemics, those of Chapter 2 and Chapter 74. In fact, Chapter 74 turned out a knowledge epidemic. Two others (Chapters 36 and 73) are in their early exponential phase, marked by increasingly larger conference proceedings. While there is a large number of outliers in the data, the vast majority emanates from the indexing of conference proceedings.

Hypothesis 2: **Outliers observed during the growth of a field are caused by influential papers.** This hypothesis is partially supported. In light of the fact that the majority of outliers are due to conference proceedings, technically speaking, the hypothesis cannot be supported. However, in light of the information epidemics found, the hypothesis is supported. The outlier of Chapter 2 arose due to increased activity in mathematical physics in the middle 1980s. About half of the abstracts carry at least one four major descriptors that relate to work on nonlinear dynamics. However, there is no unifying theme among them, and there is no influential work that they all or a majority of them cite. Of the top seven citation getters, none forms a group with another and none has a dominant work in the citation profile. Their prominence is due to several decades of work rather than a single work or a small group of works that attract attention and influence the field.

On the other hand, the epidemics in Chapters 36, 73 and 74 are all due to a group of influential authors whose works dominate their respective fields and attract a large number of workers to bring in their contributions. This is reflected in the large number of citations obtained by the influential authors and their works. The works are of current origin, and a small number dominates the citation profile.

Hypothesis 3a: **Within the observed outliers, changes in the growth patterns are reflected in the number of works published, the duration of the growth and the number of research groups active in it.** This hypothesis is supported, especially by the evidence from Chapters 36, 73 and 74. In Chapter 36, as the exponential phase of the curve progresses, the conferences attract larger numbers of contributions, a larger number of authors as well as larger numbers of institutions participating (see Table 7, page 151). In Chapter 73, the same is true, as evidenced from Tables 8 and 9 (pages 154-155). In Chapter 74, the second outlier is much larger than the first. It also contains contributions from a larger number of authors and a larger number of institutions (see Tables 14-18, pages 181-186). The information epidemic of 1987 gave rise to a knowledge epidemic in that the sudden rise to a higher base level of abstracts has been sustained at least until the end of 1994.

Hypothesis 3b: **Within the observed outliers, the increased activity in a given specialty area is reflected in the increased proportion of conference papers abstracted** (when compared to journal articles). This hypothesis is also supported. Figures 17, 23 and 27-28 (pages 152, 168 and 188-190, respectively) all show that as a field grows, so does the number of conference papers. Chapters 36, 73 and 74 are all chapters that grew significantly from 1977 to 1987. Comparing the first three-year period for 1977-1979 with the last three-year period of 1985-1987, the number of conference papers in Chapter 36 went from two to 139 (70 times), in Chapter 73 they went from 67 to 594 (9 times), and in Chapter 74 they went from 483 to 945 (twice).

Overall, then, most of the hypotheses put forward in this thesis are supported. However, the most important one, that information epidemics are frequent in physics, is not supported.

7.2. Implications for the Study of Growth

First, as stated before, literature growth in physics does not necessarily follow either an exponential or a linear pattern. It varies and it fluctuates. The mechanism largely depends on the time period and the subject matter at hand. While the ultimate trend over a number of years may be monotonic (unidirectional), growth in the short-term is characterized by continuous ups and downs. All the figures in Appendix A bear this out.

Some, such as Chapters 29, 44, 67, 74, and 77, are essentially flat; their monthly output reflects little variance from the base and is only occasionally highlighted by a spike. Others, such as Chapters 7, 12, 13, 41, 79, 81, and 98 have larger variance, but remain within the ± 3 standard deviation envelope. A few, such as Chapters 82 and 87 show a zig-zag pattern not unlike Bottle and Rees' liquid crystal literature.[208] Chapter 36 is in its initial stages of exponential growth (in support of Price, May and Line, among others), but Chapter 67 represents linear

---

[208]Bottle and Rees, *Liquid Crystal Literature*, 117.

growth (even if interrupted by triennial outliers) in support of Oliver and King. Altogether, the figures in Appendix A represent a microcosm of the literature cited in Section 2.1.1. (pages 20-24) above. For all these reasons, the linear versus exponential dichotomy must now be increased to include a third class: hypergrowth, or in the parlance of this thesis, information epidemics.

In this respect, recent efforts at curve-fitting to explain the growth mechanisms of certain literatures may fall short of satisfactory solutions. Curve fitting, while useful, at times represents an oversimplification for a whole field, a forcing of complex phenomena and the resulting data into a single and simplistic pre-conceived mathematical equation. It would therefore be more efficacious for such efforts to move in the direction of simulation models with the use of simultaneous equations.

Another result from this study is the fact that *Physics Abstracts* increased its output by 59% between 1977 and 1987 (see Section 5.2, p.109). However, not all chapters showed concomitant growth. Most increased their output, a number stayed practically constant, and one (Chapter 35) decreased. On one hand, it is not unusual that the sum total of the abstracts in the chapters should have gone up. *Physics Abstracts* increased its own coverage of the journal literature during that time. The number of journals abstracted by the index went up from 2520 in 1977 to 3654 in 1987 (up 45%).[209] On the other hand, the increase in the chapters cannot be

---

[209]The list does not provide the actual numbers. The figures were arrived at counting three random pages of each list and multiplying their average by the total number of pages in each list.

ascribed to a change in policy because *Physics Abstracts* indexes and abstracts **almost all** physics journals. Thus, the increased coverage is a reflection of the increasing number of physics journals published, and that is directly due to the increasing number of publications submitted by researchers. As such, the increasing number of journals in a field in itself reflects the growth of the literature of that field. Recent evidence shows that, in fact, the start of a new journal can be used as an indicator for the start of a new specialty.[210]

The fact that *Physics Abstracts* went up by 59% between 1977 and 1987 is no longer surprising once it is compared to other major scientific databases. A quick survey of the growth for the same time period shows that several major databases such as BIOSIS, GEOREF, Chemical Abstracts, MEDLINE, Sociological Abstracts and Economic Literature Index grew at about the same rate. These are the largest databases in their fields and they all show similar factors for eleven years of growth.

The number of scientific journals in the world between 1977 and 1987 also went up. The total number of serials covered by *Ulrich's International Periodical Directory* between 1977 and 1987 went up by 81% from 60,000 to 108,590. The size of the index during the same period went from one volume to three, from 2096 pages to 4933.[211] The number of periodicals listed under *physics* increased from 457 in

---

[210]L. Leydesdorff et al., "Tracking Areas of Strategic Importance Using Scientometric Journal Mappings," *Research Policy* 23 (1994): 217-29.

[211]*Ulrich's International Periodicals Directory*, 17th ed. (New York: Bowker, 1977); *Ulrich's International Periodicals Directory*, 27th ed. (New York: Bowker, 1987).

1977 to 946 in 1987 (107%).[212] While it is not possible to obtain accurate figures for the number of scientific serials being published at any one time, there is no doubt that physics journals showed a healthy increase during the same period.

In sum, then, the overall growth of *Physics Abstracts* is due to the growth of the literature of physics per se and not to a change of policy on the part of editors to increase their coverage of already established journals. Therefore, the growth of the individual chapters of *Physics Abstracts* is also due to the growth of particular fields of physics. This very fact renders it safe to follow the output of *Physics Abstracts* without concern about its journal coverage.

It is clear, then, that the classical postulates on linear or exponential growth are no longer accurate enough for the literature of physics and, by extension, for scientific literatures. Science activity, as reflected in its publishing, is a dynamic process and alters its course as advances and interests change. An examination of the charts for the chapters shows that not all chapters grew during that time. In other words, stating that the physics literature grew during this period of time is accurate but represents an oversimplification. One suspects that the dynamics observed in the literature of physics is also valid for the literature of science overall.

Another interesting result is that it is enough to follow journal publications in order to trace the growth patterns of a field. In fact, separating journal articles from conference papers leads to a more genuine picture of the growth of a given field.

---

[212]This excludes all the see-references to other topics such as mathematics or astronomy as well as abstracting and indexing journals.

Conference papers and other publications increase the variance in the data but do not alter the direction of the growth. Figure 17 (page 152), Figure 23 (page 168), and particularly Figure 27 (page 188) dramatically support this position. When one separates the journal articles from conference papers, one sees that the apparent ups and downs of a given field are rather due to conference participation than to fluctuations in journal publications. Given that journals are peer reviewed and sustain an archival function for the physics literature, they seem to be the more reliable indicator of the true growth trend. On the other hand, although many conference papers are peer reviewed, the quality of the peer review is not as rigorous as it is for journal articles. For example, Figure 28 on page 190 shows the results of the superconductivity analysis extended to the end of 1994. It is interesting that the outlier of 1987 turned out to be a knowledge epidemic, something that was sustained and kept growing in subsequent years. Between 1987 and 1994 the publication rate remained high, the average number of journal articles being at about three times the rate prior to 1987. Thus, all the excitement, the changes, the successes and failures in superconductivity prior to 1987 gives nothing more than a flat curve, seemingly inconsequential and innocuous compared to what was to come after 1987.

Thus, the contribution that a knowledge epidemic makes to the growth of knowledge is that it pushes knowledge production to a new base level. The curve between the old and the new level may be a linear, sigmoid/logistic, or even a threshold/step function such as in Figure 28. In other words, it may be slow or it may be rapid. In the case of a knowledge epidemic it represents very rapid growth: a new dynamic, showing unusual patterns that reflects science in its accelerated mode. If the epidemic is successful, that is if the novelty brought about by the group of influential works is accepted and spreads rapidly, the new base level remains sustained and forms a new base level for the next epidemic or jump in growth. In the meantime, following the four-step model for the growth and development of scientific specialties outlined in Section 2.1.4, the epidemic gives rise to a new subspecialty, a new field or altogether a new science.

In the case of superconductivity, the knowledge epidemic has represented a new science, in the sense that studies of superconductivity have been transformed fundamentally. For example, the number of researchers increased from 100 in 1985 to 5000 in 1987, a third of them being in Japan and Europe, and most of them working in government laboratories.[213] A large number of university departments took up work in high temperature superconductivity. Federal funding in the United States jumped from a few millions prior to 1987 to US$130 million in 1990, just three years after Bednorz and Muller's discovery.[214] The United States government also agreed to increase university funding by at least US$15 million over and above existing

---

[213]C. DeBresson, "Predicting the Most Likely Diffusion Sequence of a New Technology through the Economy: The Case of Superconductivity," *Research Policy* 24 (1995): 685-705.

[214]U.S. Congress. Office of Technology Assessment. *High-temperature Superconductivity in Perspective* (Washington, DC: Government Printing Office, 1990), 5.

levels over several years starting in 1990. Needless to say, the scientific output increased dramatically.

All this lends credence to Kuhn's contention that revolutionary works open new perspectives on nature and fundamentally transform the activity of a given field. It also supports his arguments in favour of the process of speciation, that is that new scientific specialties emerge as a result of revolutionary change.[215] Kuhn calls these changes *crises*,

> crucial symptoms of the speciation-like process through which new disciplines emerge, each with its own lexicon, and each with its own area of knowledge. It is by these divisions, I've been suggesting, that knowledge grows.[216]

7.3. The Prevalence of Epidemics

A second major result of this work is that information epidemics are unusual occurrences. The survey of fifty-eight chapters over 132 months (eleven years) yielded only four cases, of which one (Chapter 74) is a knowledge epidemic, one (Chapter 2) is an information epidemic and two are at the early exponential phase of their growth (Chapters 36 and 73). As mentioned above, this result fails to bring much support to the first hypothesis.

On page 32 above it was stated that sudden discoveries or surprising publications may offer surprising solutions to old problems, and in the rush to rework

---

[215]T.S. Kuhn, "The Road Since Structure," *PSA 1990* 2 (1990): 3-13.

[216]Ibid., 9.

and publicize old facts in a new light it should not be surprising to find information epidemics. This is still a hope despite the somewhat meagre findings of this work.

One case where the expectation for an information epidemic did not materialize involves the discovery of the W and Z particles in the early 1980s. Their existence had been predicted in 1973 by Hasert, but they were not discovered until 1983.[217] It was thus expected that this discovery would create a large amount of excitement and lead to an information epidemic. The expectations did not materialize. Figure A11 in Appendix A shows that Chapter 13 did not show any acceleration in growth in 1983 when the technical articles announcing the discovery of the W and Z particles were published. Section A1385 (Hadron-induced high- and super-high-energy interactions, energy > 10 GeV) or Section A1385K (Inclusive reactions, including total cross sections, (energy > 10 GeV)) which carried those abstracts did not grow either. A look at previous years' *Physics Abstracts* showed that the only time those section showed any major growth was during the 1973-1976 period that followed the publication of Hasert's works. This means that what induces an information epidemic is not the confirmation of a prediction but the prediction itself and its acceptance by others in the field. The publication of Hasert's works gave rise to an information epidemic that lasted four years. During that time he was cited several hundred times. However, the publication of the papers announcing the discovery at the CERN Collider in Europe did not give rise to any epidemic, even though it too has been

---

[217]Robert P. Crease and Charles C. Mann, *The Second Creation: Makers of the Revolution in 20th Century Physics* (New York: Macmillan, 1986), 348.

cited several hundred times. Thus, what gives rise to an information epidemic in physics primarily is theoretical interest. Once the confirmation comes, the field is ready to be finalized, and interest wanes.

The expectation of surprise is not limited to this author alone. From time to time the popular literature makes announcements of sensational discoveries that turn out to be anything but discoveries. This has probably been best described by Sheldon Glashow in his autobiography:

> "The discovery of the j/psi particle was a big surprise, and even the existence of charmed particles surprised some of us. As if that were not enough, there were the tau lepton, the upsilon particle and its associated beauty particles. ... The 1980s has been a remarkably quiet decade. ... the existence of the W and Z bosons (like the existence of the antiproton) cannot really be thought of as something unexpected. ... The lack of a fundamental but unanticipated discovery in this decade has not been for wont of trying. Quite a number of surprises were *reported* in the 1980s. The trouble is that none of them seem really to be there. Perhaps a list of such nondiscoveries will suffice: magnetic monopoles, neutrino oscillations, neutrino masses (twice: Russian and Canadian), zeta particle, no-neutrino double-beta decay, muons from Cygnus X-3, proton decay ..."[218]

None of the above mentioned "discoveries" were detected in *Physics Abstracts* as having produced outliers, and none of the four cases found in this study are present in Glashow's list.

In a way, this reflects the fact that the theoretical understanding of high energy physics is largely established and that the foundational questions have been largely

---

[218]S. Glashow, *The Charm of Physics* (New York: American Institute of Physics, 1991), 189-90.

settled and that a finalization is now taking place. Big science in particle physics, to use Derek Price's term, is largely becoming outdated due to the fact that the foundations of modern physics no longer need be tested and re-tested. They are well established. Thus, *the scientific enterprise is now largely involved in the creation of novelty - in the design of objects that never existed before and in the creation of conceptual frameworks to understand the complexity and novelty that can emerge from the known foundations and ontologies.*[219] This is precisely what the literature reflects. Where such novelties become significant, information epidemics take place: in cluster physics, in new semiconductor devices and high temperature superconductors.

The most successful of these novelties is the case presented by Chapter 74, "Superconductivity". The initial Bednorz and Muller paper[220] surprised the physics community and provoked a rush to confirm the work and to look for other compounds at yet higher critical temperatures (T_c). As these were found, the number of workers in the field grew and so did their output. Figure 28 on page 190 clearly shows that the output of superconductivity physics prior to 1987 looks rather insignificant when compared to the activity after 1987.

The case of superconductivity points to another fact: support for the original Goffman hypothesis. On pages 32 and 33 above a differentiation was made between

---

[219]Silvan S. Schweber, "Physics, Community and the Crisis in Physical Theory," *Physics Today* 46 (1993): 34-40.

[220]Bednorz and Muller, *Possible High Tc Superconductivity*.

the *Contagion Model* and the *Catalyst Model* of epidemic processes. Superconductivity is the best case yet to support the Contagion Model. The 1986 paper of Bednorz and Muller acted as an infectious agent that was rapidly transmitted in the population of solid state physicists and physics in general. The rapidity of the transmission is clearly reflected in the large numbers of citations obtained by the article within a short time period. Given the importance of superconductivity in science, technology and medicine, there is no doubt as to the large number of *susceptibles* in society. The latency period (publication lag time) was also shortened by the favourable attitude of many editors of physics journals who sped up the review process for newly submitted articles on high temperature superconductivity. Thus, given the initial tension that existed in the superconductivity community in 1986, the publication of a surprising and significant paper acted as a spark that produced an explosion in high temperature superconductivity work. Within two years, the vast majority of works published in superconductivity were dealing with the new family of high temperature materials discovered by Bednorz and Muller. This is also attested to by the fact that the two authors shared the Nobel prize for physics the year following the publication of their paper.

It should be stated that the Catalyst Model of epidemics is also supported by the results of this study. The Catalyst Model proposes that the confluence or the cumulation of a number of factors within a short time period can give rise to an epidemic. The results of Section A36.40 (Atomic and Molecular Clusters) and Section

A73.40L (Semiconductor contacts, heterojunctions) point toward the presence of a Catalyst Model in the growth dynamics of the physics literature. As such, this represents a novel contribution from this thesis. The two will be discussed in some detail here.

The development of cluster physics goes back to the 1960s. However, it is in the early 1970s, with advances in instrumentation and the possibility of producing large quantities of atomic clusters that the field started to take off. New techniques such as metal vapour synthesis, extended X-ray absorption fine structure spectroscopy (EXAFS), and the development of laser ablation techniques for the synthesis of increasingly larger clusters moved the field further.[221] The first conference on clusters took place in 1977 but was not indexed by *Physics Abstracts*. It brought together researchers (mainly of European origin) from diverse areas looking for a unified approach to the study of cluster physics. In addition, the usefulness of potential findings in cluster physics to diverse fields as lubrication, catalysis and photography brought a diverse group of workers together that produced an interdisciplinarity in a confined field of study, thus attracting more workers and inducing them to produce more. Theoretical difficulties added to the attraction of the field. To date the major question remains: *How many metal atoms must a cluster contain before its properties are indistinguishable from the bulk?*[222] The recent

---

[221]M. Moskovits, *Metal Clusters* (New York: Wiley, 1986), 3.

[222]Ibid., 2.

discovery of a third fundamental form of carbon, buckminsterfullerenes (also called *buckyballs*), by Smalley and Kroto, has its origins partly in their work on clusters and the development by Smalley of a special instrument for its synthesis.[223] Although their purpose was to better understand the presence of certain molecules in stars and synthesize them here on earth, their understanding of the physics and chemistry of these molecules owes its origins to their background in cluster research.

The confluence of all these factors in the late 1970s and early 1980s produced the exponential rise in the literature of cluster physics that is reflected in Section A36.40 of *Physics Abstracts*. Even though the field is rising exponentially and has not described a complete epidemic curve as yet, the fact that it is interspersed with a number of outliers all due to the proceedings of the same conferences over a number of years and the fact that these conferences keep growing from year to year makes this a bona fide information epidemic.

The factors leading to the growth of the literature of Section A73.40L are somewhat different. The first instance goes back to the surprising discovery by Klaus Von Klitzing in 1980 of the quantized Hall effect and his discovery of a new standard of electrical resistance at the atomic level. For this he was awarded the Nobel prize in physics for 1985. Given the theoretical importance of the discovery as well as the practical consequences for semiconductor physics and given the immense place of semiconductor technology in industry worldwide, the field immediately attracted a

---

[223]Jim Baggot, *Perfect Symmetry: The Accidental Discovery of Buckminsterfullerene* (Oxford: Oxford University Press, 1994).

large number of new workers. In 1982 Tsui, Stormer and Gossard at Bell Labs in New Jersey discovered the **fractional** quantum Hall effect.[224] However, while Von Klitzing's discovery could be explained adequately, the theoretical explanation for the fractional effect presented more serious difficulties. An additional factor that helped Von Klitzing in his discovery was the availability of new instrumentation to prepare the new semiconducting materials needed to demonstrate the physical phenomenon. The growth of Section A73.40L thus represents a complex pattern where the necessity to advance theory in physics is intertwined with the necessity to produce new semiconductor materials with which the experiments can be conducted and the phenomena can be understood. This also explains the reason why so many of the abstracts in Section A73.40L also carry abstract numbers for Section B2530B (semiconductor junctions) of *Electrical and Electronics Abstracts*. The same papers that have consequences for physical theory also have consequences for the synthesis of new semiconducting materials. What is interesting here is that a single surprising paper did attract new workers and did give rise to a new specialty, but before the particular specialty represented by the paper (Von Klitzing's work) could take off, theoretical difficulties had been overcome and the phenomenon had been explained. Instead, workers turned their attention to a related problem, the fractional quantum Hall effect. Still newer semiconductor materials were needed to perform experiments, and theoretical difficulties only grew larger. Given the importance of the

---

[224] "Fractional Quantum Hall Effect Indicates Novel Quantum Liquid," *Physics Today* 36 (July 1983): 19-22.

phenomenon, as mentioned above, the field and its literature continued to grow. Thus, the group of top workers obtained from the analysis of Section A73.40L is the group that followed Von Klitzing and applied itself to the solution of the fractional problem. This group is largely led by names such as Stormer, Tsui, Ando and Bastard that appear in the top ranks of Table 12 (page 170). These authors consistently obtain large numbers of citations. They are very active, they present papers at all the major international conferences, and they are authors of several papers cited more than one hundred times, as reflected by the *Science Citation Index*. Similar to Section A36.40 this section is also in its early stages of exponential growth, occasionally marked by outliers due to major and increasingly larger conferences.

Different as the fields are, the dynamics of Sections A36.40 and A73.40L show several commonalities: they both represent fields of great practical significance and commercial opportunities, they both have theoretical difficulties that are as yet unsolved, both developed as a result of advances in instrumentation and the availability of increasingly cheaper materials, and both flourished as a result of international participation by researchers. These may well represent the factors that are needed by the Catalyst Model to produce an information epidemic. These will be further discussed in Section 7.6 below (page 233) on fast growing literatures.

Another point that needs to be made concerns the fact that most outliers found on the time series charts are due to the indexing of conference proceedings. This means that one of the original precepts of this study needs to be reversed: when one finds an outlier on a growth curve, the likelihood is higher that it will be due to the indexing of conferences rather than the presence of an information epidemic. Of the eighty-two outliers listed in Table 2, only two represent outright information epidemics: those of Chapter 2 (Mathematical Methods in Physics) and Chapter 74 (Superconductivity). Thus, the chance of finding an information epidemic among the outliers present in the physics literature during the 1977-1987 period is no more than some two and a quarter percent. It should be mentioned, in passing, that there is one instance in the literature where the presence of spikes in a graph is ascribed to conferences. The caption under Figure 3.1 of *Drifting Continents & Colliding Paradigms* by John Stewart reads: *Peaks before 1968 are due to publications from symposiums.*[225]

When it comes to the presence of influential works in information epidemics, it seems clearer now that these works do not guide or influence a field in isolation. The original premise of this work was that information epidemics arise due to the extraordinary influence of a given work. Now, that has to be reworded to say that it arises due to the influence of a **group of works**. The epidemics found so far are dominated by groups of papers and groups of authors. This also supports Goffman's contention that what guides epidemics is a group of works.[226] There was no single work or a homogeneous group of works leading the epidemic in Chapter 2. There was

---

[225]John A. Stewart, *Drifting Continents & Colliding Paradigms: Perspectives on the Geoscience Revolution* (Bloomington: Indiana University Press, 1990), 46.

[226]Goffman and Newill, *Generalization of Epidemic Theory*, 225.

no single work providing the spark for the cluster work in Chapter 36, although the field spent a decade and a half in gestation before starting its exponential rise. The spark in Chapter 73, with the work on heterostructures, was provided by Von Klitzing, but the epidemic is due to work on two-dimensional electron clouds, and that, as mentioned above, is led by an international group of authors. In view of the fact that this group is the one most actively working on the solution to a difficult theoretical problem, it is the one to dominate the citation rankings. This is also consistent with the finding that citation scores are affected by the size of the activity in a given subfield; those who share an intellectual focus with other groups tend to obtain a larger number of citations than those *that work more on their own.*[227] Finally, the stimulation in Chapter 74 was provided by Bednorz and Muller, and they are largely responsible for the origin of the epidemic. The list of citations is also led by a group, the one mentioned in Table 17 (page 185 above): Cava, Chu and Wu.

Another significant factor with information epidemics is that the group of influential works obtains an unusually high number of citations. As mentioned in Section 4.7 above (page 96), less than half a percent of all papers in the *Science Citation Index* ever receive more that one hundred citations.[228] As it is, to obtain a large number of citations it is necessary to be exceptionally productive and to publish

---

[227]H.F. Moed, "Bibliometric Measurement of Research Performance and Price's Theory of Differences Among Sciences," *Scientometrics* 15 (1989): 473-83.

[228]Garfield, *The Most Cited Papers of All Time*, 52.

in international journals scanned by the *Science Citation Index*.[229] As Table 17 (page 185) showed, the five papers leading the superconductivity epidemic obtained several hundred citations in their first year alone. The *Science Citation Index*, on December 31st 1995, lists 5481 citations for the 1986 Bednorz and Muller paper. In fact, all the leaders of the epidemics are highly influential in their specialties. The top five authors of Table 8 (page 154) all have works cited more than one hundred times in the first five years of their publication. The same is true for the top five authors of Table 12 (page 170).

The analysis of the corporate sources in the outliers suggests that outbursts of activity and a high velocity of communication are not attained unless major universities and/or corporate research centers become interested and contribute at the same time. In Chapter 36 cluster physics does not begin to accelerate until the Berlin and Heidelberg conferences of 1984 and 1986, respectively, at which contributors from the IBM Corporation and the Max Planck Institute present several papers (see Table 9, page 155). Whereas the largest numbers of contributors at the first two conferences (Lausanne and Konigstein) came from École polytechnique fédérale de Lausanne, Université de Paris 11, and the Kernforschungszentrum Karlsruhe, the largest numbers of papers at the latter two (Berlin and Heidelberg) came, in addition to the above, from the Max Planck Institute in Stuttgart, the University of Constance,

---

[229]T. Luukkonen et al., "The Measurement of International Scientific Collaboration," *Scientometrics* 28 (1993): 15-36.

the KFA Julich GmbH corporation and the University of Saarland in Saarbrucken. Cluster physics in the 1970s and early 1980s was definitely a European centered activity. Despite the prominence of American universities in world research, only two universities are present in the list of top 20 institutions: Stanford University and Penn State University.

In contrast to Chapter 36, the author affiliations in the outliers of Chapter 73 are more diverse (see Table 13, page 171). Five of the top 10 are of American origin. Five nations from three continents are present among the sixteen names represented with having made four or more contributions. The names represent the largest American and French corporations and some of the largest universities in each country.

This is even more evident from Chapter 74 and Tables 15 and 18 on pages 182 and 186, respectively. The two tables represent the before and after the publication of the 1986 Bednorz and Muller paper. Both tables are dominated by American and Japanese institutions. However, Table 18 represents far a greater volume, a higher diversity, and a larger number of contributions from other countries such as India and China. The total sample of author affiliations carries 175 names for the first outlier and 306 for the second one.

All this is to say that an information epidemic will not take place unless the spurt in growth is accompanied by a growth in the number of participating institutions as well as the arrival of the major institutions that have a stake in the results of the research being conducted. One of the earliest participants in cluster work was the Eastman Kodak Corporation, arguably the largest maker of photographic film in the world, with a stake in materials affecting photography. Leading the pack of contributors to work on heterostructures are laboratories from AT&T, IBM and Thomson CSF, again the largest corporations of the kind likely to benefit from advances in semiconductor work.

All epidemics found here share several major characteristics. One relates to instrumentation and new materials. Table 19 below illustrates the commonalities:

Table 19
Factors common to the information epidemics identified

Chapter no.	Instrument	Material
2	computers	simulation software
36	laser ablation	clusters
73	chemical vapour deposition	heterostructures
74	(not applicable)	oxides

In addition, all have theoretical consequences that are significant. Chapter 2 relates to work on nonlinear dynamics and the hope for a new description of nature. Chapter 36 reflects significant consequences for photographic materials, catalysts and understanding a new form of matter, buckminsterfullerenes. Chapter 73 provides a new standard for resistance and a new explanation for the behaviour of electrons between layers of semiconducting materials. Finally, in Chapter 74 the reasons for the phenomena are not clear. Experiments for now are being conducted by trial and error while a comprehensive theory is being sought. Given the social significance of superconductivity, finding the underlying theory for the new family of materials is of primary importance.

Another major commonality among the epidemics is that they all emanate from (to use Derek Price's term) *little science*. All come from fields where the activity is centered among small groups of individuals working in a variety of institutions. The instruments they use, while expensive, are not beyond the reach of any major university department or corporate research center. The materials, once the methodology and/or the instrumentation to synthesize or acquire them cheaply has been worked out, are not difficult to make or obtain. The experiments can be run by the researchers themselves, and the results can be analyzed in place. International collaboration comes from collegiality and the sharing of common research interests, not out of necessity to work on the only instrument(s) or facilities available. This is in sharp contrast to *big science* such as particle physics or astronomy which requires instrumentation costing many millions of dollars, run by a multitude of engineers and scientists, where time for experiments must be booked months in advance and collaboration is necessitated by the impossibility of working alone or in a small group. For example, the papers confirming the discovery of the W and Z particles in

1983 carried between 110 and 150 authors.[230] The recent discovery of the tau lepton carries 440 authors.[231] Most of these *authors* are scientists and engineers who made it possible to run the experiments but are not the theoretical physicists who conceived, designed and analyzed the experiments.

One further commonality is that all the four are reflections of *enabling science*. They not only signal success in their own right, but they also provide theoretical arguments to help move other fields forward as well as to enable engineering applications toward the manufacture of industrial and consumer products. Thus, the data and information obtained from the work in these fields are useful to many other related and less related fields.

Finally, there is here the need to distinguish between works that are "puzzle generating" and those that are "puzzle solving". A puzzle generating work is one that other researchers can build on or can use in their work. A puzzle solving work is one that produces results that are expected or that fit into prevailing theory.[232] For example, the information epidemics found here are all puzzle generating. Information epidemics are obtained if the research activity in a field results in a significant puzzle

---

[230]For example, G. Arnison et al. "Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c/sup 2/ at the CERN SPS Collider," *Physics Letters B* 126B (1983): 398-410.

[231]CDF Collaboration, "Evidence for Top Quark Production in pp Collisions at Square Root (s)=1.8 TeV," *Nuclear Physics B, Proceedings Supplements* 39 (1995, no.B-C): 343-7.

[232]Cole, *Making Science*, 47.

generating work where looking for solutions keep researchers working and publishing intensively. Wherever puzzles are solved, information epidemics do not take place unless a new puzzle is found and propositions for its solution are put forward. In other words, those works that resolve theoretical tensions or confirm theoretical predictions provide the beginning of the end for an epidemic. BCS theory in superconductivity led to a drying up of funds, the confirmation of the W and Z particles did not lead to an information epidemic, and recently, the results obtained from the Cosmic Background Explorer satellite (COBE) have *meant the sudden death for a number of cosmological theories.*[233]

Given the results of this study, it may now be possible to classify information epidemics as the reflection of three types of *anomalous science* (in the sense of Kuhn):

1. internally generated, especially through theoretical advances, such as nonlinear dynamics, heterostructures, and superconductivity,
2. externally precipitated, such as clusters and superconductivity (in relation to the profit seeking motive),
3. conferences, as an artefactual group behaviour of scientists.

Further work will be needed to either support this classification or to modify it.

---

[233]John Horgan, "COBE's Cosmic Cartographer," *Scientific American* 267, July (1992): 34-41.

7.4. Changing Patterns of Communication

The presence of information epidemics in literatures indicates that the patterns of communication may be changing, especially in physics. The dominant model over the last twenty-five years has been that of Garvey and his coworkers.[234] His model of scientific communication is largely a linear process where work initiated, say in a laboratory, goes through the stages of preliminary report, technical report, preprint, journal publication, and abstracting journal until it enters the core (if at all) in the monographic literature and finally the encyclopedia. This may now be changing. Partly influenced by sociologists of science, especially constructivists, science is increasingly being seen as being socially shaped and changing form depending on the audiences that it addresses.[235] It is being seen as gaining in complexity, losing the linear process in favour of a more complicated process where all forms of communication interact and lead to each other, and where an information glut ends up adding to instability and confusion about facts.[236] Increasingly, the journal literature is losing its function as the formal harbinger of new developments in a specialty and is taking on the archival function. The electronic journal is changing many of the traditional functions of the journal as a tool for technical communication. Work that is significant and needs to have a priority claim made over other colleagues is still

---

[234]Garvey, *Communication.*

[235]Bruce V. Lewenstein, "From Fax to Facts: Communication in the Cold Fusion Saga," *Social Studies of Science* 25 (1995): 403-36.

[236]Ibid., 425-9.

submitted as a printed journal article, and the peer review system ensures that material of high quality gets published. However, the news function is now partially taken over by the conference literature but also increasingly by electronic mail and the facsimile machine (fax). For example, the paper announcing cold fusion was disseminated by fax long before the technical article was published.[237] However, not all of this is as yet reflected in the 1977-1987 data that is the subject of this study.

What emerges from the patterns of communication in the 1977-1987 data is that the literature output is dominated by journal articles and conference papers. As mentioned above, separating journal articles from conference papers has provided a more genuine picture of the growth of a given field. The case of superconductivity as well as cluster and semiconductor physics support this argument. Especially in physics, the journal article has now become an archival medium, mainly useful for documenting past achievements and establishing priority. The communication function as such has now been overtaken by electronic mail and in part by the conference literature.[238]

So far as conference proceedings go, they have now become a byproduct of the group communication processes. On the one hand, in most science and engineering fields conference proceedings constitute a major source of information.

---

[237]Bruce V. Lewenstein, "Cold Fusion and Hot History," *Osiris* 7 (1992): 135-63.

[238]Ginsparg, *First Steps*, 391.

For example, there are over 500 conferences held in physics every year.[239] A survey by Oseman found that in the mid-1980s the British Library Document Supply Centre was receiving more than twenty thousand conference proceedings a year.[240] Of those, about two thirds were published as proceedings, one third were published in journals.

Conference proceedings as a publication format present several problems. For the contributor, a conference paper takes time to write (even if it is produced at a much faster pace than a journal article), it is not adequately refereed, it is rarely cited, and therefore does not carry the same weight on one's resume as a journal publication. Due to delays in the publication of most proceedings, conference papers function mainly as a written record of a conference. Reliable documentation is still available in journal articles.[241] For the collection developer in a research library, conference proceedings are very expensive, they are not always indexed, and not every conference results in the publication of proceedings. When it does, the delay often amounts to well over a year. When the proceedings are included in a journal subscription, they raise the subscription price considerably. Most publishers do not

---

[239]Barschall and Haeberli, *Conference Proceedings in Physics*, 564.

[240]Robert H. Oseman, "The Growth and Value of Conference Literature," in *New Horizons for the Information Profession* ed. H. Dyer and G. Tseng. (London, Taylor and Graham, 1988), 124-37.

[241]Robert S. Allen, "Physics Information and Scientific Communication: Information Sources and Communication Patterns," *Science and Technology Libraries* 11 (1991): 27-38.

offer the option of not subscribing to the proceedings section. For the user, conference papers are difficult to find, they do not carry enough details, and they do not necessarily report original work. In many respects, conferences simply provide opportunities for face to face communication and discussion of authors' works, but so far as the advancement of a given specialty goes, they do not represent the format of choice. This function is now increasingly being fulfilled by the electronic medium, especially the Internet.

An informal survey by David Pendlebury of the Institute for Scientific Information in 1993 showed that of the 1.3 million conference papers in the *Science Citation Index* database, only 51 were cited more than fifty times and 761 were cited more than twenty times. In contrast, 1.06% or 348,537 journal articles in the 1945-1988 cumulation of the database were cited more than fifty times and 842,950 (2.56%) were cited more than 25 times (the data for more than twenty times is not available).[242] Thus, scientists go to conferences, they communicate and publish their communications in proceedings but they don't cite those communications. Their informal communication most likely takes place through the electronic medium, and the formal communication through the journal article and the distribution of preprints.

One may ask, then, if scientists are communicating adequately through modern technology, why do they have the need to go to conferences? The answer is provided by the conference literature itself: that it is at conferences where the really useful and

---

[242]David Pendlebury, personal communication, 6 August 1993; Garfield, *Essays*, 13: 46.

significant exchange between researchers takes place; that is where new ideas are communicated and many new others develop as the result of communication. It is at conferences that members of invisible colleges or small networks meet face to face and exchange ideas and results and get new ideas. Communication theory suggests that *a field whose members communicate more frequently and across greater distances and/or institutional boundaries is more likely to attract resources for growth.*[243] Gordon Conferences which are small and elite gatherings by invitation only have gained notoriety partly because of their informal atmosphere but also because no proceedings or any other records are published as a result of the meetings.[244]

In a way, conferences represent anomalous behaviour because they cause a concentration of communication behaviour among scientists. However, this *anomaly* should not be interpreted as a confirmation of Kuhn's ideas on anomalous science, but simply seen as a producer of outliers during the growth of literature. Anomalous science is a most dynamic science: due to contacts and increased communication it is younger, it is less consolidated, and it is very stimulating. However, the fact that more than 97% of the outliers represent conference papers should not be taken to mean that physics between 1977 and 1987 was made of nothing but anomalous science. Quite the contrary. If it had been dynamic and anomalous, the data would

---

[243]Leah A. Lievrouw, "Communication, Representation, and Scientific Knowledge: A Conceptual Framework and Case Study," *Knowledge and Policy: The International Journal of Knowledge Transfer and Utilization* 5 (1992): 6-28.

[244]Ibid., 22.

have revealed many more information epidemics. The fact that scientists are concentrated in a place for a short period of time does not necessarily mean that they are in the process of transforming their field.

In some burgeoning fields some of the ideas taken up at conferences may lead to growth. In some others they do not amount to much. This difference is reflected in the data obtained from this study. In contrast to those chapters with information epidemics, some chapters remained entirely flat, only punctuated by occasional outliers that were due to regularly held conference proceedings. Four examples will suffice: Chapter 29 (Experimental methods and instrumentation for elementary-particle and nuclear physics), Chapter 67 (Quantum fluids and solids; liquid and solid helium), Chapter 75 (Magnetic properties and materials), and Chapter 77 (Dielectric properties and materials). Chapter 29 contains the biennial *Particle Accelerator Conference*, Chapter 67 contains the triennial *International Conference on Low Temperature Physics*, Chapter 75 contains the triennial *International Conference on Magnetism*, and Chapter 77 contains the biennial *International Meeting on Ferroelectricity* as well as the *European Meeting on Ferroelectricity*. All four remained essentially flat over eleven years, seemingly unchanged despite the occasional high attendance (as evidenced by the number of contributions) at conferences. Thus, they all show that knowledge in a field is not always susceptible to stimulation. In the short-term it can be pushed by conferences, but unless a theoretical or experimental breakthrough takes place subsequently, activity usually gets back to the base level. Occasional step changes do not always perturb the stability of a field; that perturbation has to come from within.

In this respect, the superconductivity graph in Figure 28 (page 190) may at first appear remarkable and surprising, but on second thought, it makes a lot of sense: researchers will advance a number of ideas at conferences (after all, they all have to be original) but only those that pan out and are widely accepted (in part, through peer review) get published in journals. This control means that only what is regarded as significant gets published in the journal literature. What is highly significant over time gives rise to major growth patterns. And herein lies the significance of information epidemics. They reflect ideas or currents that hit the essence, the overall activity of a field until participants have had the time and the opportunity to judge them and decide whether these ideas are of consequence to the core knowledge of the field and whether they are worth pursuing. This is clearly reflected in the dynamics of Chapters 36 and 73 which have evolved in parallel with major conferences. As the fields grow so does the conference participation in them. As success breeds success, the intensity of a field rises, work accelerates and the output increases.

One source of stimulation in science may come from professional associations that are often involved in the diffusion of ideas and innovations. In physics some of these major association are the American Physical Society (APS), the American Institute of Physics (AIP), the European Physical Society and the International Union for Pure and Applied Physics (IUPAP). Of the four examples of flat chapters punctuated by regular conferences mentioned above (Chapters 29, 67, 75 and 77) all have their conferences sponsored by one or more of these associations. In addition, the *Particle Accelerator Conferences* of Chapter 29 are also sponsored by the Institute of Electrical and Electronics Engineers (IEEE).

There is evidence from the literature that the degree of involvement by the professional associations depends for the most on three factors: professional development activities, communication networks, and size.[245] These associations diffuse knowledge to their members through their own publications (especially journals and conference proceedings) and create networks that link their members to other organizations. While the literature does not mention it, practice indicates that the profits obtained from conferences provide significant motivation for associations to organize meetings that are as large as possible. The profit motive for many associations may be one reason why conferences are held regularly even in fields that do not grow significantly over a decade or from conference to conference.

Finally, the role of popular magazines and newspapers in the creation and sustenance of information epidemics cannot be overlooked. Given the coverage of physics by the popularizing media such as magazines like *Scientific American* or *New Scientist* and newspapers like *The New York Times* with its Tuesday science section, it is also likely that information epidemics do indeed occur, that sudden excitement still grips certain fields of study, but that such study is far from the norm and that this

---

[245]Sue Newell and Jacky Swan, "Professional Associations as Important Mediators of the Innovation Process," *Science Communication* 16 (1995): 371-87.

analysis at the chapter level is too coarse to identify them adequately. One must at the same time point out that the popularizing media itself influences scientific currents. Thus, an additional factor is that of value judgement, that is what becomes labelled as significant by the lay scientific press.[246]

Public attitudes toward science and technology are largely shaped by the media (i.e. papers, magazines, radio and television) and, worse, by the tabloids. In fact, those who are influenced by the popularizing press (but not necessarily the tabloids) are not only the public but researchers as well. A recent study by Phillips et al. has demonstrated that *researchers are more likely to cite papers that have been publicized in the popular press.*[247] Articles in the *New England Journal of Medicine* that were covered by *The New York Times received a disproportionate number of scientific citations in each of the 10 years after the Journal articles appeared.*[248] It is not inconceivable that some of the popularity of some subjects, and of fields, is due to the influence of the popular press on the researcher. Admittedly, successful cases like high temperature superconductivity and buckyballs are rare, but so are failures such as cold fusion. Journalists and popularizers at times take a few sensational cases and

---

[246]Dorothy Nelkin, *Selling Science: How the Press Covers Science and Technology* (New York: Freeman, 1987).

[247]D.P. Phillips, et al., "Importance of the Lay Press in the Transmission of Medical Knowledge to the Scientific Community," *New England Journal of Medicine* 325, 17 October (1991): 1180-3.

[248]Ibid., 1180.

blow them out of proportion. These are dramatic examples of exciting science but by no means do they typify how science works on the average.

7.5. Why do Information Epidemics Occur?

Section 7.3 discussed how and how often information epidemics occur in the physics literature of 1977-1987. This section will seek to provide possible explanations as to why they occur.

One of the first causes that comes to mind is federal legislation as a source of scientific and technical innovation, and, by extension, of information epidemics. Recent and well publicized examples include the elimination of CFCs from the atmosphere, pollution-related technology, and health-related research. This is also true of the U.S. space program. However, there is no evidence for this in the case studies found so far. To the contrary, looking at legislation as a source of epidemics puts the cart before the horse. Information epidemics, and especially knowledge epidemics, first wake people up to something significant that is happening in a specialty and thus precede any public action.

The same goes for the general influx of funding. What happens before and after epidemics would suggest more a scientific or technical cause than a financial one. In most countries it is national or federal programs that provide the vast majority of the funding for research and innovation. For example, in the United States the National Science Foundation and in Canada The National Research Council of Canada and The Social Sciences and Humanities Research Council provide strategic theme and block grants and thus determine areas that get funding. It is quite possible that such areas may show significant growth over a number of years as a result of piecemeal but concerted effort. However, one would not expect dramatic financial input unless there were serious scientific or technical reasons to begin with. There is no evidence as yet that any of the information epidemics in this study arose due to any a priori targeted funding. On the contrary, as the OTA report mentioned above suggests, the influx of money into high temperature superconductivity in the United States came well after the Bednorz and Muller discovery. The increasing attention paid to so-called AMO (Atomic, Molecular and Optical Science), and especially cluster physics, in the United States came in the 1980s, well after the start of the exponential phase of Chapter 36.[249]

For these reasons, the causes of information epidemics should be seen to reside in scientific and technical advances, especially in the form of surprises and in more complex human social and psychological factors. Surprises have been mentioned above, and the best example remains that of superconductivity. Technological advances have also been mentioned, for example, in the progress of cluster physics and the fractional quantum Hall effect. Technical advances enable researchers to conduct certain experiments or make crucial materials cheaper to obtain, but, most

---

[249]National Research Council. Panel on the Future of Atomic, Molecular, and Optical Sciences. *Atomic, Molecular, and Optical Science* (Washington, DC: National Academy Press, 1994).

importantly, they speed up the gathering of data. Continuing advances in computer technology and data processing as well telecommunications technology allow the rapid analysis of data and exchange of information. Thus, the crucial advantage conferred by materials and technology is not their existence or availability per se but the fact that they enable others to join the fray much more rapidly and cheaply. This is why the information epidemics found by this study suit the appellation *enabling technology*.

Francis Crick remarks that *it's possible to discern the speed of work in a given field by the way researchers use the term 'recently'. 'Recently' in neurobiology means the last 2 or 3 years, ... but in molecular biology it usually means the last 2 or 3 weeks!*[250] Technique becomes the *great leveler: a new technique can act as a democratic influence, bringing parity into a field where individuals are separated widely by their technical skills. What was once enormously difficult, and could be done only by the most highly skilled scientists, can now be done by almost anyone.*[251] Evidence for this comes from Von Klitzing whose first experiments were very difficult, being run at zero degree K, from high temperature superconductivity with easy to produce materials, and from cluster physics with a new laser ablation instrument to produce fullerenes cheaply. In general, then, *in the absence of an essential technique a researcher or a field flounders, developing elegant theories that*

---

[250]Stephen S. Hall, "How Technique is Changing Science," *Science* 257, 17 July (1992): 344-9.

[251]Ibid., 349.

cannot be decisively accepted or rejected - no matter how many intriguing circumstantial observations are available. But with a key technique in hand, the individual and field move ahead at almost terrifying speed, finding the right conditions to test one hypothesis after another.[252] This is certainly what happens with information epidemics.

Another argument comes from the work of Diana Crane who states that models exhibiting breadth but not testability can attract considerable interest, but are most likely to exhibit spurts of growth followed by a sharp decline.[253] Cold fusion provides the perfect example here.

The social and psychological factors in the causation of information epidemics are far more complex to ascertain. They are not the focus of this study and therefore it is not appropriate here to digress too far into these topics. However, it is worth attempting a possible explanation based on recent work concerning mass behaviour, threshold levels and social networks.

One possibility is that what an influential work does is to lower the threshold or resistance to diverse opinion that may hinder collective progress. Those with the lowest thresholds are (by definition) the most ready to react to the influential work, and they themselves produce other works in the same order which lowers the threshold for many other susceptible researchers in the field. With thresholds lowered

---

[252]Ibid., 345.

[253]Crane, *An Exploratory Study of Kuhnian Paradigms*, 37.

and thus with many barriers against communication lowered or removed, resistance in the pathways of networks is lowered or eliminated. More researchers join in, a critical mass develops and the field becomes ripe for take-off. Publishing becomes easier in journals where editors make increased allowance for the new theory (or craze). This fits with evidence from superconductivity in 1987 and 1988 and with low journal rejection rates by Zuckerman and Merton.[254] In addition, the influence overshoots the confines of invisible colleges or existing networks. By contact with other networks, new workers flow into the field, networks lose their rigidity and confining role. The resulting free for all ends up producing a large amount of work that in the published literature gets reflected as an information epidemic. In terms of communication among researchers, *the hustle and bustle of electronic mail, proliferating research literature, publication and citation indicators, and frequent international conferences, creates a general atmosphere of urgency* that becomes part and parcel of crowd behaviour during an information epidemic.[255] In sum, an influential or exciting work lowers thresholds, eliminates network barriers, increases contacts and with the ensuing increased communication, attracts new actors into the field.

---

[254]H. Zuckerman and R.K. Merton, "Patterns of Evaluation in Science: Institutionalization, Structure and Functions of the Reference System," *Minerva* 9 (1971): 66-100.

[255]John Ziman, *Prometheus Bound: Science in a Dynamic Steady State* (Cambridge: Cambridge University Press, 1994), 143.

The publication of the 1986 Bednorz and Muller paper was so surprising that Paul Chu at Houston changed his approach immediately, and almost overnight converted his laboratory to verify and confirm the B&M paper. Subsequently, he and his co-workers (K.W. Wu among them) synthesized new materials that were conductive at even higher temperatures. Here is another argument why legislation or funding cannot be accepted as causes of information epidemics. It takes too long to draw up legislation, and obtaining funding requires prior application. Information epidemics ensue all too rapidly for these two to be of any significance.

Due to the scarcity of information epidemics in the data, one needs to be cautious with generalizations. These results from the literature of physics within a restricted time period notwithstanding, information and knowledge epidemics may be more the norm in current science than the evidence so far has demonstrated, in part due to advances in materials and instrumentation, in part due to the increasing globalization and interdisciplinarity of scientific activities that force countries and research groups of diverse origins to compete with one another. On one hand, increasingly collaborative (but also competitive) activity on the international scene and, on the other, the shortage of funding, manpower as well as driving forces of major ideas may be resulting in the crowding of many workers into a few active and exciting research areas, with the consequence that a larger number of publications appear around the same subject specialties within a short period of time. Again, information epidemics in physics are scarce, but they may be more plentiful in other fields such as computer science, engineering, chemistry, medicine or molecular biology. There is evidence that this is indeed the case in molecular biology.[256]

7.6. A Model for Fast Growing Literatures

The characteristics of a fast growing literature enumerated in Section 1.1 (page 11) above were summarized as: rapid citation impact, high citation frequency, and swift diffusion.[257] However, the causes of these events had not been hypothesized. A vector was missing. The risk of looking for causality in social phenomena notwithstanding, it can now be stated that in some fast growing literatures the origins can be traced to one or a group of influential (sometimes seminal) articles that, as stated before, attract a new group of workers and induces them to publish extensively.

The characteristics of a fast growing area of science have been described previously by De Mey.[258] He puts the emphasis on the fact that most references in articles are recent, the articles are short, they carry few references but a high percentage of self-references, references are predominantly to journal publications (outcome of the work of Small and Crane[259]), there is a low rejection rate from

---

[256]B. Balmer and B. R. Martin, "Who's Doing What in Human Genome Research?" *Scientometrics* 22 (1991): 369-77.

[257]Hurt and Budd, *Modelling the Literature of Superstring Theory*, 472.

[258]De Mey, *The Cognitive Paradigm*, 119-24.

[259]Henry Small and Diana Crane, "Specialties and Disciplines in Science and Social Science: An Examination of Their Structure Using Citation Indexes," *Scientometrics* 1 (1979): 445-61.

journals (outcome of the work of Zuckerman and Merton[260]) that the proportion of secondary literature in the total is very small (outcome of the work of Menard[261]) and that *information disseminates fast and efficiently.*[262]

The results of this work largely support these propositions. In the three cases of Chapters 36, 73 and especially Chapter 74 all three factors of Budd and Hurt are valid. As mentioned above, the group of influential works diffuse very rapidly and obtain a high number of references. The authors themselves often publish other highly cited works. Although De Mey's criteria were not tested specifically for the results of this work, a cursory examination of the epidemics tends to support his arguments. Most articles are short, refer to recent articles, and carry a large number of self-references. One reason for the small number of references is that many are conference submissions and as such are limited in space and usually do not carry as many references as journal articles. In the case of journal articles in the epidemic of Chapter 74, many carried a large number of references (that is more than 20). The authors themselves give a large number of citations to those in the influential group leading the epidemic. In that respect it is difficult not to agree with Stephen Cole's contention that those scientists at a research frontier derive most of their information

---

[260]Zuckerman and Merton, *Patterns of Evaluation in Science*, 66.

[261]Menard, *Science: Growth and Change*, 1971.

[262]De Mey, *The Cognitive Paradigm*, 123.

from and thus cite a small number of active scientists in the same research frontier.[263] It seems that there is a two tier system, and that most scientific work is 'bread-and-butter' science rather than ground breaking work.

As mentioned in the case of superconductivity, the fact that journal editors facilitated the publication of new work on superconductivity supports Zuckerman and Merton's findings on low rejection rates.[264] It is also true that the proportion of secondary literature is small. The difference between the total number of abstracts in each epidemic and the sum of journal articles and conference papers, that is the number of other items, is rather small. Most are composed of technical reports. The proportion of remnants such as monographs and monograph chapters are even smaller. They only increase in numbers as the information epidemic becomes a knowledge epidemic and the new field becomes institutionalized into a new specialty area, as in the case of superconductivity after 1987.

Based on the above mentioned arguments and results, it is now possible to formulate a preliminary model for fast growing literatures. A fast growing literature is one where journal articles grow exponentially over a number of months, one that is led by a group of highly influential works that obtain several hundred citations within a short period of time after publication, and where the doubling of the output in the field takes place very rapidly.

---

[263]Stephen Cole, "The Hierarchy of the Sciences?" *American Journal of Sociology* 89 (1983): 111-39.

[264]Zuckerman and Merton, *Patterns of Evaluation in Science*, 66.

The path for the spread of information epidemics is paved by a group of leaders and their influential works. Significant works and novelties remain important only if they are tested and accepted by the leaders of their fields. If something new cannot also be accepted by the leading institutions and research groups, then it is not significant and durable enough, and it will wither and be discarded until further evidence can be brought to bear in favour of it (as in the case of cold fusion).

It can therefore be hypothesized that if an influential work gives rise to at least two other highly cited works within a short period of time, then the field will attract new interest and it will go epidemic. A highly cited group of works, with the core leadership it provides, plus the international participation it invites, will give rise to an information epidemic. So long as the core leadership is renewed periodically, and theoretical difficulties remain in the field, the information epidemic will persist until a solution is found. If, in addition, the epidemic has major industrial and financial implications or brings about changes in the fundamental theory, then the information epidemic will turn into a knowledge epidemic which will then give rise to a new specialty. The works that make up the core leadership will be characterized by a rapid citation impact, high citation frequency and swift diffusion.

7.7. Significance of the Results

This dissertation makes several contributions to information science and related fields.

1. With respect to information science, it provides an approach to the description of scientific information and the evolution of a field over time by understanding the data from *Physics Abstracts*.

2. It provides a contribution for those interested in understanding the growth of science. It shows that it is better to examine it in smaller intervals of monthly data as opposed to yearly output of data. None of the results here would have been observed (other the superconductivity) if the data had not been disaggregated from yearly into monthly periods. As such, it provides quick identification of epidemic areas and allows to have a sensitive test for fast growing areas. Concept of information epidemics may become the model for the growth and decline of fields of research.

3. It provides an increased understanding of information epidemics as a useful and usable concept in information science. As such, it brings to the study of fast growing literatures a model that can now be tested. The results of this work can be used toward theory building in information science.

4. It confirms and extends the epidemic model in the growth of literatures. Not only is there a contagion effect but there is also a catalyst effect, and the catalyst effect may be more widespread.

5. With respect to the sociology of science, this study shows that anomalies do occur as part of the behaviour of normal science. Here it is possible to obtain a statistical description of anomalies in science. Therefore, this thesis offers a methodological tool for explaining abnormal science. The interpretation here is based on the data obtained

from the science itself. For example, so far no one has dealt with conferences as abnormal science.

6. With respect to the history of science, it provides a confirmation of physics as a mature science. Especially the *big science* portion of physics is now in a decline, and excitement is expected to come from *little science* and the production of novel materials.

7. With respect to the philosophy of science, it provides added confirmation for Kuhn's ideas on normal versus anomalous science and on how changes in paradigms are reflected in scientific literature. It also adds to the understanding of *normal science* by showing that normal science punctuated by growth spurts can result in significant progress.

8. With respect to science policy it provides a methodology for the identification of those successful areas that deserve attention and the influx of funds. It also allows the evaluation of the effectiveness of certain science policies, especially in the effectiveness of certain stimulative programs.

7.8. Alternative Interpretations of the Results

It is possible that the epidemics ascribed to results obtained from this study are in fact accidental growth points that accompany all stochastic and noisy data and that so-called information epidemics are nothing more than attaching significance to spurious data. For example, one may argue that it is not appropriate to pay any importance to information epidemics simply because the results observed are merely comparative: those fields with information epidemics are simply producing more. Also, not all fields get the same amount of funding. It is possible that those producing information epidemics have initially and for unrelated reasons attracted more funding to begin with, and results obtained from a totally unexpected angle happen to be popular temporarily. However, it is possible to counter that the epidemics observed are not between fields of study but the different states of one field before and after a causative mechanism. This suggests that the cause may be more a scientific or technical one. The argument becomes a bit of a chicken-and-the-egg problem until the causative mechanisms have been clearly identified. Still, one would not expect to find an influx of funds (or even new workers) before the occurrence of something scientifically or technically significant. Even if the results came from an unexpected source, the fact remains that they are significant by the very fact that the field grows significantly, new workers come in and certain publications become influential leaders.

It could even be that a discovery seemingly revolutionary may in fact be cumulative. As Landsberg puts it

> the advance of science, which sometimes seems to proceed by a big leap, can by careful historical study often be seen to actually be the result of a slow step by step advance. ... The big steps in the advance

> of science are often big only because the little intermediate steps have been forgotten or ignored.[265]

Another problem is one of superficiality: The study should have gone deeper into subspecialty areas. However, if this had taken place, even more epidemics could have been found, such as in the explosion of a supernova in 1987 and on superstrings. Nevertheless, this is a valid criticism and is one that deserves further work. Future work needs to cover the period since 1987 and all of the sections and subsections of *Physics Abstracts*. At the same time, it should be pointed out that the more detailed the specialty examined, the less literature there is available in it and the more the analysis becomes susceptible to mistakes.

One further argument may be inspired from Diana Crane:

> one implication might be that social factors are much more important in recruitment into research areas ... than cognitive factors. In other words, such fields attract new members because they become fashionable rather than because of their scientific potential.[266]

This is true, except that it is difficult to see how a fad can be sustained over a number of years at the intensity shown by growth in cluster work or the quantum Hall effect, not to speak of superconductivity. These fields are fashionable because of their scientific potential, and this has been demonstrated sufficiently.

---

[265]Peter T. Landsberg, "Problems of Explanation in Physics," in *Beyond Belief: Randomness, Prediction and Explanation in Science* ed. John L. Casti and A. Karlqvist (Boca Raton, FL: CRC Press, 1991), 55-64.

[266]Crane, *Invisible Colleges*, 90.

Finally, it has so far been assumed that all journals are equally available and are equally timely in arrival (see page 67 above). Of course, this is not true. Therefore, one possibility of information epidemics may be due to a grouping of a number of journals. Some of them may arise due to an inordinate lag time between the publication of a journal and its being abstracted and added to the database. It may also be due to some reason such as staff changes or sick leave due to which several issues of a journal cumulate until their indexation all at once. Such grouping is possible but rare. The one such instance in this study was with Chapter 93 (Geophysical observations, instrumentation, and techniques). It is doubtful even there that one or two journals have the power to give false positives or to change the course of a discipline. It takes far more than a journal or two to make an information epidemic.

Overall, then, while alternative explanations are possible, they cannot be sustained in face of the results obtained here. The results are not voluminous, but nor are they spurious.

7.9. Future Directions

There is a large amount of work that flows from the approach adopted and from the results obtained here. The first study that needs to be undertaken is testing Garfield's citation classics; do they give rise to epidemics? Citation classics cover practically all fields of science and belong to scientists of various capacities, many of them Nobel prize winners. If information epidemics arise due to influential works, it should be possible to test the reverse hypothesis that an influential body of authors and an influential set of works (as evidenced from their citation patterns) should also have given rise to information epidemics over time. A series published by ISI Press provides numerous examples.[267]

Do Nobel prize winners create information epidemics? The answer is at least a partial yes, since two cases examined in this thesis involve Nobel prize winners: Von Klitzing for the quantum Hall effect and Bednorz and Muller for superconductivity. However, can this be generalized (or even generalized for certain fields) or are these cases here coincidental? Most Nobel prizes are awarded years after the significant discoveries were made. In that respect, some types of knowledge may disseminate rapidly, but the long term implications are difficult to assess.

The second immediate task is to study all sections and subsections of *Physics Abstracts* and bring them up to date into the 1990s and find out whether there are many more information epidemics hidden within the data of chapters and whether the ratio of epidemics versus conferences changes by going to a higher level of detail. In addition, the database of growth statistics in *Physics Abstracts* can be supplemented with other databases of growth statistics in medicine, chemistry, electrical engineering, computer science and mathematics - areas where there are well-defined classification systems. These, together, can act as a science and technology watch

---

[267]Arnold Thackray, *Contemporary Classics in Physical, Chemical, and Earth Sciences* (Philadelphia: ISI Press, 1986).

program that can in the future, with the amelioration of techniques, be used as forecasting tools for the growth of current science.

Third, information epidemics should be sought among major scientific news items such as superstrings and the explosion of Supernova 1987A. Was the voluminous nature of these news just hype or was there substance to them?

A knowledge epidemic must be started by someone, by a leader who published a germinal work or puts forward an exciting idea. Is it possible to find this leader? Do certain people become leaders consistently, by being at the forefront of important work, by being at the right place at the right time and by publishing influential work continuously? There are major figures in physics such as John Bardeen who received two Nobel prizes in physics or P.W. Anderson (Nobel prize in physics for his work on superconductivity) who remain highly active, original and productive throughout their careers. What characteristics (other than genius) do these scientists possess that makes them consistent leaders? Is there any connection between their being an editor, being on the editorial board, on accreditation boards, on national advisory panels? Are these people vectors of epidemics: are they the ones that nurture key papers from which an epidemic emerges? There is no evidence one way or another provided by this study, but evidence from the field of nonlinear dynamics suggests that the so-called *chaos* revolution in the 1980s took place in large part due to the efforts of James Yorke who rediscovered Ed Lorenz' work and publicized it widely.[268] The

---

[268]James Gleick, *Chaos: Making a New Science* (New York: Viking, 1987), 65.

availability of technology then allowed corroboration and further advances. There is evidence suggesting that when an older work gets rediscovered and starts receiving new attention, it may be the harbinger of a new spurt of growth in the field.[269]

Another area of future work involves the possibility of prediction or anticipation. There are clear differences among terms such as prediction, forecasting, foresight, and anticipation.[270] These differences notwithstanding, is it possible that, once a specialty or an area of research starts going epidemic, it might be feasible to foresee how far will it go and how long will it take? While one may not be able to tell with certainty, it should be possible to model some curves that have started their exponential phase and follow them until they are past their mid-point, the so-called inflection point. Once the curve crosses over its inflection point, it should be possible to extrapolate the s-curve and wait for data to confirm the topping process. Such an approach has been used in innovation diffusion research.[271]

What predisposes works to become influential? Can one distinguish them on the basis of content, authorship or sponsorship? It is possible to put factors together and test via discriminant analysis to see how strongly factors discriminate with respect to independent variables. Is it also possible that those who generate epidemics have

---

[269]James R. Bright, "Improving the Industrial Anticipation of Current Scientific Activity," *Technological Forecasting and Social Change* 29 (1986): 1-12.

[270]Ben R. Martin and John Irvine, *Research Foresight: Priority-Setting in Science* (London: Pinter Publishers, 1989), 4.

[271]Nigel Meade, "Forecasting Using Growth Curves - An Adaptive Approach," *Journal of the Operational Research Society* 36 (1985): 1103-15.

more money, work at more prestigious universities or go to more conferences? High temperature superconductivity work can be done with very cheap materials, but the place where it was discovered (IBM Laboratories in Zurich) has (or maybe had) considerable funding and prestige.

There are also a number of other scientometric studies that can be undertaken, such as co-citation and social network analysis of epidemic areas before and after information epidemics with the purpose of identifying further causative factors. The usefulness of citation clustering in addition to classification in the study of the mathematical logic literature has already been suggested by Wagner-Dobler.[272] Do authorship and networks change during an information epidemic? Could a co-word analysis indicate the cognitive paths on the way to novelty and epidemic occurrences?

There is the study of scatter: one would expect to get less scatter of citations from an information epidemic (where workers are concentrating on a core literature) than from a major outlier made of all the papers of a conference. One can thus expect a relationship between the degree of focus shown in a research front versus that shown in the field as a whole. Thus, it would be instructive to follow growth and at the same time follow the number of citations from the core literature.

One further hypothesis to test involves the validity of Lotka's law during an information epidemic. Wagner-Dobler recently indicated that the validity of Lotka's

---

[272]R. Wagner-Dobler and J. Berg, "Regularity and Irregularity in the Development of Scientific Disciplines: The Case of Mathematical Logic," *Scientometrics* 30 (1994): 303-19.

law may be dependent on the stages of growth during an epidemic process.[273] In the initial stages of an epidemic there are few authors or pioneers. As the exponential phase is entered, new authors are attracted, and the pioneers shows enormous progress. In the exhaustion stage, the progress is not a rapid for the pioneers, and many entrants start leaving. Thus, the value of the contributions per time unit is not distributed evenly in a Lotkan manner. Is this valid for the epidemics at hand? Once the data from this thesis are updated to cover 1995 and beyond, Wagner-Dobler and Berg's hypothesis may be tested for agreement with the data of this study.

There is a number of factors indicating the relationship between quantity and quality in scientific production that can be tested in the data obtained from this study. Some examples are Rousseau's Law that states that the number of important articles in a field is the square root of the number published in it[274], the lambda quality level that shows an exponential relationship between categories of quality in a given literature and a citation test thereof[275], and a test of the Ortega Hypothesis in the epidemic literature and whether *a relatively small number of physicists produce work that becomes the base for future discoveries in physics.*[276]

---

[273]R. Wagner-Dobler and J. Berg, "The Dependence of Lotka's Law on the Selection of Time Periods in the Development of Scientific Areas and Authors," *Journal of Documentation* 51 (1995): 28-43.

[274]Rescher, *Scientific Progress*, 96-98; Hans Werner Holub et al., "The Iron Law of Important Articles," *Southern Economic Journal* 58 (1991-92): 317-28.

[275]Rescher, *Scientific Progress*, 98-103.

[276]Cole and Cole, *Ortega Hypothesis*, 372.

One further possibility is to examine whether a sample of major and significant journals could not give the same results as a whole survey. It might be sufficient to take a small number of core journals covering two-thirds to three-quarters of one of the epidemic fields from this thesis, following the output by subject over an appropriate period of time in order to see whether one can discern the same information epidemics in the data. In other words, is it possible to trace a number of core journals in order to discover information epidemics? If the answer is in the affirmative, it may be possible to simplify the methodology developed here further by identifying a few core journals and following fields with those rather than having to resort to exhaustive index and database searches. The list of core journals could belong to the core list published by the *Journal Citation Reports* of the *Science Citation Index*.

Another set of studies would be sociological in nature. One qualitative study would involve contacting physicists in different specialties, showing them the results of this study and asking them whether they see things the same way. Preliminary contacts made with specialists in cluster physics and the quantum Hall effect at McGill University and Université de Montréal suggest that the specialists agree with the approach to information epidemics and recognize the top names in ranked lists as belonging to those most active and significant contributors in their fields. However, they do not think the results to be of any consequence for the specialties as such.

The communication structure of information epidemics could be studied by following electronic communication and the value of the preprints available through the Internet, too early for the 1977-1987 literature but potentially valuable for literature starting with the 1990s. Are epidemics in the printed literature reflected in electronic scientific communication? Does electronic communication engender or contribute to epidemic behaviour? Are communication networks during an epidemic different from *normal science*? How does electronic mail affect the growth of literature? What is the role of technical reports during an information epidemic or are they forgotten entirely? What happens to co-authorship networks with electronic mail? Some work suggests that electronic mail does not stimulate new relationships but enhances the impact of ties already established.[277]

One further deficiency that this work brings to evidence is that much of the current understanding of the communication patterns in science resides on the work of Garvey and his co-workers. There has been no in-depth work on the changing patterns of communication in science since Garvey. Given the gap of some thirty years since his work and the changes in communication technology, it would be instructive to investigate how physicists in the 1990s communicate their ideas and obtain new ideas.

---

[277]Kathleen Carley and Kira Wendt, "Electronic Mail and Scientific Communication: A Study of the Soar Extended Research Group," *Knowledge: Creation, Diffusion, Utilization* 12 (1991): 406-40.

A large quantity of statistical work has been devoted in recent years to the understanding of the growth models of various scientific literatures. This has largely involved curve-fitting exercises with various degrees of success. Another angle to adopt in such studies would be to undertake computer modelling in order to better understand whether the emergence of information epidemics can be explained in terms of populations, critical mass, and threshold levels and how collective behaviour during moments of excitement can result in the rapid emergence of a body of literature. This model could be further used to test the veracity of the explanations offered in Section 7.5 above. A simulation of Kuhnian paradigm changes by a system dynamics model recently demonstrated that feedback processes are heavily involved.[278] Although such models do not require the presence of competition theories, a test of the appropriateness of Lotka-Volterra equations could be integrated into a study of the current competitive theories concerning the mechanisms of high-temperature superconductivity.

How are threshold levels lowered in response to influential works? What are the mechanisms involved? There is no doubt that, in addition to scientific and theoretical motivations, a number of cognitive factors are also involved. How such change comes about, how cognitive factors are involved in the maintenance of a certain crowd behaviour, how they sustain information and/knowledge epidemics, and

---

[278]John D. Sterman, "The Growth of Knowledge: Testing a Theory of Scientific Revolutions with a Formal Model," *Technological Forecasting and Social Change* 28 (1985): 93-122; Jason Wittenberg, "On the Very Idea of a System Dynamics Model of Kuhnian Science," *System Dynamics Review* 8 (1992): 21-33.

finally how saturation leads to fragmentation in a field of study and to the formation of a new specialty are all important questions that would further the understanding of the causes underlying information epidemics, and also in many ways understanding related types of populous behaviour such as fads, fashions, and riots.[279]

Finally, as a highly speculative venture, it would be instructive to undertake certain quantitative studies in order to increase the level of understanding of the dynamics of scientific literatures. Scientific growth models are qualitatively analogous to ecological or economic growth models. Some of the dynamics have been shown to be due to the intrinsic nature of the data and not necessarily externally driven. Is it possible that some of the spikes observed, and consequently some of the information epidemics, are due to internal dynamic reasons such as intermittency or are they random shocks?[280] Whatever the answer, a test would require voluminous data of the kind obtained in this study. Yearly data would just not suffice. Even conceptually, it is often rather difficult to ascribe physical causes to social data. Due to the mathematical sophistication and the methodological rigour necessary, such work would have to be undertaken with great care and would perforce be a multidisciplinary study.

---

[279]Michael Hammond, "Finite Human Capacities and the Pattern of Social Stratification in a Knowledge Society," in *The Knowledge Society* ed. G. Bohme and N. Stehr (Dordrecht: Reidel, 1986), 31-50.

[280]Bernardo Huberman, Social Intermittency, Preprint, Xerox Corporation, 1993.

CHAPTER 8

CONCLUSIONS

In direct response to the stated objectives, this study has conclusively shown:

1. the prevalence of information epidemics in the physics literature between 1977 and 1987,

2. the presence of a new growth mechanism in short term literature growth entitled *information epidemics*,

3. the utility of micro-level studies in the growth of scientific literature,

4. the utility of the triangulation method in analyzing information epidemics.

It is now possible to answer the questions posed previously on pages 2 and 24 clearly: 1. Yes, the yearly attention to data does result in extreme smoothing. To paraphrase Diana Crane, the empirical study of information epidemics in science should proceed with the location of exceptionally rapid growth , as indicated by the number of publications abstracted monthly.[281] 2. Yes, some of the spikes **are** caused by groups of influential papers that affect the practice of a given specialty. 3. (from page 24): Yes, **the outliers**, that reflect a small number of the spikes, **are** meaningful in terms of the growth of knowledge. In some cases, such as that of superconductivity, they give rise to knowledge epidemics and to a structural change in the future of a given field of study. At the same time, the model presented on page 4 is also supported.

---

[281]Crane, *Fashion in science*, 441.

This study also provides empirical confirmation for the steady-state thesis, especially as it pertains to physics between 1977 and 1987. It shows that the pace of major discoveries was rather slow during this period, and most success came from smaller science that was capable of producing novel materials and procedures. Thus, information epidemics, while they exist, are unusual and emanate from specific circumstances.

A domain needs to be *fertile* for change, and this fertility and resultant change can be brought about by either of two circumstances. There could be historical reasons, where the field of study is of long standing interest and progress would have many useful social consequences. The reason could also be theoretical where existing theory is inadequate or where a large and encompassing theory does not exist, and new work has the potential of either solidifying existing theory or being nefarious to its continuing existence. Thus, information epidemics will tend to develop in areas where either there are serious theoretical problems or where there is major potential for finalization. The information epidemics found in this study cover both.

In large part, the rapidity in the genesis of an information epidemic will follow one of two models at hand, either the contagion model or the catalyst model. So far, as it pertains to the results of this work, it is difficult to mix the two models. However, it is possible to find a mechanism where work in a given specialty will develop slowly along many lines, will follow a catalyst model until it crosses a threshold due to a surprising discovery or a major analytical effort, and will then follow the contagion model. The reverse would be more difficult to qualify. If a field of work following a contagion model slows down and then settles to follow a catalyst model, it would be regarded as a typical information epidemic where an exciting idea or discovery did not work out and those initially attracted to work and publish in it left the specialty for other more interesting pursuits. Once the work accumulates and crosses a threshold, a renewed epidemic process begins and new actors are attracted once again.

This study also found that the information epidemics identified share several characteristics, especially as they pertain to advances in instrumentation and/or to the availability of cheaper and more plentiful materials. They all are *enabling sciences*, in the sense that their results give rise to progress in other fields as well as technological opportunities. They all share serious theoretical concerns and can be qualified as puzzle generating domains. Where an existing puzzle is solved, one would not expect to find epidemic behaviour. In other words, confirmations do not produce epidemics. At the same time, one would expect a field's publication numbers to shoot up when it comes up against theoretical difficulties.

Overall, the epidemics identified reflect the rapid transmission of information as a result of the spread of an exciting idea and the resultant growth of the specialty. The direction of growth of a specialty, it was found, could be determined solely by paying attention to journal articles. Conference papers and other document types confer a spurious quality to the data and remain spikes of little consequence. This result also confirms the new and increasing archival value of the journal in communication among scientists as well as the classical importance of the conference paper for the announcement and sharing of ideas and discoveries. It is seen that if certain discoveries are of serious import, attention turns to the publication of journal articles and the securing of priority in research through the periodical medium. It is the printed periodical medium that provides science with one of its pillars of stability.

There are several implications from this study. For the scholarly understanding of science, it provides evidence for a third major mechanism of growth of literature. In terms of theory building, it provides a model for the mechanisms in the genesis of fast growing literatures that can be further tested. In terms of the professional practice of librarianship, it provides the librarian with potential information on significant authors and institutions of research. In its implications for future research, it provides a framework for the quantitative study of literature dynamics.

Future work will need to cover an increasing number of sciences, fields and specialties as well as bring up to date the data obtained for this study. A research program of *science watch* carries much potential for historical understanding as well as close range anticipation.

It is comforting to ascertain that most of these results could not have been obtained by adhering to the analysis yearly numbers of publications. The advantage obtained from the monthly charts of publications indexed by *Physics Abstracts* was vital for the results of this study.

This study was designed to reflect the state of science in a domain as it is evolving. Modern science is at the same time a slower as well as a far more dynamic activity than seen in the reflection of biographies and histories. On the one hand, much of science still consists of routine work. Moments of excitement are infrequent. In physics they were rare between 1977 and 1987, and the literature reflects that. At the same time, unusual and seemingly important discoveries are made quite often. They are increasingly being covered by the news media. Fields of specialty are born and die every day. Scientists in quest of interesting research questions move in and out of domains all the time. From year to year groups of graduates target different fields selectively. The number of scientists active and the volume of work produced ensures that there is some interesting result coming from somewhere all the time. Much of science is still seen as a harbinger of better times ahead. Thus, the occasional publication causes temporary or longer term excitement because it fuels future hopes. It awakens in the scientist as well as the public a feeling of progress and a possible solution to a longstanding societal problem. A yearly model for following this continuous flux would miss the momentum and most of the dynamics of change. As it is, a certain lag time in publishing and indexing gives even *current* bibliometric studies an air of historicity. Nevertheless, short term attention cuts the information loss considerably and increases the sensitivity of the methodology used in the study of information epidemics.

CHAPTER 9

BIBLIOGRAPHY

Allen, Robert S. "Physics Information and Scientific Communication: Information Sources and Communication Patterns," *Science and Technology Libraries* 11 (1991): 27-38.

Anderson, Frances. *New Approaches to Research Policy Using Bibliometrics.* St. Foy, Québec: Conseil de la science et de la technologie, 1987.

Anderson, P.W. *The Economy as an Evolving Complex System.* Toronto: Addison-Wesley, 1988.

Anderson, Richard C., Francis Narin, and Paul McAllister. "Publication Ratings versus Peer Ratings of Universities." *Journal of the American Society for Information Science* 29 (1978): 91-103.

Anthony, L.J., H. East and M.J. Slater, "The Growth of the Literature of Physics." *Reports on Progress in Physics* 32 (1969): 709-67.

Arnison, G. et al. "Experimental Observation of Lepton Pairs of Invariant Mass Around 95 GeV/c/sup 2/ at the CERN SPS Collider," *Physics Letters B* 126B (1983): 398-410.

Avramescu, A. "Modelling Scientific Information Transfer." *International Forum on Information and Documentation* 1 (1975): 13-9.

________. "Coherent Informational Energy and Entropy." *Journal of Documentation* 36 (1980): 293-312.

Baggot, Jim. *Perfect Symmetry: The Accidental Discovery of Buckminsterfullerene.* Oxford: Oxford University Press, 1994.

Balmer, B. and B. R. Martin. "Who's Doing What in Human Genome Research?" *Scientometrics* 22 (1991): 369-77.

Barschall, H.H. and W. Haeberli. "Conference Proceedings in Physics." *College and Research Libraries* 53 (1992): 563-6.

Bartholomew, D.J. "Continuous Time Diffusion Models with Random Duration of Interest." *Journal of Mathematical Sociology* 4 (1976): 187-99.

Beck, Leonard N. "Soviet Discussion of the Exponential Growth of Scientific Publications." In *The Information Conscious Society: Proceedings of the Thirty-Third ASIS Annual Meeting, Philadelphia, 11-15 October 1970*. Edited by J.B. North. Philadelphia: American Society for Information Science, 1970, 5-17.

Bednorz, J.G. and Kurt A. Muller. "Possible High Tc Superconductivity in the Ba-La-Cu-O System." *Zeitschrift fur Physik B - Condensed Matter* 64 (1986): 189-193.

Blalock, Hubert M. *Social Statistics*. Rev. 2d ed. New York: McGraw-Hill, 1979.

Bohme, G. & N. Stehr, eds. *The Knowledge Society: The Growing Impact of Scientific Knowledge on Social Relations*. Dordrecht: Reidel Publishing Company, 1986.

Bottle, R.T. and M.K. Rees. "Liquid Crystal Literature: A Novel Growth Pattern." *Journal of Information Science* 1 (1979): 117-9.

Bright, J.R. "Improving the Industrial Anticipation of Current Scientific Activity." *Technological Forecasting and Social Change* 29 (1986): 1-12.

Broad, William J. "The Publishing Game: Getting More for Less." *Science* 211 (1981): 1137-9.

Brodie, B.C. *Psychological Inquiries: In a Series of Essays, Intended to Illustrate the Mutual Relations of the Physical Organisation and the Mental Faculties*. 2nd. ed. London: Longman, Brown, Green, and Longmans, 1856.

Bruckner, E., W. Ebeling and A. Scharnhorst. "The Application of Evolution Models in Scientometrics." *Scientometrics* 18 (1990): 21-41.

Budd, J.M. and C.D. Hurt. "Superstring Theory: Information Transfer in an Emerging Field." *Scientometrics* 21 (1991): 87-98.

Builova, N.M., E.K. Zakharova and N.V. Akshinskaya. "How Information Publications Reflect Studies on High-Temperature Superconductivity." *Nauchno-Tekhnicheskaya Informatsiya, Ser. 1* 17 (8, 1990): 22-8

Bujdoso, Erno and Tibor Braun, "Publication Indicators of Relative Research Efforts in Physics Subfields." *Journal of the American Society for Information Science* 34 (1983): 150-5.

Burbeck, Stephen L., Walter J. Raine, and J. Abudu Stark, "The Dynamics of Riot Growth: An Epidemiological Approach." *Journal of Mathematical Sociology* 6 (1978): 1-22.

Burrell, Quentin L. "Predictive Aspects of some Bibliometric Processes." In *Informetrics 87/88: Select Proceedings of the First International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval, Diepenbeek, Belgium, 25-28 August 1987.* Edited by Leo Egghe and Ronald Rousseau. Amsterdam: Elsevier, 1988, 43-63

Carley, Kathleen and Kira Wendt. "Electronic Mail and Scientific Communication: A Study of the Soar Extended Research Group," *Knowledge: Creation, Diffusion, Utilization* 12 (1991): 406-40.

CDF Collaboration. "Evidence for Top Quark Production in pp Collisions at Square Root (s)=1.8 TeV," *Nuclear Physics B, Proceedings Supplements* 39 (1995, no.B-C): 343-7.

Chang, Hans and Dennis Dieks, "The Dutch Output of Publications in Physics." *Research Policy* 5 (1976): 380-96.

Chubin, Daryl E., A.L. Porter, and F.A. Rossini. "'Citation Classics' Analysis: An Approach to Characterizing Interdisciplinary Research" *Journal of the American Society for Information Science* 35 (1984): 360-8.

________ and Edward J. Hackett. *Peerless Science: Peer Review and U.S. Science Policy.* Albany, NY: SUNY Press, 1990.

Clark, K.E. *America's Psychologists: A Survey of a Growing Profession.* Washington, D.C.: American Psychological Association, 1957.

Close, Frank. *Too Hot to Handle: The Story of the Race for Cold Fusion.* Toronto: Penguin Books, 1992.

Cohn, S. "The Effect of Funding Changes Upon the Rate of Knowledge Growth in Algebraic Topology, 1955-75." *Social Studies of Science* 16 (1986): 23-59.

Cole, F.J. and N.B. Eales. "The History of Comparative Anatomy." *Science Progress* 11 (1917): 578-96.

Cole, Jonathan R. and Stephen Cole. "Measuring the Quality of Sociological Research: Problems in the Use of the Science Citation Index." *American Sociologist* 6 (1971): 23-9.

________ and Harriet Zuckerman, "The Emergence of a Scientific Specialty: The Self-exemplifying Case of the Sociology of Science." In *The Idea of Social Structure: Essays in Honor of R.K. Merton*. Edited by Louis A. Coser. New York: Harcourt-Brace-Jovanovich, 1975, 139-74.

Cole, Stephen. "The Hierarchy of the Sciences?" *American Journal of Sociology* 89 (1983): 111-39.

Cole, Stephen. *Making Science: Between Nature and Society*. Cambridge, MA: Harvard University Press, 1992.

________ and Jonathan R. Cole. "Scientific Output and Recognition: A Study in the Operation of the Reward System in Science." *American Sociological Review* 32 (1967): 377-90.

Cozzens, Susan E. *Literature Based Data in Research Evaluation: A Manager's Guide to Bibliometrics*. London: Science Policy Support Group, 1990.

Crane, Diana. "Fashion in Science: Does it Exist?" *Social Problems* 16 (1968): 433-41.

________. *Invisible Colleges: Diffusion of Knowledge in Scientific Communities*. Chicago: University of Chicago Press, 1972.

________. "An Exploratory Study of Kuhnian Paradigms in Theoretical High Energy Physics." *Social Studies of Science* 10 (1980): 23-54.

Crease, Robert P. and Charles C. Mann. *The Second Creation: Makers of the Revolution in 20th Century Physics*. New York: Macmillan, 1986.

Cronin, Blaise. "When is a problem a research problem?" In *Applying Research to Practice: How to Use Data Collection and Research to Improve Library Management Decision Making*. Edited by L.S. Estabrook. Urbana-Champaign, IL: University of Illinois Graduate School of Library and Information Science, 1992, 117-32.

Czerwon, H.J. "Scientometric Indicators for a Specialty in Theoretical High-energy Physics: Monte Carlo Methods in Lattice Field Theory." *Scientometrics* 18 (1990): 5-20.

DeBresson, C. "Predicting the Most Likely Diffusion Sequence of a New Technology through the Economy: The Case of Superconductivity," *Research Policy* 24 (1995): 685-705.

De Mey, Marc. *The Cognitive Paradigm: An Integrated Understanding of Scientific Development.* Chicago: University of Chicago Press, 1992.

Dietz, Klaus. "Epidemics and Rumours: A Survey." *Journal of the Royal Statistical Society, Ser. A* 130 (1967): 505-28.

Dou, Henri, P. Hassanaly, L. Quoniam, R. Blanc and J. Pullino. "La physique à Aix-Marseille: indicateurs et évolution à partir de l'analyse automatique de la base de données INSPEC." *Revue francaise de bibliometrie* 3 (1988): 49-85.

_______, _______ and _______. "Infographic Analytical Tools for Decision Makers: Analysis of the Research Production of the Sciences." *Scientometrics* 17 (1989): 61-70.

Easingwood, Christopher J. and Simon O. Lunn. "Diffusion Paths in a High-tech Environment: Clusters and Commonalities." *R&D Management* 22 (1992): 69-80.

Edge, David. "Quantitative Indicators of Communication in Science: A Critical Review." *History of Science* 17 (1979): 102-34.

Egghe, L. and I.K. Ravichandra Rao. "Classification of Growth Models Based on Growth Rates and its Applications" *Scientometrics* 25 (1992): 5-46.

Etzkowitz, H. et al. "The Paradox of Critical Mass for Women in Science," *Science* 266 (1994): 51-4.

Fan, David B. "Ideodynamics: The Kinetics of the Evolution of Ideas." *Journal of Mathematical Sociology* 11 (1985): 1-23.

_______. "Ideodynamic Predictions for the Evolution of Habits." *Journal of Mathematical Sociology* 11 (1985): 265-81.

Fleischmann, Martin and Stanley Pons. "Electrochemically Induced Nuclear-Fusion of Deuterium." *Journal of Electrochemical Chemistry and Interfacial Chemistry* 261 (1989): 301-8.

Fractional Quantum Hall Effect Indicates Novel Quantum Liquid, *Physics Today* 36 (July 1983): 19-22.

Frumau, Coen C.F. "Choices in R&D and Business Portfolio in the Electronics Industry: What the Bibliometric Data Show." *Research Policy* 21 (1992): 97-124.

Garfield, Eugene. *Essays of an Information Scientist*. Philadelphia: ISI Press, 1977-1994, 17 vols.

________. *Citation Indexing*. New York: Wiley, 1979.

________. "Is Citation Analysis a Legitimate Evaluation Tool?" *Scientometrics* 1 (1979): 359-75.

________. "The Most-Cited Papers of All Time, SCI 1945-1988, Part 4: The Papers Ranked 301-400." *Current Contents* 21 (27 May 1991): 5-16.

________, Morton V. Malin and Henry Small. "Citation Data as Science Indicators." In *Toward a Metric of Science: The Advent of Science Indicators*. Edited by Yehuda Elkana, Joshua Lederberg, Robert K. Merton, Arnold Thackray and Harriet Zuckerman. New York: Wiley, 1978, 179-207.

Garvey, William D. *Communication: The Essence of Science*. Oxford: Pergamon Press, 1979.

Ginsparg, Paul. "First Steps Toward Electronic Research Communication," *Computers in Physics* 8 (1994): 390-6.

Girifalco, Louis A. *Dynamics of Technological Change.* New York: Van Nostrand Reinhold, 1991.

Glashow, Sheldon. *The Charm of Physics*. New York: American Institute of Physics, 1991.

Gleick, James. *Chaos: Making a New Science*. New York: Viking, 1987.

Goffman, William. "Mathematical Approach to the Spread of Scientific Ideas: The History of Mast Cell Research." *Nature* 212 (1966): 449-52.

________. "An Application of Epidemic Theory to the Growth of Science Symbolic Logic from Boole to Godel." *Progress of Cybernetics* 3 (1970): 971-84.

________. "A Mathematical Method for Analyzing the Growth of a Scientific Discipline." *Journal of the Association for Computing Machinery* 18 (1971): 173-85.

________ and Vaun A. Newill. "Generalization of Epidemic Theory: An Application to the Transmission of Ideas." *Nature* 204 (1964): 225-8.

Granovetter, Mark S. "The Strength of Weak Ties," *American Journal of Sociology* 78 (1973): 1360-80.

Green, Michael B. and John H. Schwarz. "Anomaly Cancellations in Supersymmetric d=10 Gauge-Theory and Superstring Theory." *Physics Letters B* 149 (1984): 117-22.

Griffith, Belver C. and Nicholas C. Mullins. "Coherent Social Groups in Scientific Change." *Science* 177 (1972): 959-64.

________, Henry Small, Judith A. Stonehill, and Sandra Dey. "The Structure of Scientific Literatures II: Toward a Macro- and Microstructure for Science." *Science Studies* 4 (1974): 339-65.

Gross, P.L.K. and E.M. Gross. "College Libraries and Chemical Education." *Science* 66 (1927): 385-9.

Haitun, S.D. "Stationary Scientometric Distributions, Part II: Non-Gaussian Nature of Scientific Activities." *Scientometrics* 4 (1982): 89-94.

Hall, D.H. "The Interface Between Geoscience and Industry: A Case Study of the Interaction Between Research and the Discovery and Mining of Ores for Nuclear Fuels." *Scientometrics* 11 (1987): 199-216.

________. "Rate of Growth of Literature in Geoscience from Computerized Databases." *Scientometrics* 17 (1989): 15-38.

________. "The Science-industry Interface: Correlation of Time Series of Indicators and their Spectra, and Growth Models in the Nuclear Fuel Industry." *Scientometrics* 24 (1992): 237-80.

Hall, Stephen S. "How Technique is Changing Science," *Science* 257 (17 July 1992): 344-9.

Hammond, Michael. "Finite Human Capacities and the Pattern of Social Stratification in a Knowledge Society," In *The Knowledge Society*. Edited by G. Bohme and N. Stehr (Dordrecht: Reidel, 1986), 31-50.

Hawkins, Donald T. "Unconventional Uses of On-line Information Retrieval Systems: On-line Bibliometric Studies." *Journal of the American Society for Information Science* 28 (1977): 13-8.

________. "The Literature of Noble Gas Compounds." *Journal of Chemical Information and Computer Science* 18 (1978): 190-9.

Hicks, Diana and Dave Crouch. "Can Bibliometrics Measure Up?" *Physics World* 3 (September 1990): 27-8.

Holton, Gerald. "Models for Understanding the Growth of Research." *Daedalus* (1962): 94-131.

________. *Thematic Origins of Scientific Thought: Kepler to Einstein.* Cambridge, MA: Harvard University Press, 1973.

Holub, Hans Werner et al. "The Iron Law of Important Articles," *Southern Economic Journal* 58 (1991-92): 317-28.

Horgan, John. "COBE's Cosmic Cartographer," *Scientific American* 267 (July 1992): 34-41.

Huberman, Bernardo. *Social Intermittency*, Preprint, Xerox Corporation, 1993.

Hulme, E. Wyndham. *Statistical Bibliography in Relation to the Growth of Modern Civilization.* London: Butler & Tanner, 1923.

Hurt, C.D. and J.M. Budd. "Modelling the Literature of Superstring Theory: A Case of Fast Literature." *Scientometrics* 24 (1992): 471-80.

Institution of Electrical Engineers. *INSPEC User Manual.* London: Institution of Electrical Engineers, 1983.

________. *INSPEC List of Journals and Other Serial Sources*, 1991/2. London: Institution of Electrical Engineers, 1991.

Keenan, Stella and Pauline Atherton. *The Journal Literature of Physics.* New York: American Institute of Physics, 1961.

________ and F.G. Brickwedde. *Journal Literature Covered by Physics Abstracts in 1965*. New York: American Institute of Physics, 1968.

Kermack, W.O. and A.G. McKendrick, "A Contribution to the Mathematical Theory of Epidemics." *Proceedings of the Royal Society A* 115 (1927): 700-21.

King, Donald W., Dennis D. McDonald, and Nancy K. Roederer. *Scientific Journals in the United States: Their Production, Use and Economics*. Stroudsburg, PA: Hutchinson Ross Publication Co., 1981.

Knorr-Cetina, Karin D. *Science Observed: Perspectives on the Social Study of Science*. London: Sage Publications, 1983.

Kochen, Manfred. *Integrative Mechanisms in Literature Growth*. Westport, CT: Greenwood Press, 1974.

Krassa, Michael A. "Social Groups, Selective Perception, and Behavioral Contagion in Public Opinion," *Social Networks* 10 (1988): 109-36.

Kuhn, Thomas S. *The Structure of Scientific Revolutions*. 2nd ed. Chicago: University of Chicago Press, 1970.

________. "Second Thoughts on Paradigms." In *The Structure of Scientific Theories*. Edited by Frederick Suppe. Urbana, IL: University of Illinois Press, 1974, 589-612.

________ "The Road Since Structure," *PSA 1990* 2 (1990): 3-13.

Kunz, M. "Time Spectra of Patent Information." *Scientometrics* 11 (1987): 163-73.

Lancaster, F.W. and Ja-Lih Lee. "Bibliometric Techniques Applied to Issues Management: A Case Study." *Journal of the American Society for Information Science* 36 (1985): 389-97.

Landsberg, Peter T. "Problems of Explanation in Physics," In *Beyond Belief: Randomness, Prediction and Explanation in Science*. Edited by John L. Casti and A. Karlqvist. Boca Raton, FL: CRC Press, 1991, 55-64.

Lawani, Stephen M. "Citation Analysis and the Quality of Scientific Productivity." *BioScience* 27 (1977): 26-31.

________. "On the Relationship Between Quantity and Quality of a Country's Research Productivity." *Journal of Information Science* 5 (1982): 143-5.

Lazarsfeld, Paul. "The Sociology of Empirical Social Research," *American Sociological Review* 27 (1962): 757-67.

Le Coadic, Yves F. "Information Systems and the Spread of Scientific Ideas." *R&D Management* 4 (1974): 97-111.

________. "Modelling the Communication, Distribution, Transmission or Transfer of Scientific Information." *Journal of Information Science* 13 (1987): 143-8.

Lemaine, Gerard, Roy Macleod, Michael Mulkay and Peter Weingart, eds. *Perspectives on the Emergence of Scientific Disciplines*. Paris: Mouton, 1976.

Lewenstein, Bruce V. "Cold Fusion and Hot History," *Osiris* 7 (1992): 135-63.

Lewenstein, Bruce V. "From Fax to Facts: Communication in the Cold Fusion Saga," *Social Studies of Science* 25 (1995): 403-36.

Leydesdorff, L. et al. "Tracking Areas of Strategic Importance Using Scientometric Journal Mappings," *Research Policy* 23 (1994): 217-29.

Lievrouw, Leah A. "Triangulation as a Research Strategy for Identifying Invisible Colleges Among Biomedical Scientists." *Social Networks* 9 (1987): 217-48.

Lievrouw, Leah A. "Communication, Representation, and Scientific Knowledge: A Conceptual Framework and Case Study," *Knowledge and Policy: The International Journal of Knowledge Transfer and Utilization* 5 (1992): 6-28.

Line, Maurice B. "The Half-life of Periodical Literature - Apparent and Real Obsolescence." *Journal of Documentation* 26 (1970): 46-52.

Lotka, Alfred J. "Contributions to the Analysis of Malaria Epidemiology." *American Journal of Hygiene* (3, Suppl.1, 1923): 1-121.

Luukkonen, T. et al. "The Measurement of International Scientific Collaboration," *Scientometrics* 28 (1993): 15-36.

McAllister, Paul R. and Francis Narin, "Characterization of the Research Papers of U.S. Medical Schools." *Journal of the American Society for Information Science* 34 (1983): 123-31.

Machlup, Fritz. *The Production and Distribution of Knowledge in The United States.* Princeton, NJ: Princeton University Press, 1962.

Magyar, George. "How Trustworthy are Experimental Facts?" *European Journal of Physics* 2 (1981): 244-9.

Mahajan, Vijay and Yoram Wind. *Innovation Diffusion Models of New Product Acceptance.* Cambridge, MA: Ballinger Publishing Company, 1986.

Martin, Ben R. and John Irvine. "Assessing Basic Research: Some Partial Indicators of Scientific Progress in Radio Astronomy." *Research Policy* 12 (1983): 61-90.

Martin, Ben R. and John Irvine. *Research Foresight: Priority-Setting in Science.* London: Pinter Publishers, 1989.

Martin, R. Douglas, Alexander Samarov, and Walter Vandaele. "Robust Methods for ARIMA Models." In *Applied Time Series Analysis of Economic Data: Proceedings of a Conference*. Arlington, VA: U.S. Dept. of Commerce, 1983, 153-69.

May, K.O. "Quantitative Growth of the Mathematical Literature." *Science* 154 (1966): 1672-3.

Meade, Nigel. "Forecasting Using Growth Curves - An Adaptive Approach," *Journal of the Operational Research Society* 36 (1985): 1103-15.

Meadows, A.J. *Communication in Science*. London: Butterworths, 1974.

Meadows, A.J. and J.G. O'Connor. "Bibliographical Statistics as a Guide to Growth Points in Science." *Science Studies* 1 (1971): 95-9.

Menard, Henry W. *Science: Growth and Change*. Cambridge, MA: Harvard University Press, 1971.

Menzel, Herbert and Elihu Katz. "Social Relations and Innovation in the Medical Profession: The Epidemiology of a New Drug." *Public Opinion Quarterly* 19 (1955): 337-52.

Merton, Robert K. *The Sociology of Science*. Chicago: University of Chicago Press, 1973.

Miettinen, Ole S. *Theoretical Epidemiology*. Albany, NY: Delmar, 1985.

Moed, H.F. "The Use of Online Databases for Bibliometric Analysis." In *Informetrics 87/88: Select Proceedings of the First International Conference on Bibliometrics and Theoretical Aspects of Information Retrieval, Diepenbeek, Belgium, 25-28 August 1987*. Edited by Leo Egghe and Ronald Rousseau. Amsterdam: Elsevier, 1988, 133-46.

Moed, H. F. "Bibliometric Measurement of Research Performance and Price's Theory of Differences Among Sciences," *Scientometrics* 15 (1989): 473–83.

________, W.J.M. Burger, J.G. Frankfort, and A.F.J. Van Raan. "The Application of Bibliometric Indicators: Important Field- and Time-dependent Factors to be Considered." *Scientometrics* 8 (1985): 177-203.

Moravcsik, Michael J. "Measures of Scientific Growth." *Research Policy* 2 (1973): 266-75.

________. "Phenomenology and Models of the Growth of Science." *Research Policy* 4 (1975): 80-6.

________ and P. Murugesan. "Citation Patterns in Scientific Revolutions." *Scientometrics* 1 (1979): 161-9.

Moskovits, M. *Metal Clusters*. New York: Wiley, 1986.

Mulkay, Michael J., G.N. Gilbert and Steve Woolgar. "Problem Areas and Research Networks in Science." *Sociology* 9 (1975): 187-203.

Mullins, Nicholas C. "Model for the Development of Sociological Theories." In *Theories and Theory Groups in Contemporary American Sociology*. Edited by N.C. Mullins. New York: Harper & Row, 1973, 17-35.

Nadel, Edward. "Commitment and Co-citation: An Indicator of Incommensurability in Patterns of Formal Communication." *Social Studies of Science* 13 (1983): 255-82.

Narin, Francis. *Evaluative Bibliometrics: The Use of Publication and Citation Analysis in the Evaluation of Scientific Activity*. Philadelphia: Computer Horizons Inc., 1976.

________ and Mark P. Carpenter, "National Publication and Citation Comparisons." *Journal of the American Society for Information Science* 26 (1975): 80-93.

National Board of Employment, Education and Training. *Quantitative Indicators of Australian Research*. Canberra: Australian Government Publishing Service, 1994.

National Research Council. Panel on the Future of Atomic, Molecular, and Optical Sciences. *Atomic, Molecular, and Optical Science*. Washington, DC: National Academy Press, 1994.

A. Neelameghan, "Expressions of Time in Information Science and Their Implications: An Overview," *Annals of Library Science and Documentation* 24 (1977): 13-33.

Nelkin, Dorothy. *Selling Science: How the Press Covers Science and Technology*. New York: Freeman, 1987.

Newell, Sue and Jacky Swan. "Professional Associations as Important Mediators of the Innovation Process," *Science Communication* 16 (1995): 371-87.

Oliver, M.R. "The Effect of Growth on the Obsolescence of Semiconductor Physics Literature." *Journal of Documentation* 27 (1971): 11-17.

Oliver, Pamela E. and Gerald Marwell. "The Paradox of Group Size in Collective Action: A Theory of the Critical Mass. II," *American Sociological Review* 53 (1988): 1-8.

Oseman, Robert H. "The Growth and Value of Conference Literature," In *New Horizons for the Information Profession*. Edited by H. Dyer and G. Tseng. (London, Taylor and Graham, 1988), 124-37.

Persson, O. "Measuring Scientific Output by Online Techniques." In *Handbook of Quantitative Studies in Science and Technology*. Edited by A.F.J. Van Raan. Amsterdam: North-Holland, 1988, 229-52.

Peters, H.P.F. et al. "Cognitive Resemblance and Citation Relations in Chemical Engineering Publications," *Journal of the American Society for Information Science* 46 (1995): 9-21.

Pierce, John R. "Research and Surprise." In *International Symposium on International Cooperation and Competition in Science and Technology, Tokyo, 12-13 April, 1988*. Tokyo: Engineering Academy of Japan, 1988, 1-5.

Phillips, D.P. et al., "Importance of the Lay Press in the Transmission of Medical Knowledge to the Scientific Community," *New England Journal of Medicine* 325 (October 17, 1991): 1180-3.

Pierce, Sidney J. "Disciplinary Work and Interdisciplinary Areas: Sociology and Bibliometrics." In *Scholarly Communication and Bibliometrics*. Edited by Christine L. Borgman. Newbury Park, CA: Sage, 1990, 47-58.

Polanyi, Michael. "The Republic of Science: Its Political and Economic Theory." *Minerva* 1 (1962): 54-73.

Popper, Karl R. *Conjectures and Refutations: The Growth of Scientific Knowledge*. New York: Harper & Row, 1965.

Price, Derek J. DeSolla. "Quantitative Measures of the Development of Science." *Archives internationales d'histoire des sciences* 14 (1951): 85-93.

________. *Science since Babylon*. New Haven, CT: Yale University Press, 1961.

________. *Little Science Big Science*. New York: Columbia University Press, 1963.

________. "Networks of Scientific Papers." *Science* 149 (1965): 510-5.

________. "Measuring the Size of Science." *Proceedings of the Israel Academy of Sciences and Humanities* 4 (1969): 98-111.

________. Ups and Downs in the Pulse of Science and Technology. *Sociological Inquiry* 48 (1978): 162-71.

Rebrova, M.P. and V.V. Komarov, "Some Aspects of the Scientometric Analysis of the Development of Research in the Area of Superconductivity." *Nauchno-Tekhnicheskaya Informatsiya, Ser. 1* 16 (8, 1989): 23-7.

Rescher, Nicholas. *Scientific Progress*. Pittsburgh, PA: University of Pittsburgh Press, 1978.

Rogers, Everett M. *Diffusion of Innovations*. 3rd ed. New York: Free Press, 1983.

Rosenberg, N. *Inside the Black Box: Technology and Economics*. Cambridge: Cambridge University Press, 1982.

Ross, Ronald. "Some a Priori Pathometric Equations." *British Medical Journal* (March 27 1915): 546-7.

Schwarz, John H. "Superstring Theory." *Physics Reports* 89 (1982): 223-322.

Schweber, Silvan S. "Physics, Community and the Crisis in Physical Theory," *Physics Today* 46 (1993): 34-40.

Sichel, S.H. "The GIGP distribution Model with Applications to Physics Literature." *Czechoslovak Journal of Physics B* 36 (1986), 133-7.

Siegfried, A. *Germs and Ideas*. London: Oliver and Boyd, 1965.

Small, Henry. "Macro-level Changes in the Structure of Co-Citation Clusters: 1983-1989." *Scientometrics* 26 (1993): 5-20.

________ and Diana Crane. "Specialties and Disciplines in Science and Social Science: An Examination of Their Structure Using Citation Indexes," *Scientometrics* 1 (1979): 445-61.

________ and Eugene Garfield. "The Geography of Science: Disciplinary and National Mappings." *Journal of Information Science* 11 (1985): 147-59.

________ and E. Greenlee. "Collagen Research in the 1970s." *Scientometrics* 10 (1986): 95-117.

________ and David Pendlebury. "Introducing Science Watch: A New Kind of Science Intelligence." *Science Watch* 1 (January 1990): 2.

Smith, Linda. "Citation Analysis." *Library Trends* 30 (Summer 1981): 83-106.

Stahl, Michael J. and Joseph A. Steger. "Measuring Innovation and Productivity - a Peer Rating Approach." *Research Management* 20 (1977): 35-8.

Stephan, Paula E. and Sharon G. Levin, "Inequality in Scientific Performance: Adjustment for Attribution and Journal Impact." *Social Studies of Science* 21 (1991): 351-68.

Sterman, John D. "The Growth of Knowledge: Testing a Theory of Scientific Revolutions with a Formal Model," *Technological Forecasting and Social Change* 28 (1985): 93-122;

Stewart, John A. *Drifting Continents & Colliding Paradigms: Perspectives on the Geoscience Revolution*. Bloomington: Indiana University Press, 1990.

Strang, David. "Adding Social Structure to Diffusion Models: An Event History Framework." *Sociological Methods & Research* 19 (1991): 324-53.

Sullivan, Daniel, D. Koester, D.H. White, and R. Kern. "Understanding Rapid Theoretical Change in Particle Physics: A Month-by-month Co-citation Analysis." In *Information Choices and Policies: Proceedings of the 16th ASIS Annual Meeting, Minneaolis, Minnesota, 14-18 October 1979*. New York, Learned Information, 1979, 276-85.

Sullivan, Thomas J. "The 'Critical Mass. in Crowd Behavior: Crowd Size, Contagion and the Evolution of Riots," *Humboldt Journal of Social Relations* 4 (Spring/Summer 1977): 45-59.

Tabah, Albert N. *Growth Patterns Following Sudden Discovery: The Case of Superconductivity Literature*. Montreal: McGill University, 1991.

________ "Nonlinear Dynamics and the Growth of Literature." *Information Processing and Management* 28 (1992): 61-73.

Tague, Jean, Jamshid Beheshti, and Lorna K. Rees-Potter. "The Law of Exponential Growth: Evidence, Implications and Forecasts." *Library Trends* 30 (1981): 125-50.

Taubes, Gary. *Nobel Dreams: Power, Deceit and the Ultimate Experiment.* New York: Random House, 1986.

Thackray, Arnold. *Contemporary Classics in Physical, Chemical, and Earth Sciences.* Philadelphia: ISI Press, 1986.

Thome, Helmut. "A Box-Jenkins Approach to Modeling Outliers in Time Series Analysis." *Sociological Methods & Research* 23 (1995): 442-78.

Todorov, Radosvet. "Distribution of Physics Literature." *Scientometrics* 7 (1985): 195-209.

________ "Co-Classification Analysis for Science Mapping: An Example from Superconductivity." In *Science and Technology Indicators.* Edited by A.F.J. Van Raan. Leiden: DSWO Press, 1988, 261-70.

Trofimenko, A.P. "Scientometric Analysis of the Development of Nuclear Physics during the Last 50 Years." *Scientometrics* 11 (1987): 231-50.

________ "Scientometric Analysis of the Topical Content of Scientific Research and Its Particularities," *Scientometrics* 18 (1990): 409-35.

U.S. Congress. Office of Technology Assessment. *High-temperature Superconductivity in Perspective.* Washington, DC: Government Printing Office, 1990.

Van Raan, A.F.J. "Evaluation of Research Groups." In *The Evaluation of Scientific Research.* Edited by D. Evered. NY: Wiley, 1989, 169-87.

________ and H.P.F. Peters. "Dynamics of a Scientific Field Analyzed by Co-Subfield Structures." *Scientometrics* 15 (1989): 607-20.

________ and R.J.W. Tijssen, "Numerical Methods for Information on Aspects of Science: Scientometric Analysis and Mapping." *Perspectives in Information Management* 2 (1990): 203-28.

Van Vianen, Ben and Anthony Van Raan, "Knowledge Expansion in Applied Science: A Bibliometric Study of Laser Medicine and Polyimide Chemistry." In *Dynamics of Science- Based Innovation.* Edited by H. Grupp. New York: Springer, 1992, 227-67.

Virgo, Julie "A Statistical Procedure for Evaluating the Importance of Scientific Papers." *Library Quarterly* 47 (1977): 257-67.

Vlachy, Jan. "Publication Output in Physics Subfields." *Czechoslovak Journal of Physics B* 29 (1979): 829-36.

________. "Publication Output of World Physics." *Czechoslovak Journal of Physics B* 29 (1979): 475-80.

________. "Citation Histories of Scientific Publications: The Data Sources." *Scientometrics* 7 (1985): 257-67.

Wagner-Dobler, R. and J. Berg. "Regularity and Irregularity in the Development of Scientific Disciplines: The Case of Mathematical Logic," *Scientometrics* 30 (1994): 303-19.

________ "The Dependence of Lotka's Law on the Selection of Time Periods in the Development of Scientific Areas and Authors," *Journal of Documentation* 51 (1995): 28-43.

Wallace, Danny P. and Susan Bonzi. "The Relationship Between Journal Productivity and Quality" In *ASIS '85: Proceedings of the Forty-Eighth ASIS Annual Meeting, Las Vegas, 20-24 October 1985*. Edited by C.A. Parkhurst. White Plains, NY: Knowledge Industry Publications, 1985, 193-6.

Webb E.T., D.T. Campbell, R.D. Schwartz, L. Sechrest, and J.B. Grove. *Nonreactive Measures in the Social Sciences*. 2nd ed. Boston: Houghton Mifflin Company, 1981.

Wilson, P.W. and E.B. Fred. "The Growth Curve of a Scientific Literature: Nitrogen Fixation by Plants." *Scientific Monthly* 41 (1935): 240-50.

Wittenberg, Jason. "On the Very Idea of a System Dynamics Model of Kuhnian Science," *System Dynamics Review* 8 (1992): 21-33.

Wolfram, D., C.M. Chu and X. Lu, "Growth of Knowledge: Bibliometric Analysis Using Online Database Data." In *Informetrics 89/90: Selection of Papers Submitted for the Second International Conference on Bibliometrics, Scientometrics and Informetrics, London, Ontario, 5-7 July, 1989*. Edited by Leo Egghe and Ronald Rousseau. Amsterdam: Elsevier, 1990, 355-64.

Worthen, Dennis B. "The Epidemic Process and the Contagion Model." *Journal of the American Society for Information Science* 24 (1973): 343-6.

Ziman, John. *Science in a Steady State*. London: Science Policy Support Group, 1977.

Ziman, John. *Prometheus Bound: Science in a Dynamic Steady State*. Cambridge: Cambridge University Press, 1994.

Zuckerman, H.A. "Nobel laureates in Science: Patterns of Productivity, Collaboration, and Authorship." *American Sociological Review* 32 (1967): 391-403.

________ and Joshua Lederberg. "Postmature Scientific Discovery?" *Nature* 324 (1986): 629-631.

________ and R.K. Merton. "Patterns of Evaluation in Science: Institutionalization, Structure and Functions of the Reference System," *Minerva* 9 (1971): 66-100.

Zunde, Pranas. *On Empirical Foundations of Information Science*. Atlanta, GA: School of Information and Computer Science, 1981.

APPENDIX A

Times Series Charts of *Physics Abstracts* Chapters

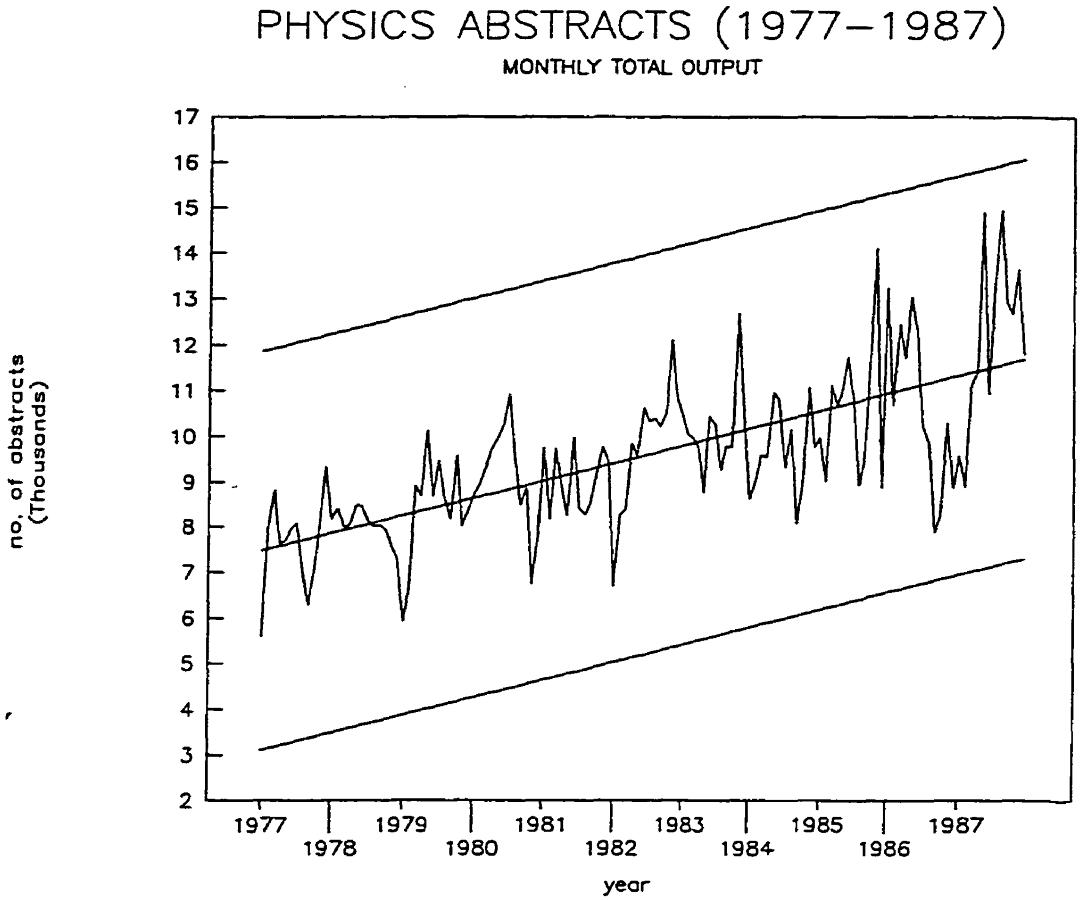


Figure A1. Total number of items published by *Physics Abstracts* per month.

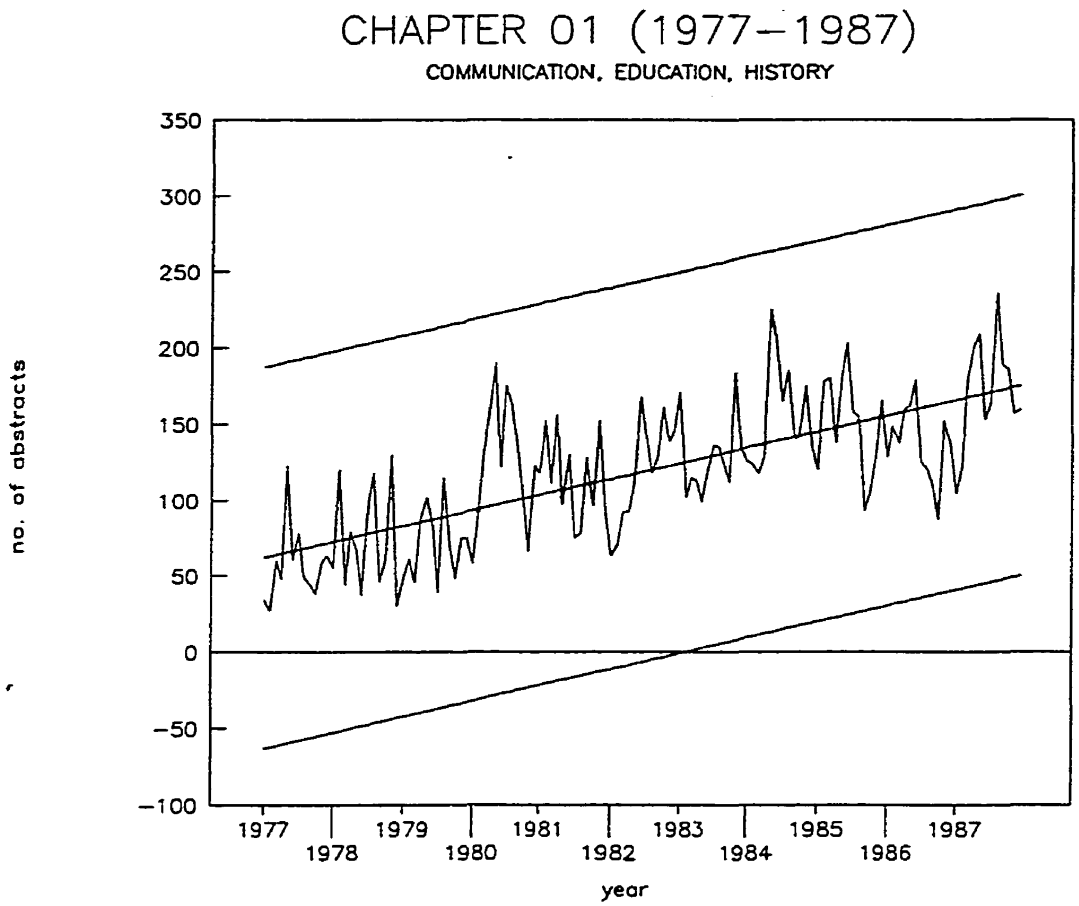


Figure A2. Chapter 1.

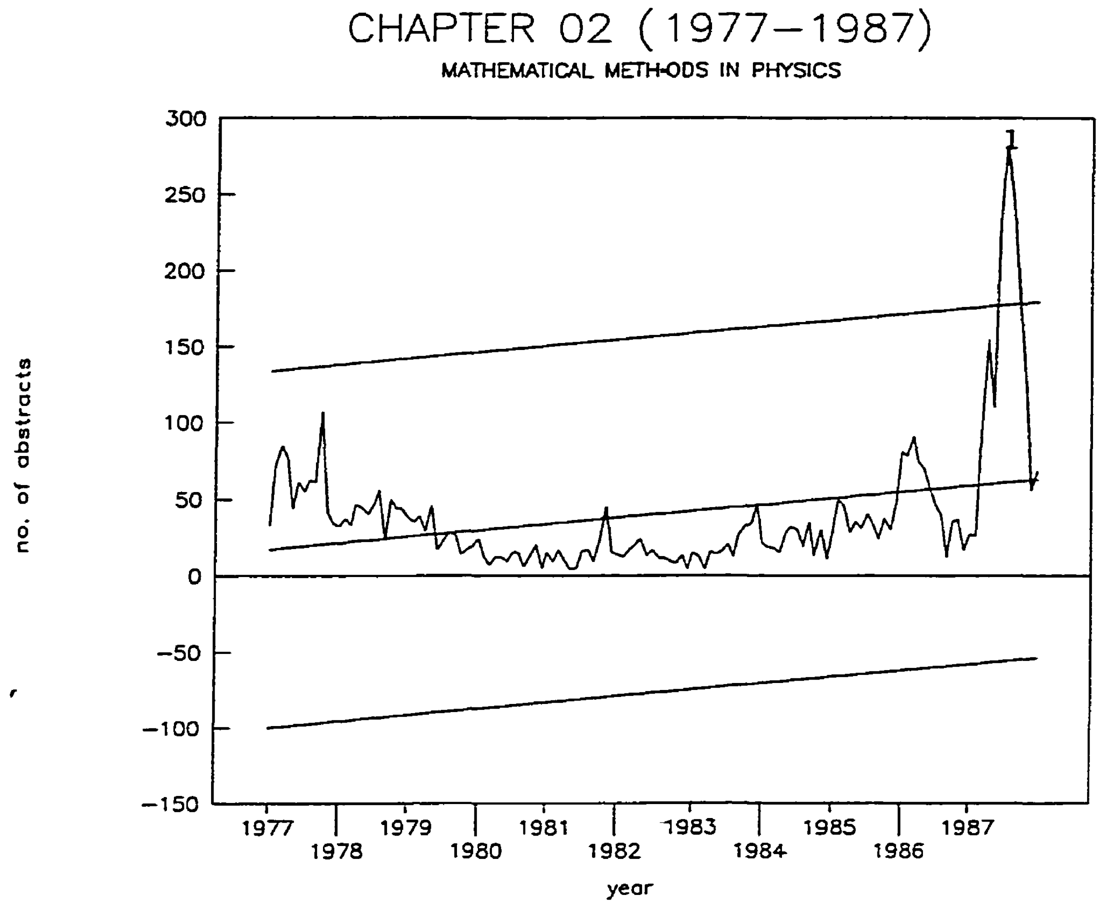


Figure A3. Chapter 2.

Outlier 1 of June-September 1987 contains 942 abstracts composed, for the most, of journal articles. There are no discernible groupings of conference proceedings or special journal issues that account for it.

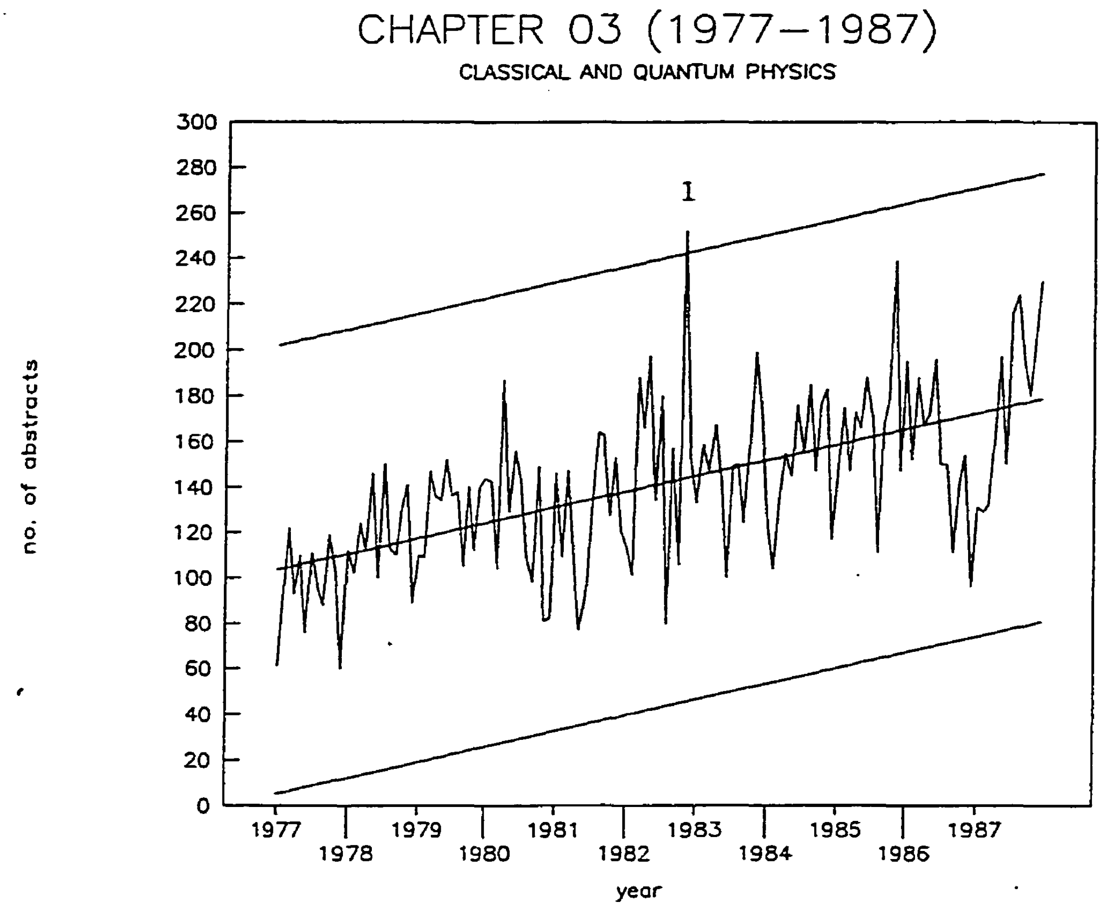


Figure A4. Chapter 3.

Outlier 1 of November 1982 contains 252 abstracts, including *Mathematical Problems in Theoretical Physics*: Proceedings of the 6th International Conference on Mathematical Physics, Berlin, August 1981 (21 notices). Papers from all conferences account for 10.7% of the contents.

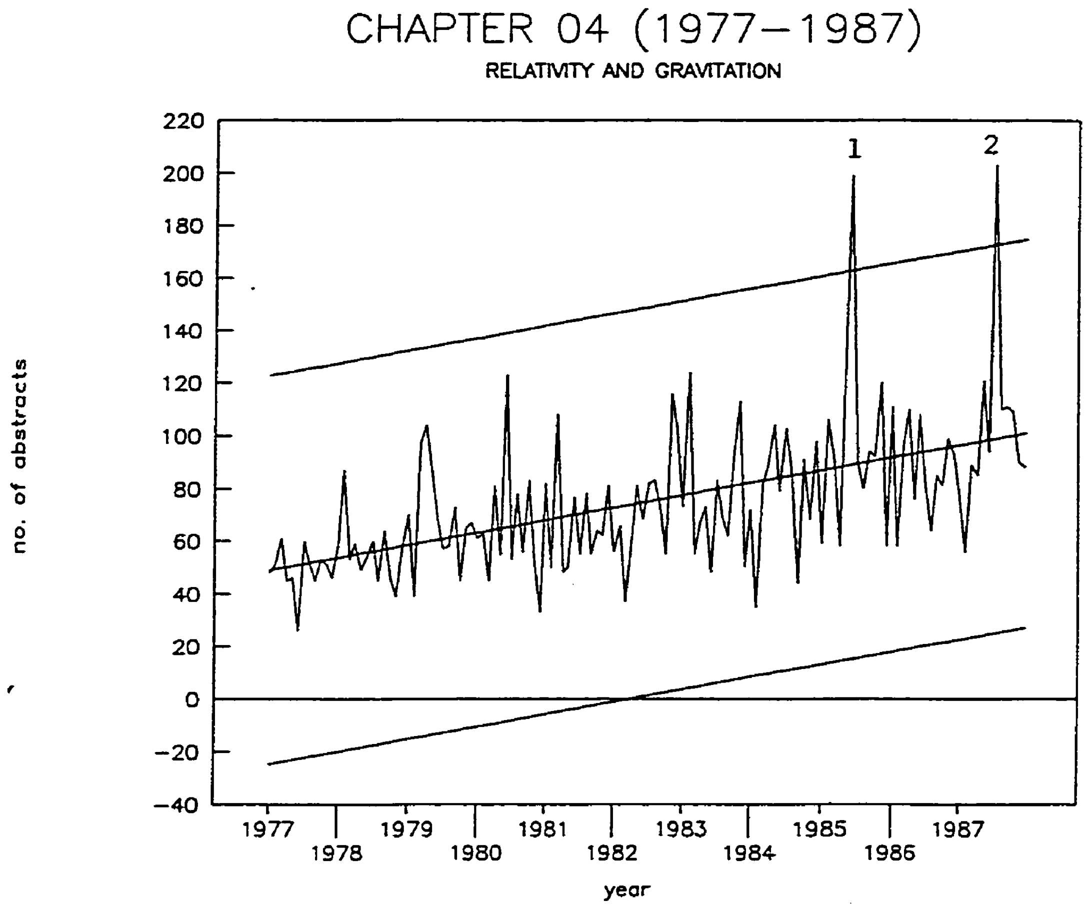


Figure A5. Chapter 4.

Outlier 1 of June 1985 contains 199 abstracts, including the *Marcel Grossman Meeting on General Relativity*, 3rd, Shanghai, September 1982 (103 notices) and the *Workshop on Klauza-Klein Theories*, Chalk River, Ontario, August 1983 (11 notices). Papers from all conferences account for 57.3% of the contents.

Outlier 2 of July 1987 contains 203 abstracts, including the *Marcel Grossman Meeting on General Relativity*, 4th, Rome, 1985 (125 notices). Papers from all conferences account for 61.6% of the contents.

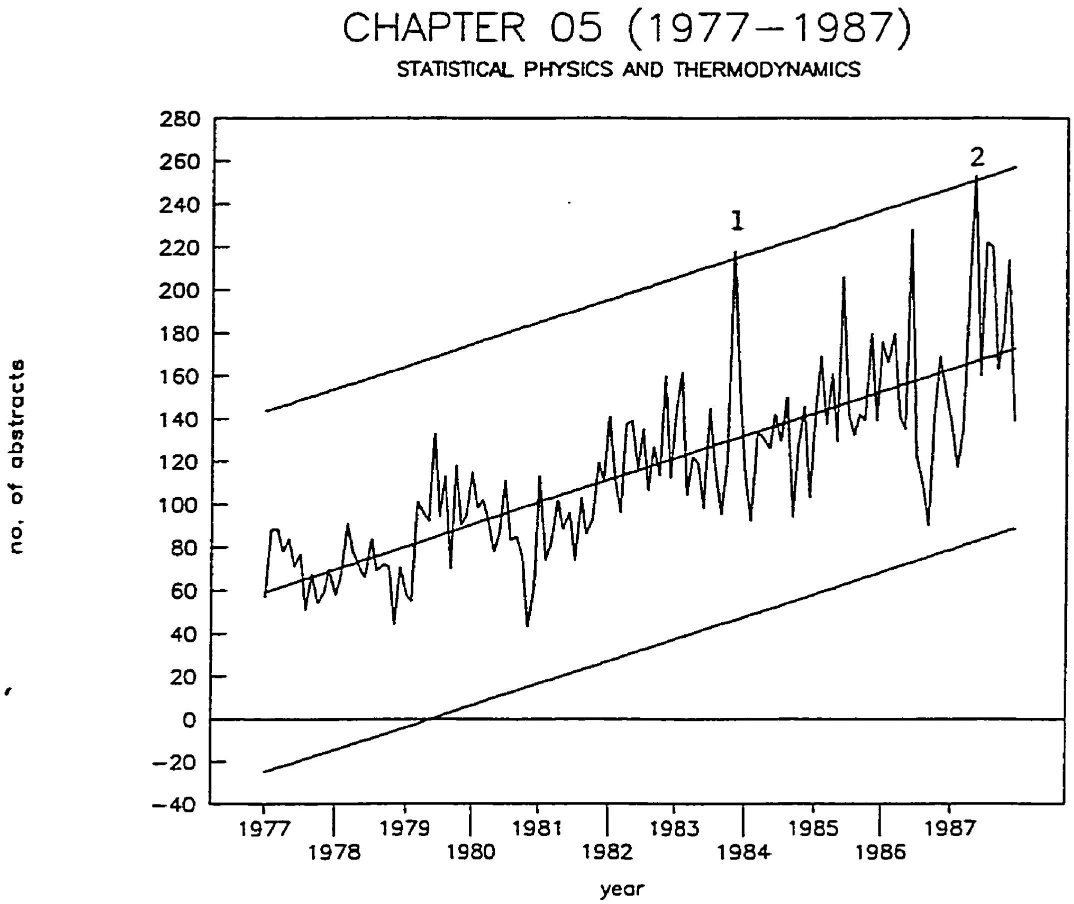


Figure A6. Chapter 5.

Outlier 1 of November 1983 contains 218 abstracts, including *Dynamical Systems and Chaos*: Proceedings of the Sitges Conference on Statistical Mechanics, Barcelona, September 1982 (29 notices). Papers from all conferences account for 13.3% of the contents.

Outlier 2 of May 1987 contains 253 abstracts, including *Statistical Physics and Dynamical Systems: Rigorous Results*, Koszeg, Hungary, September 1984 (15 notices). Papers from all conferences account for 5.9% of the contents.

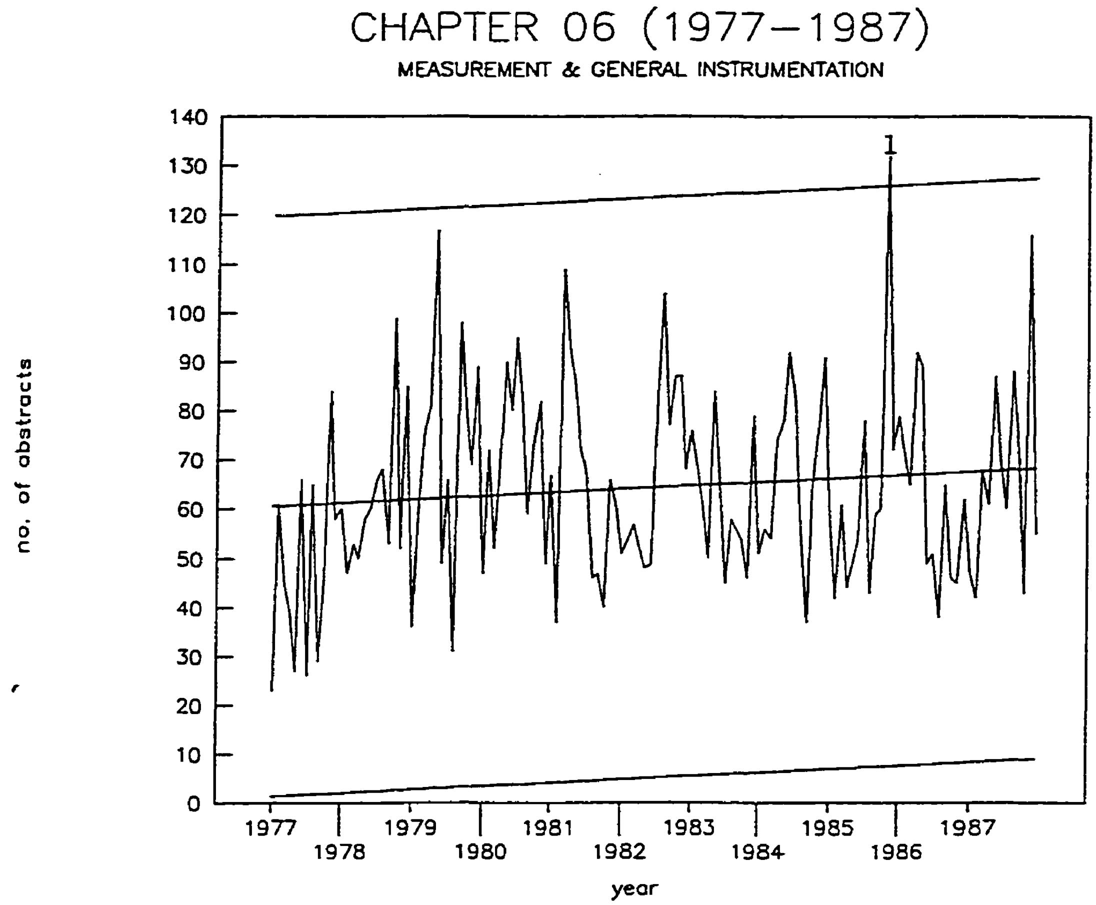


Figure A7. Chapter 6.

Outlier 1 of November 1985 contains 132 abstracts, including the *Annual Frequency Control Symposium*, 38th, Philadelphia, June 1984 (13 notices) and the *International Conference on Precision Measurement and Fundamental Constants*, 2nd, Gaithersburg, MD, June 1981 (41 notices). Papers from all conferences account for 40.9% of the contents.

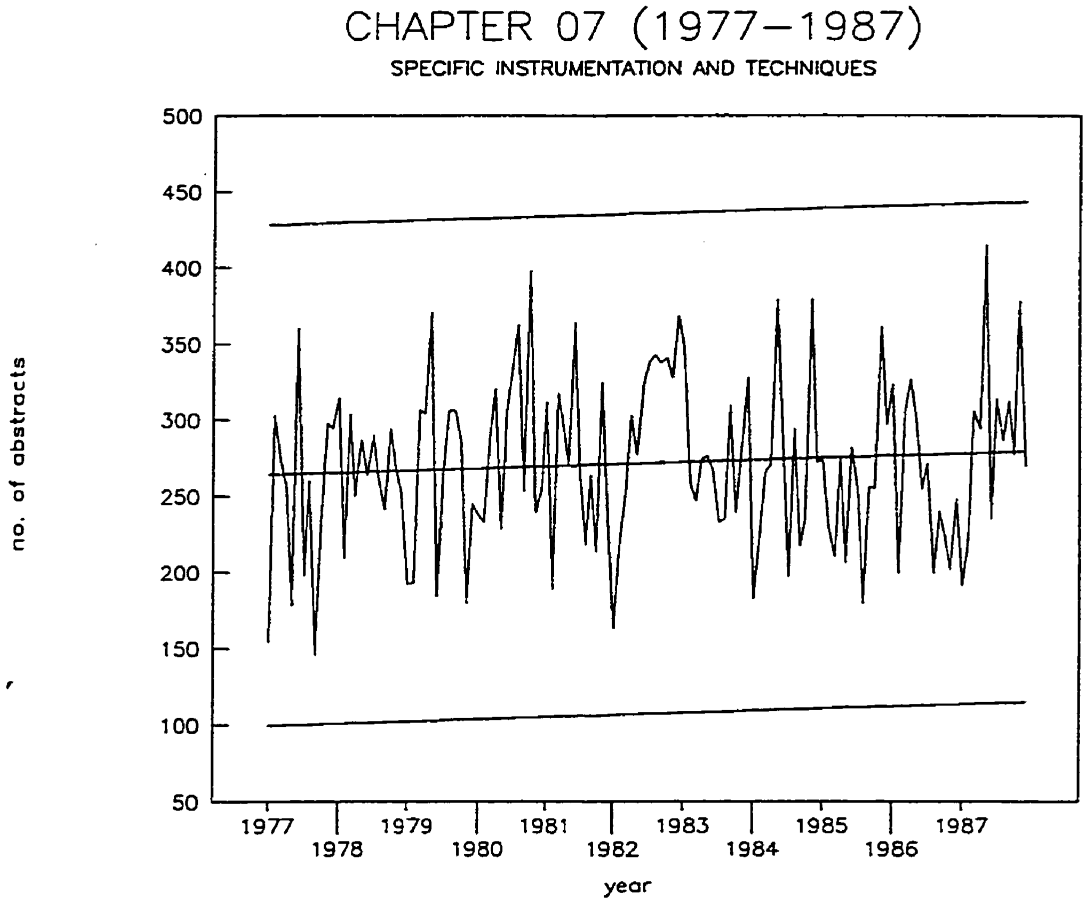


Figure A8. Chapter 7.

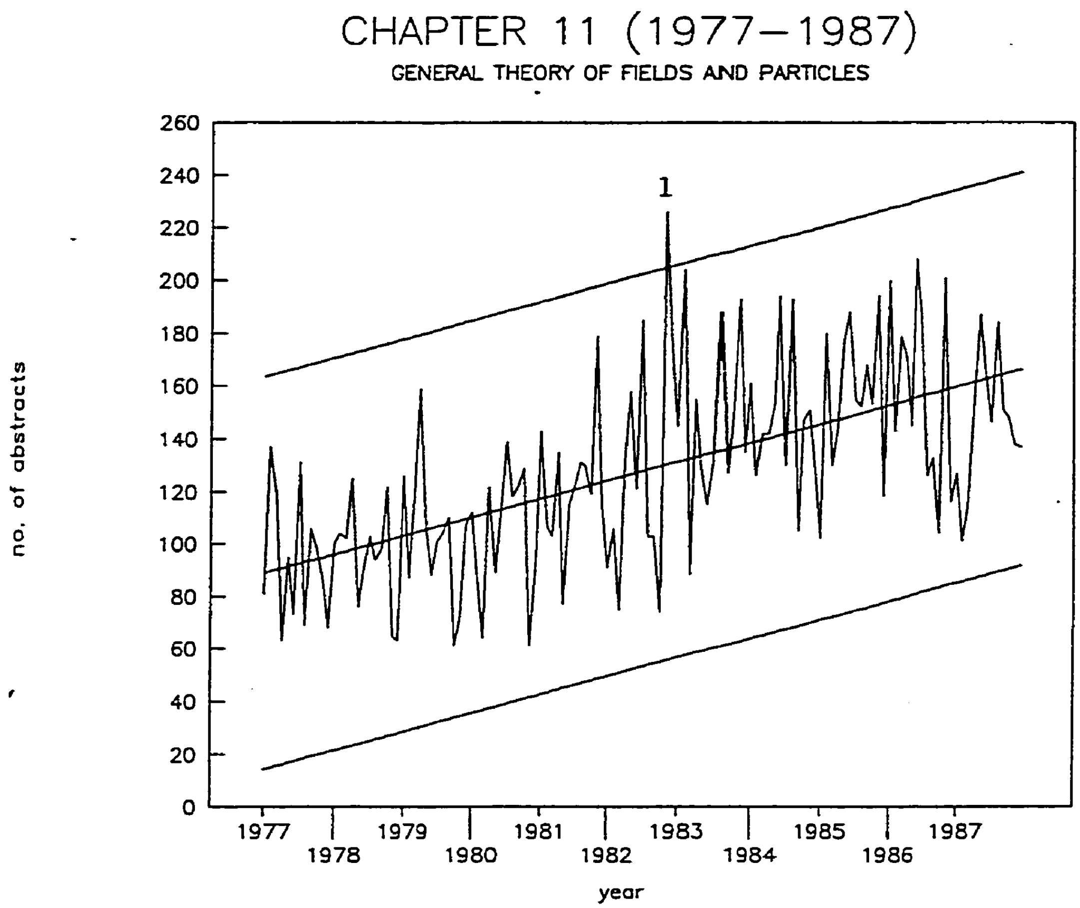


Figure A9. Chapter 11.

Outlier 1 of November 1982 contains 226 abstracts, including *Mathematical Problems in Theoretical Physics: International Conference on Mathematical Physics*, 6th, Berlin, August 1981 (19 notices). Papers from all conferences account for 8.4% of the contents.

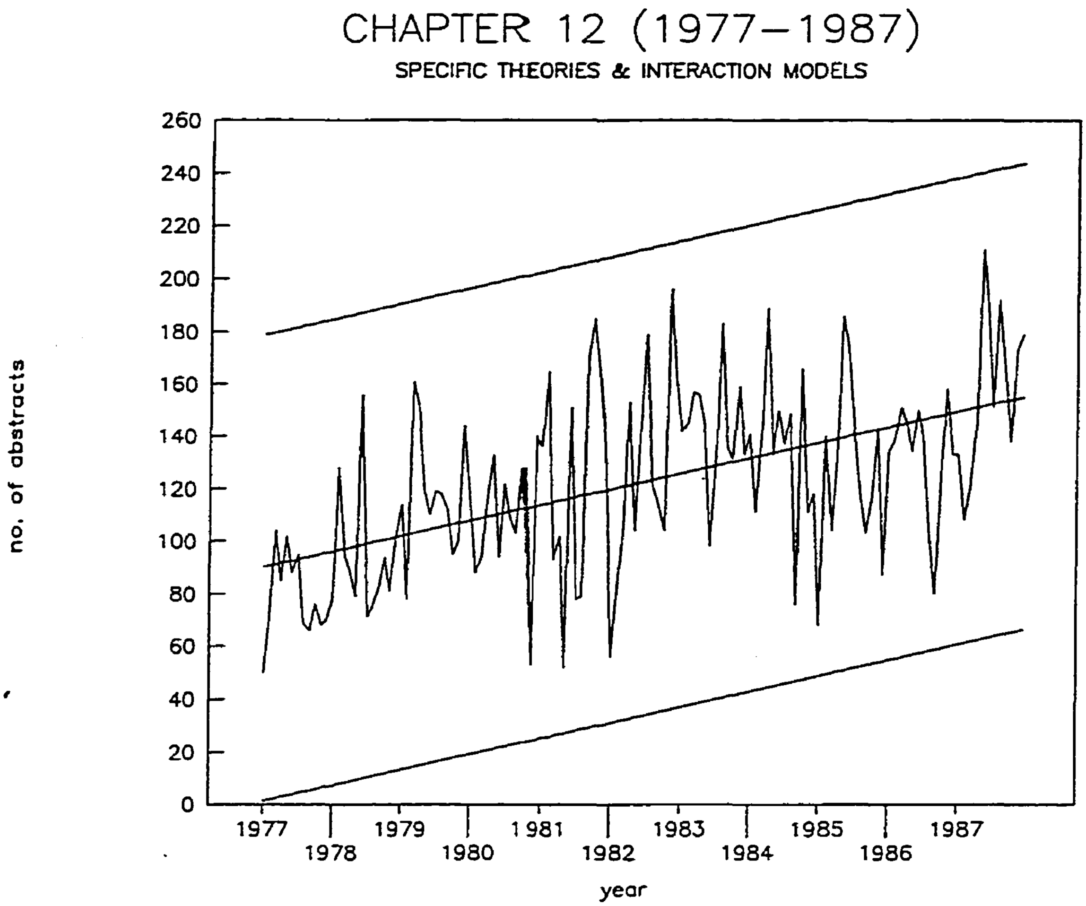


Figure A10. Chapter 12.

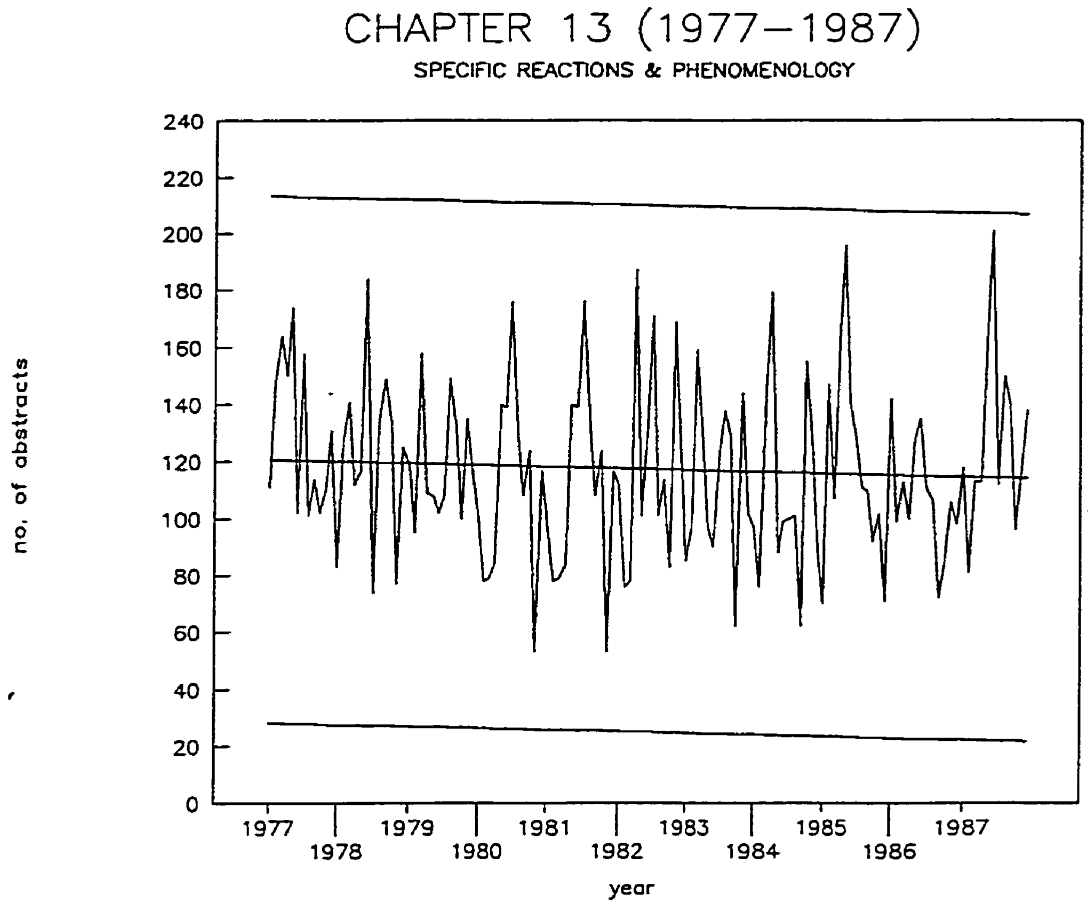


Figure A11. Chapter 13.

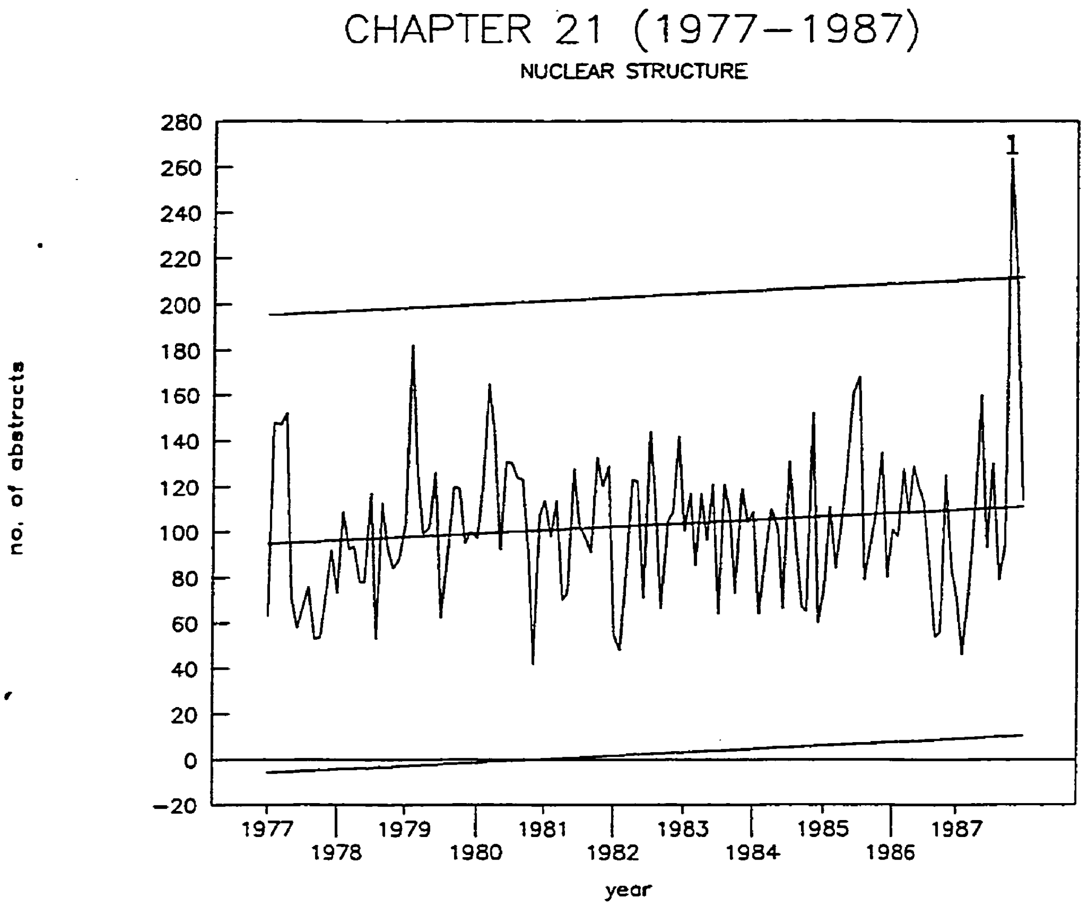


Figure A12. Chapter 21.

Outlier 1 of October 1987 contains 264 abstracts, including the *International Nuclear Physics Conference*, Harrogate, Yorks, August 1986 (141 notices) and the *International Conference on Hyperfine Interactions*, 7th, Bangalore, September 1986 (15 notices). Papers from all conferences account for 59.1% of the contents.

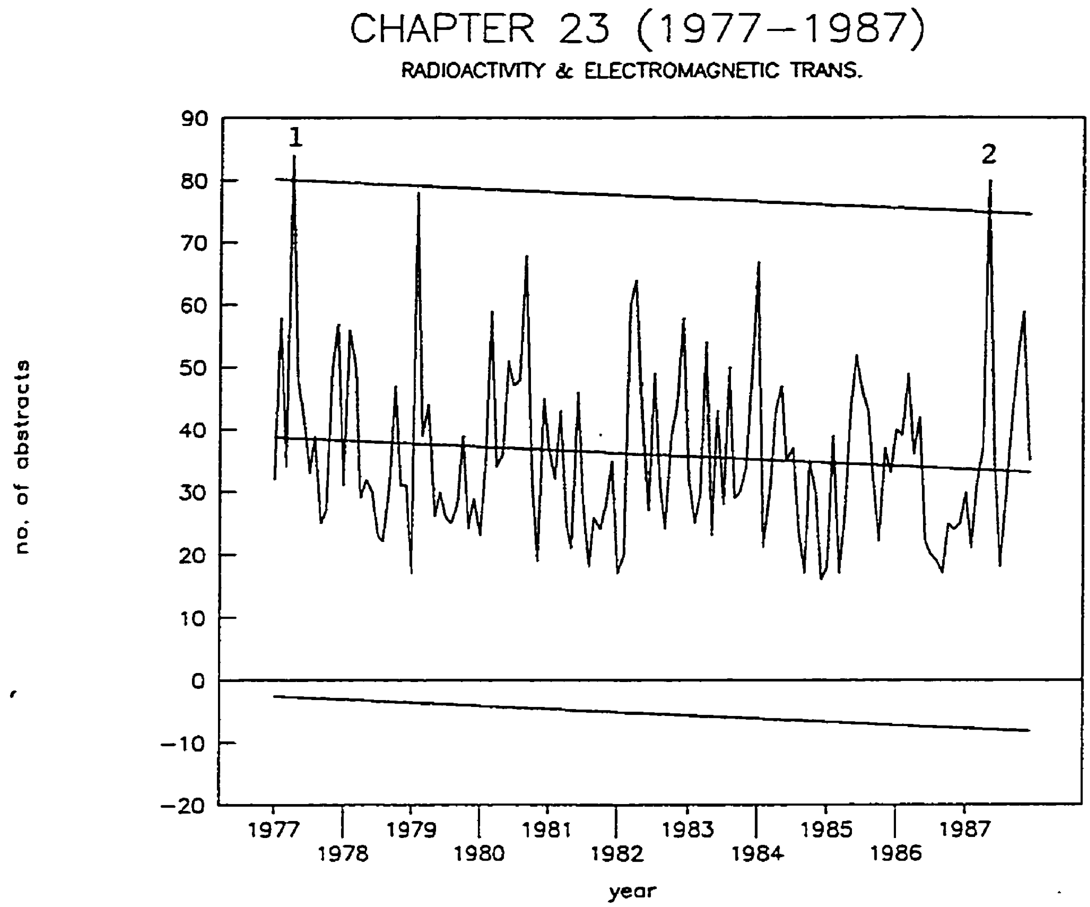


Figure A13. Chapter 23.

Outlier 1 of April 1977 contains 84 abstracts, including the *International Conference on Nuclei Far From Stability*, 3rd, Corsica, 1976 (38 notices). Papers from all conferences account for 45.2% of the contents.

Outlier 2 of May 1987 contains 80 abstracts, including the *International Symposium on Weak and Electromagnetic Interactions in Nuclei*, Heidelberg, July 1986 (38 notices). Papers from all conferences account for 47.5% of the contents.

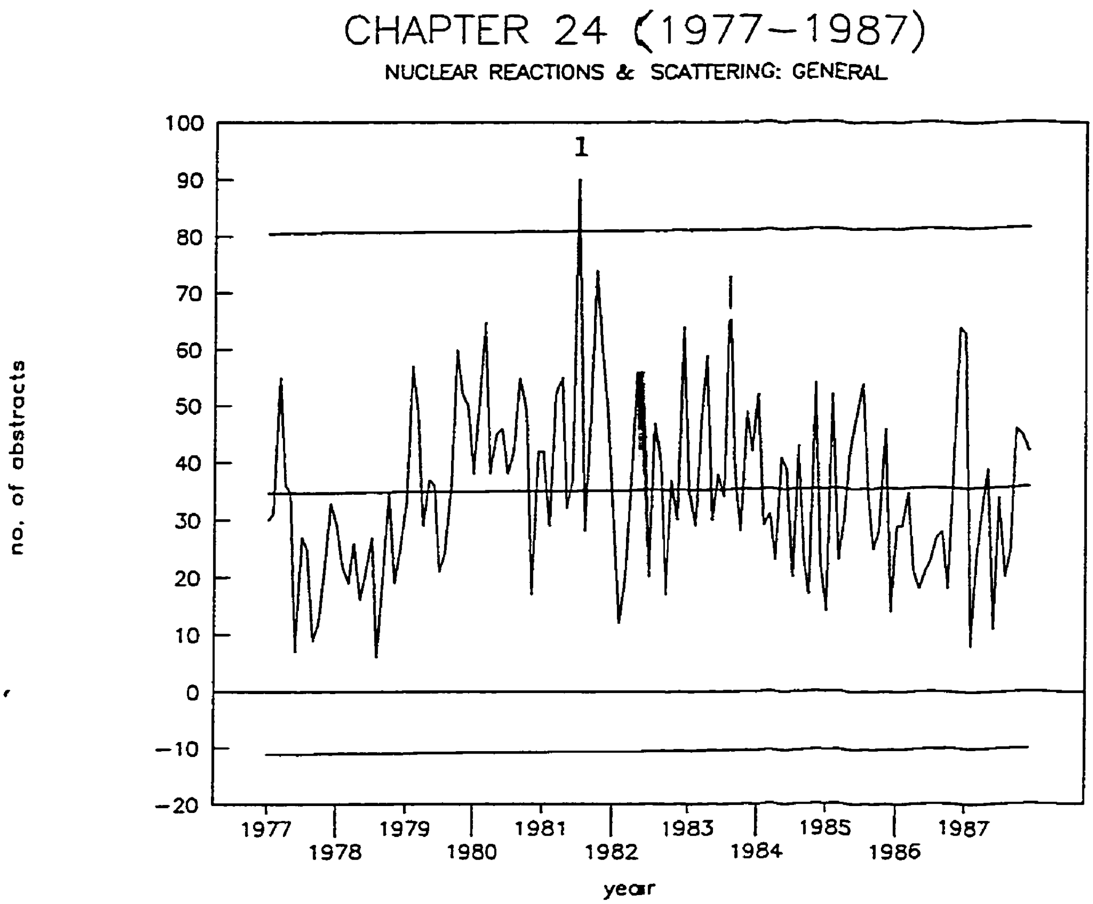


Figure A14. Chapter 24.

Outlier 1 of July 1981 contains 90 abstracts, including the *Topical Conference on Giant Multiple Resonances*, Oak Ridge, October 1979 (48 notices). Papers from all conferences account for 53.3% of the contents.

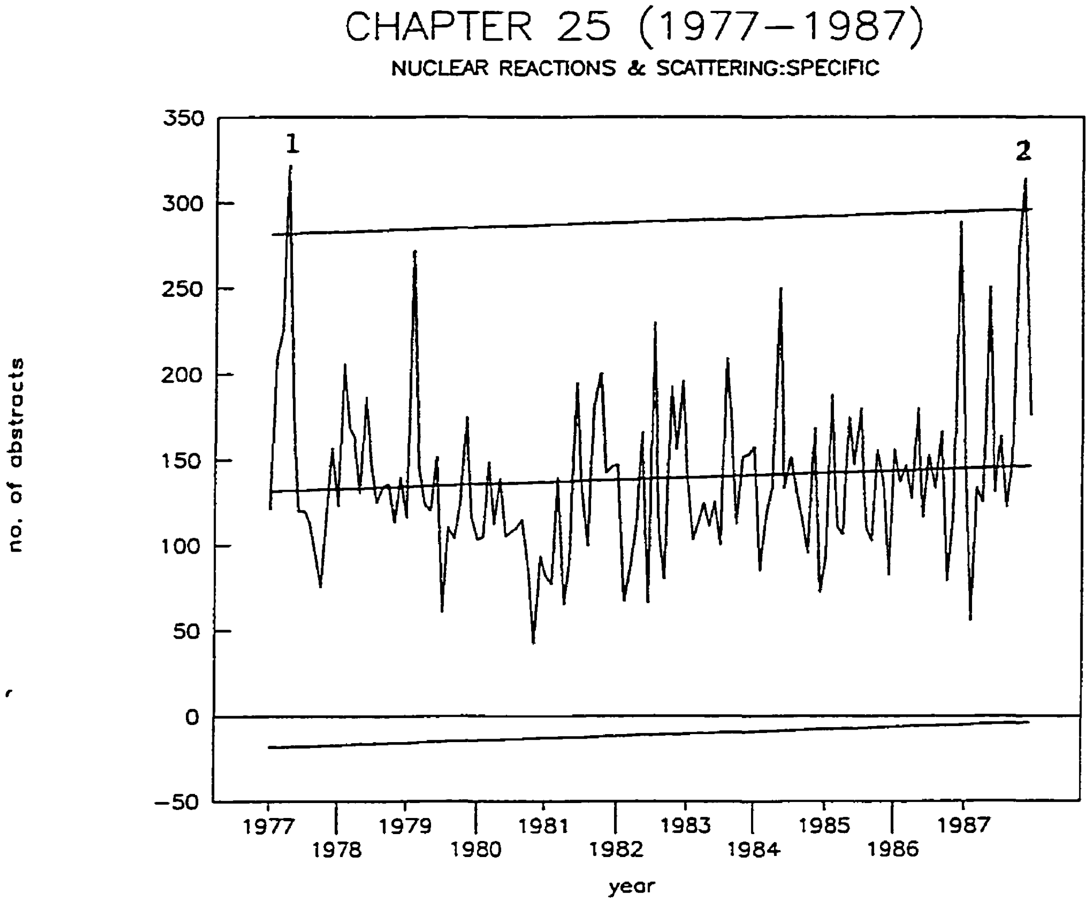


Figure A15. Chapter 25.

Outlier 1 of April 1977 contains 322 abstracts, including the *International Conference on Interactions of Neutrons with Nuclei*, Lowell, MA, 1976 (96 notices) and *Nuclear Cross Sections and Technology*, Washington, DC, 1975 (122 notices). Papers from all conferences account for 67.7% of the contents.

Outlier 2 of November 1987 contains 314 abstracts, including the *International Nuclear Physics Conference*, Harrogate, Yorks, England, August 1986 (98 notices). Papers from all conferences account for 31.2% of the contents.

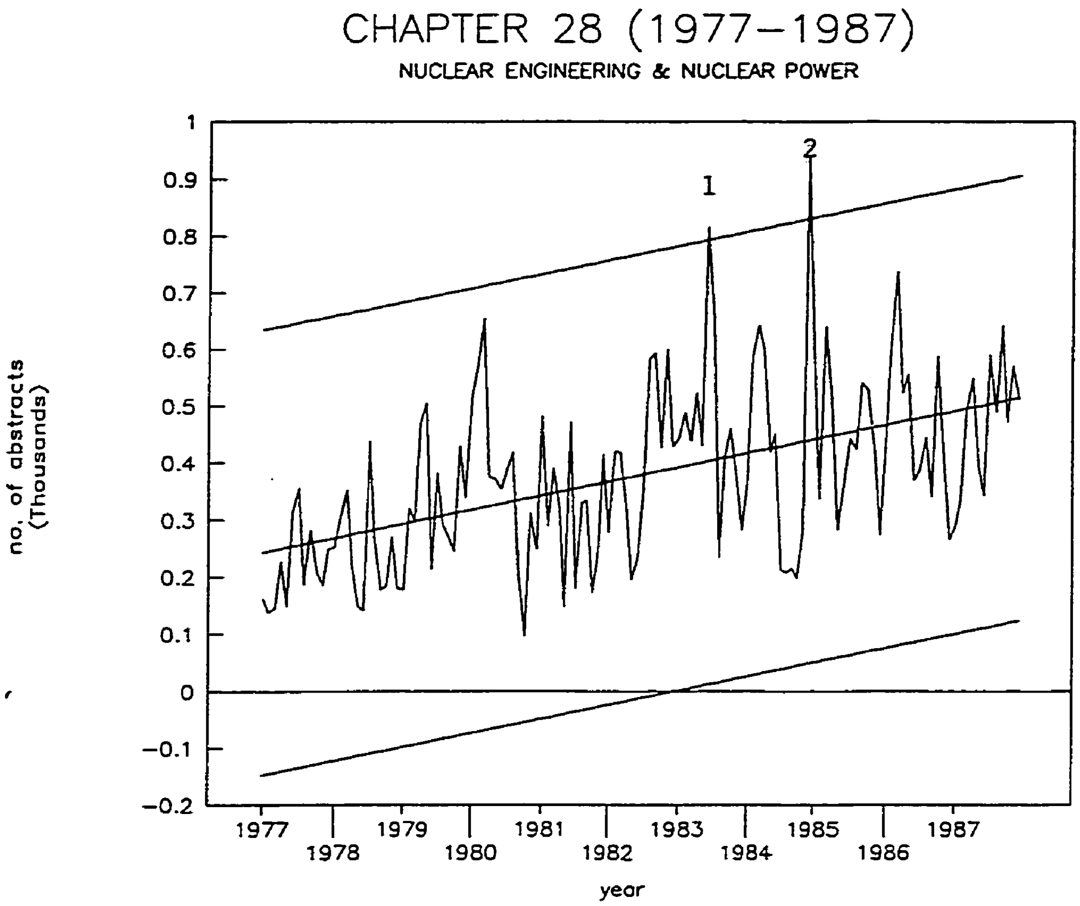


Figure A16. Chapter 28.

Outlier 1 of June 1983 contains 816 abstracts, including the following:

- *1982 Nuclear Science Symposium* and *1982 Symposium on Nuclear Power Systems*, Washington, DC, October 1982 (35 notices),
- *Actinide Recovery from Waste and Low Grade Sources*: Proceedings of the International Symposium, New York, August 1981 (22 notices),
- *ASTM-EURATOM Symposium on Reactor Dosimetry, Radiation Metrology Techniques, Data Bases and Standardization*, 4th, Gaithersburg, MD, March 1982 (93 notices),
- *Canadian Nuclear Society, 3rd Annual Conference*, Toronto, June 1982 (34 notices).

Figure A16--Continued

- *Conference on Fast, Thermal, and Fusion Reactor Experiments*, Salt Lake City, April 1982 (67 notices),
- *Fusion Technology Symposium*, 12th, Julich, Germany, September 1982 (103 notices),
- *IMACS World Congress on System Simulation and Scientific Computation*, 10th, Montreal, August 1982 (22 notices),
- *L.M.F.B.R. Safety Topical Meeting*, Lyon-Ecully, France, July 1982 (200 notices), and
- *Neutron and Its Applications, 1982: Conference to Mark the 50th Anniversary of the Discovery of the Neutron*, Cambridge, September 1982 (15 notices).

Papers from all conferences account for 72.4% of the contents.

Outlier 2 of December 1984 contains 937 abstracts, including the following:

- *Annual Meeting of the American Nuclear Society*, New Orleans, June 1984 (409 notices),
- *Annual Symposium on Safeguards and Nuclear Materials Management*, 6th, Venice, May 1984 (98 notices),
- *International Conference on Radioactive Waste Management*, Seattle, WA, May 1983 (53 notices),
- *Jahrestagung Kerntechnik '84: Annual Meeting of Nuclear Technology*, Frankfurt, May 1984 (23 notices),
- *Symposium on Fusion Engineering*, 10th, Philadelphia, December 1983 (171 notices), and
- *Topical Meeting on Fusion Reactor Materials*, 3rd, Albuquerque, NM, September 1983 (65 notices).

Papers from all conferences account for 87.4% of the contents.

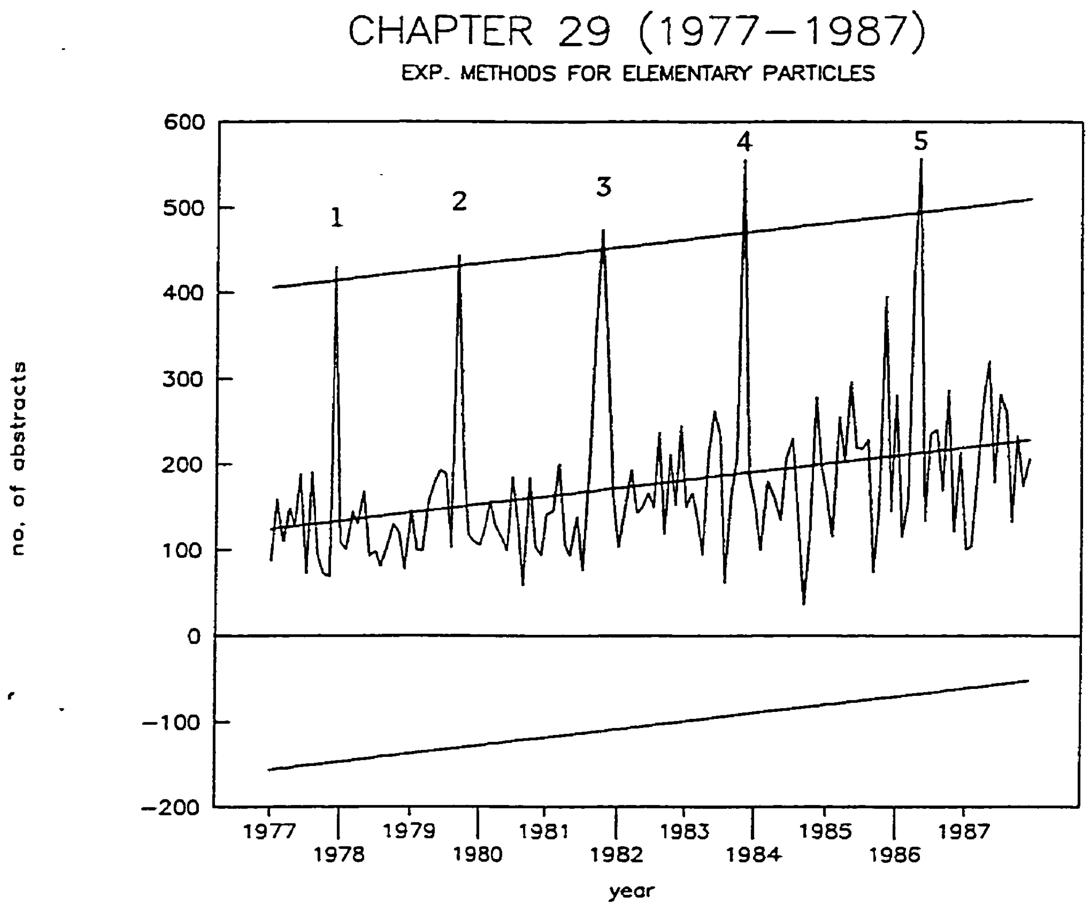


Figure A17. Chapter 29.

Outlier 1 of December 1977 contains 430 abstracts, including the 1977 *Particle Accelerator Conference*, Chicago, March 1977 (290 notices), 67.4% of the contents.

Outlier 2 of September 1979 contains 444 abstracts, including the 1979 *Particle Accelerator Conference*, San Francisco, March 1979 (326 notices), 73.4% of the contents.

Figure A17-- Continued

Outlier 3 of October 1981 contains 474 abstracts, including the following:

- 1981 *Particle Accelerator Conference*, Washington, DC, March 1981 (268 notices),
- *Fifth Tandem Conference*, Catania, June 1980 (36 notices),
- *International Conference on Experimentation at LEP*, Uppsala, June 1980 (26 notices),
- *International Conference on High Energy Accelerators*, 11th, Geneva, July 1980 (28 notices), and
- *International Conference on Polarization Phenomena in Nuclear Physics*, 5th, Santa Fe, NM, 1980 (41 notices).

Papers from all conferences account for 84.2% of the contents.

Outlier 4 of November 1983 contains 555 abstracts, including the following:

- 1983 *Particle Accelerator Conference*, Santa Fe, NM, March 1983 (426 notices),
- *International Workshop on Mercuric Iodide Nucleic Reaction Detectors*, 5th, Jerusalem, June 1982 (191 notices), and
- *Yamada Conference VI on Neutron Scattering in Condensed Matter*, Hakone, Japan, September 1982 (14 notices).

Papers from all conferences account for 82.7% of the contents.

Outlier 5 of May 1986 contains 557 abstracts, including the following:

- 1985 *Particle Accelerator Conference*, Vancouver, May 1985 (258 notices),
- *International Conference on Electrostatic Accelerator Technology and Associated Boosters*, Buenos Aires, April 1985 (56 notices), and
- *Symposium on X and Gamma Ray Sources and Applications*, 6th, Ann Arbor, MI, May 1985 (24 notices).

Papers from all conferences account for 60.7% of the contents.

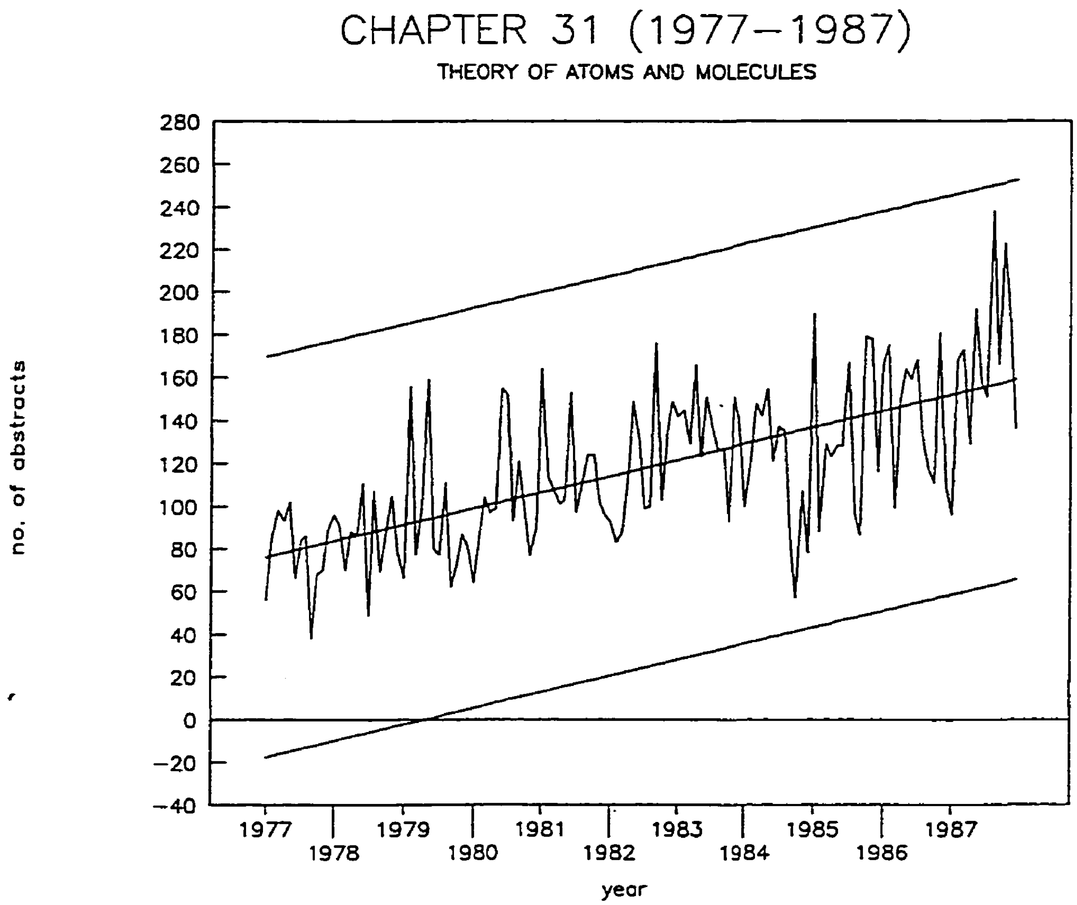


Figure A18. Chapter 31.

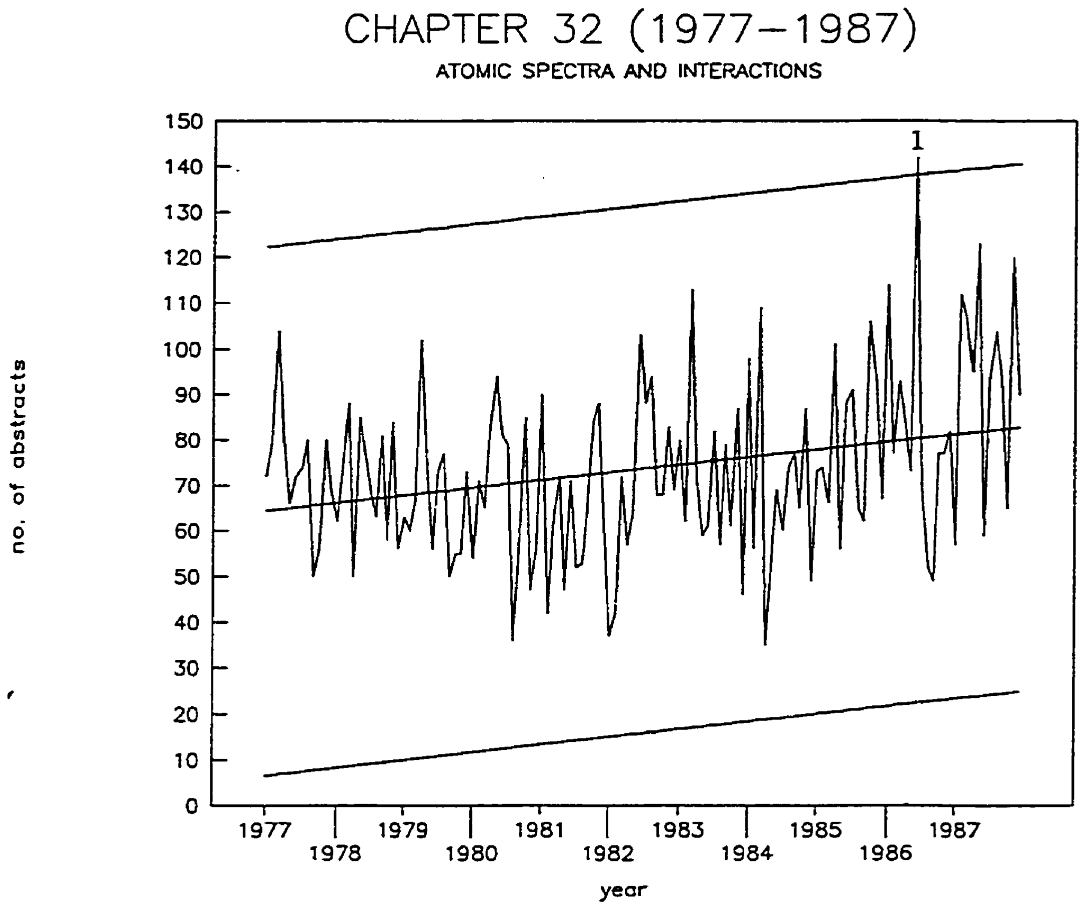


Figure A19. Chapter 32.

Outlier 1 of June 1986 contains 142 abstracts, including the *International Conference on the Physics of Electronic and Atomic Collisions*, 14th, Palo Alto, CA, July 1985 (43 notices). Papers from all conferences account for 30.3% of the contents.

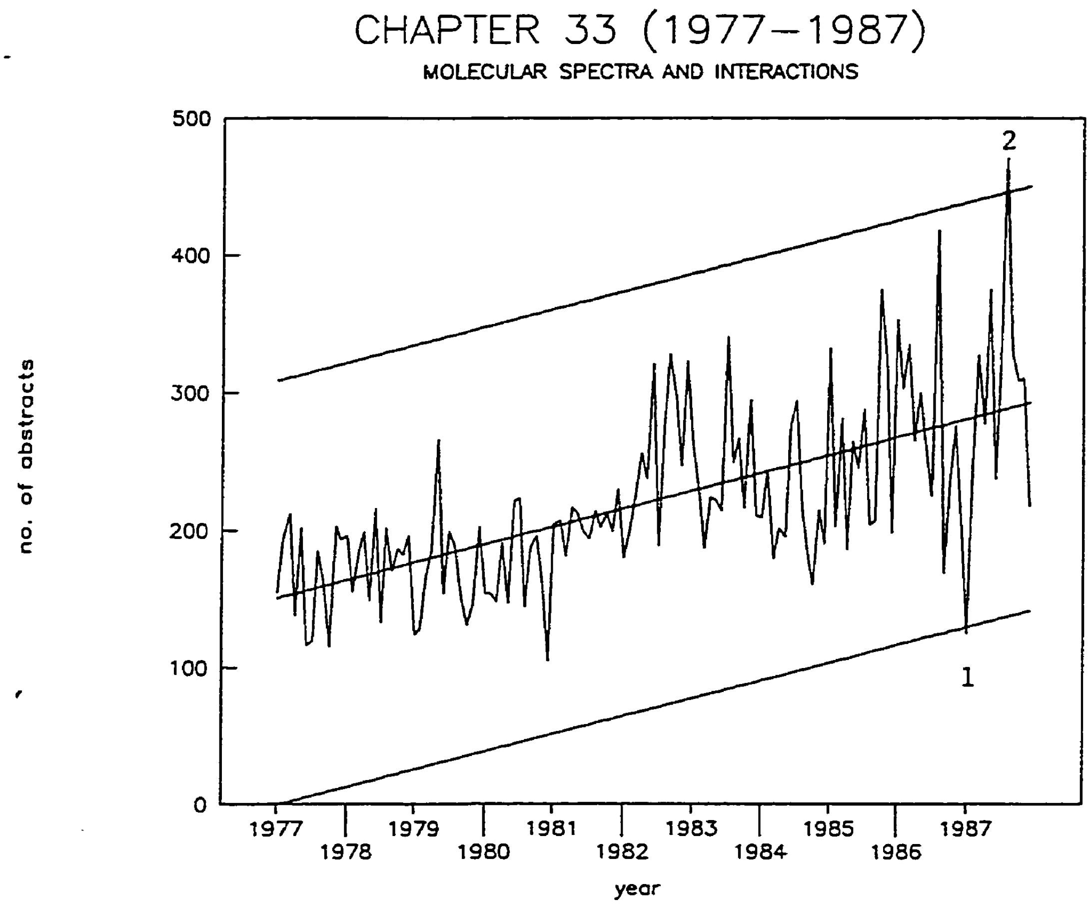


Figure A20. Chapter 33.

Outlier 1 of January 1987 is a negative outlier. It contains 125 abstracts whereas the average month over the last two years carried 287 abstracts. There are no content related characteristics to account for the sudden low volume of inclusion. Possible reasons are discussed in the text (see Section 6.2, page 110).

Outlier 2 of August 1987 contains 471 abstracts, including the *International Conference on Raman Spectroscopy*, 10th, Eugene, OR, September 1986 (88 notices) and the *International Conference on Resonance Ionization Spectroscopy*, 3rd, Swansea, Wales, September 1986 (12 notices). Papers from all conferences account for 21.2% of the contents.

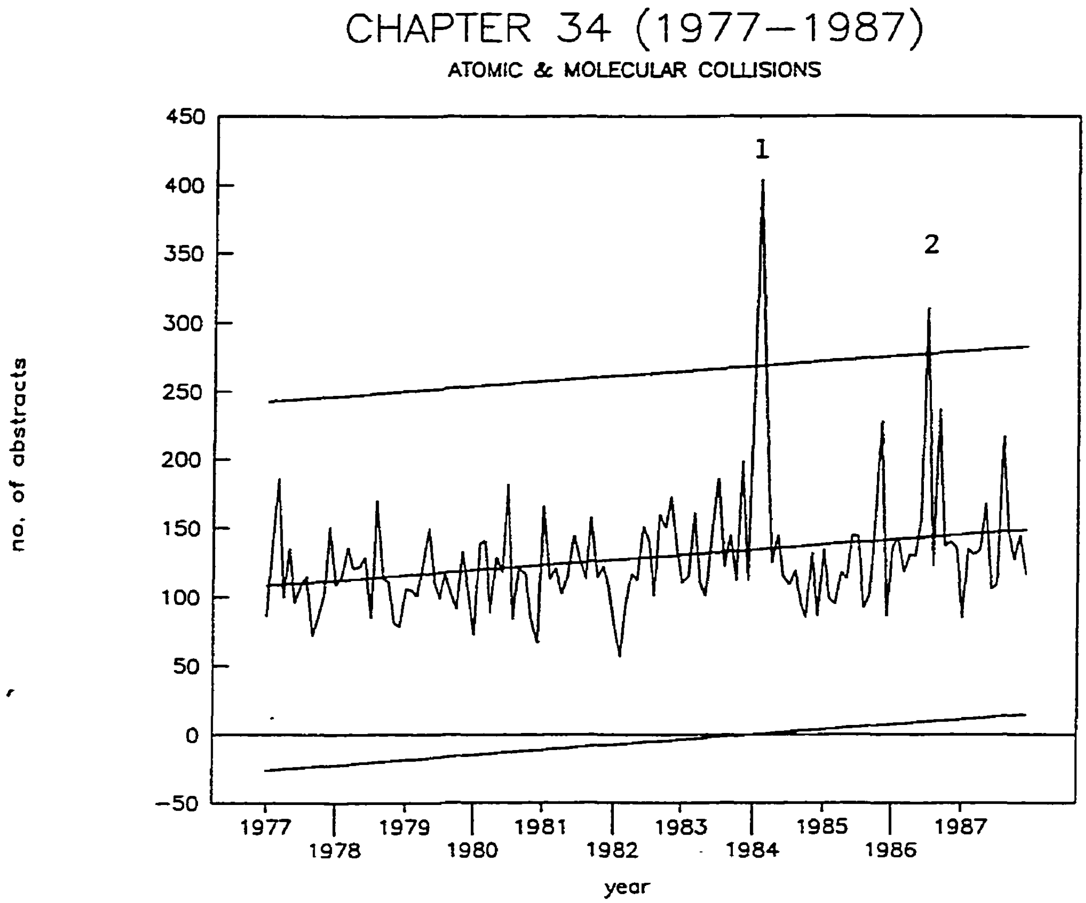


Figure A21. Chapter 34.

Outlier 1 of February 1984 contains 404 abstracts, including the *International Conference on the Physics of Electronic and Atomic Collisions*, 13th, Berlin, August 1983 (320 notices). Papers from all conferences account for 79.2% of the contents.

Outlier 2 of July 1986 contains 311 abstracts, including the *International Conference on the Physics of Electronic and Atomic Collisions*, 14th, Palo Alto, CA, July 1985 (223 notices). Papers from all conferences account for 71.2% of the contents.

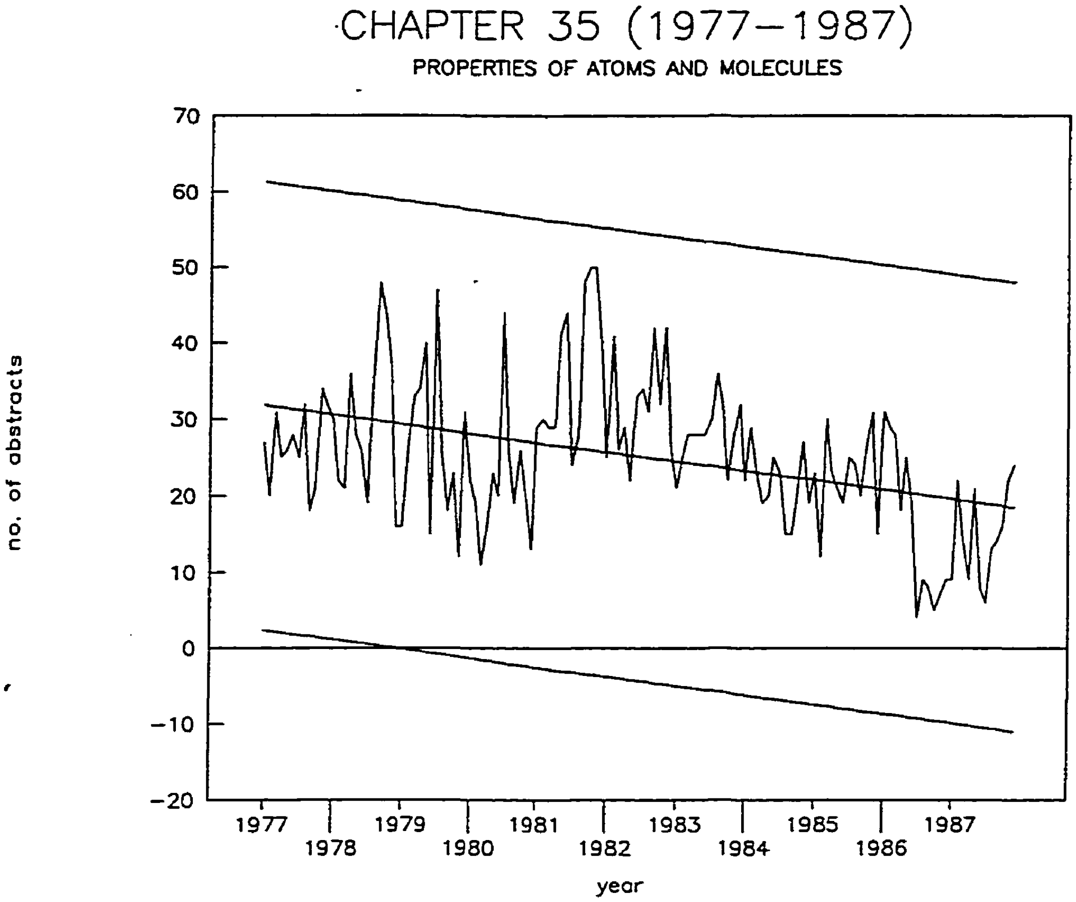


Figure A22. Chapter 35.

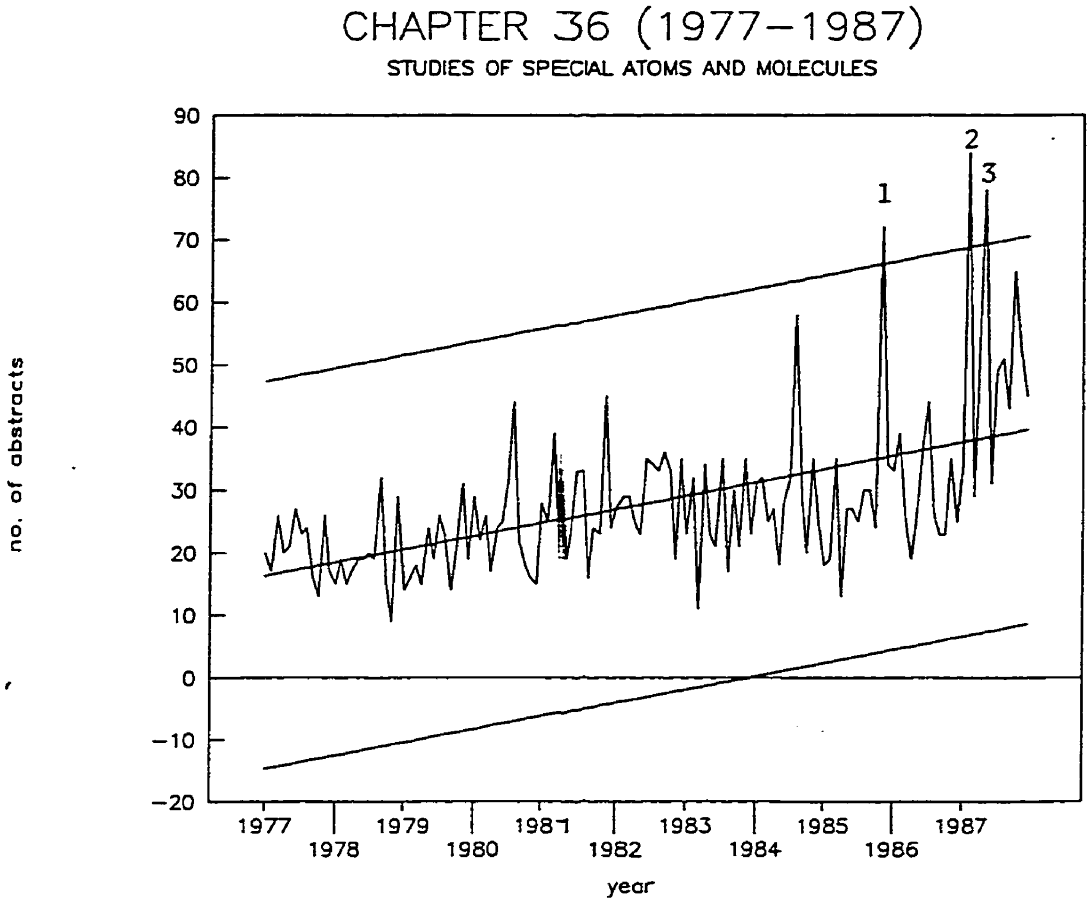


Figure A23. Chapter 36.

Outlier 1 of November 1985 contains 72 abstracts, including *International Meeting on Small Particles and Inorganic Clusters*, 3rd, Berlin, July 1984 (28 notices). Papers from all conferences account for 38.9% of the contents.

Outlier 2 of February 1987 contains 84 abstracts, including the *International Symposium on Metal Clusters*, Heidelberg, April 1986 (28 notices). Papers from all conferences account for 33.3% of the contents.

Figure A23--Continued

Outlier 3 of May 1987 contains 78 abstracts, including the *International Conference on Muon Spin Rotation, Relaxation and Resonance*, Uppsala, June 1986 (18 notices). Papers from all conferences account for 23.1% of the contents.

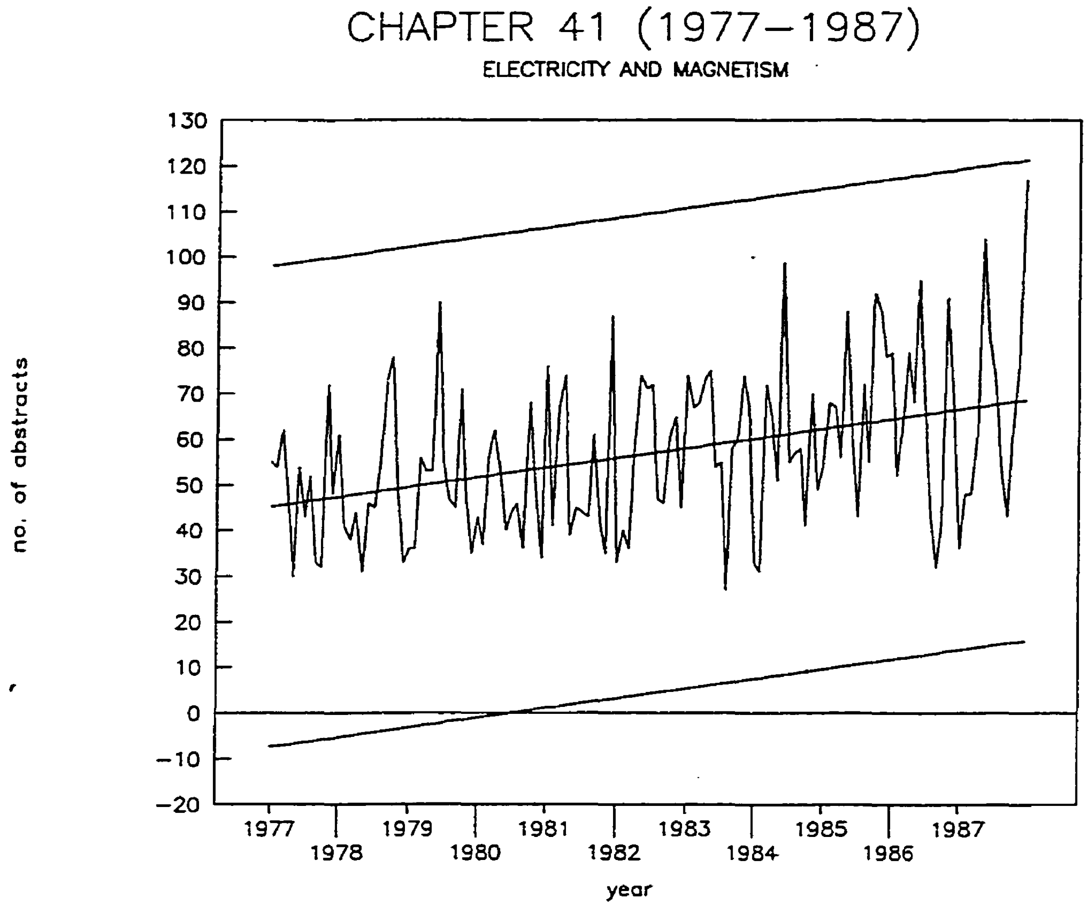


Figure A24. Chapter 41.

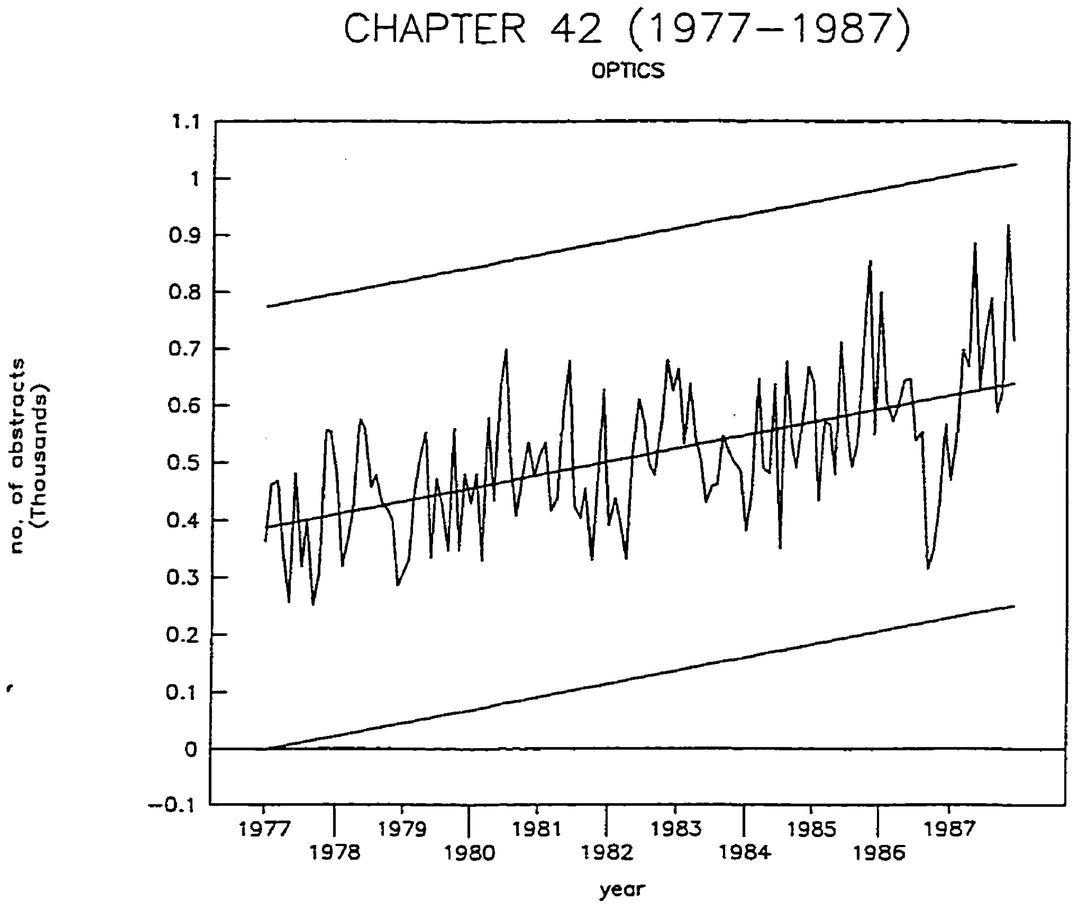


Figure A25. Chapter 42.

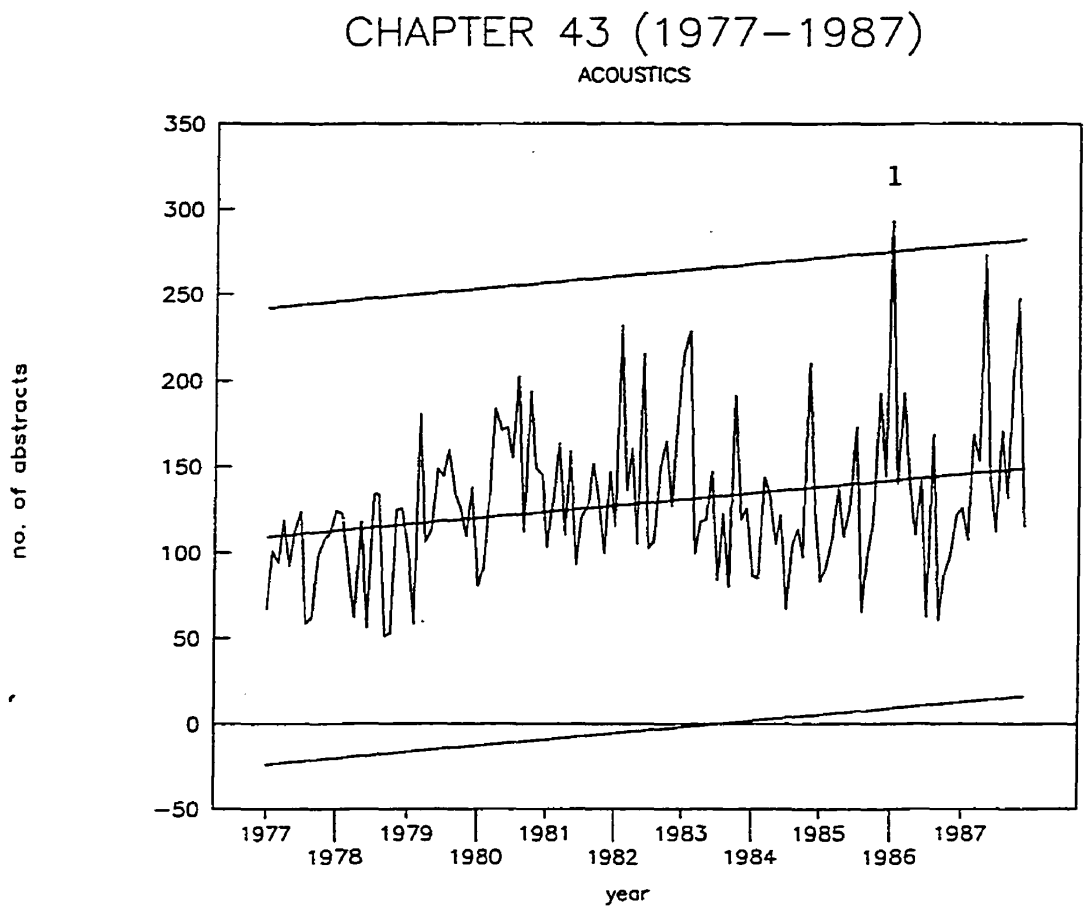


Figure A26. Chapter 43.

Outlier 1 of January 1986 contains 293 abstracts, including the *IEEE 1984 Ultrasonics Symposium*, Dallas, TX, November 1984 (107 notices). Papers from all conferences account for 36.5% of the contents.

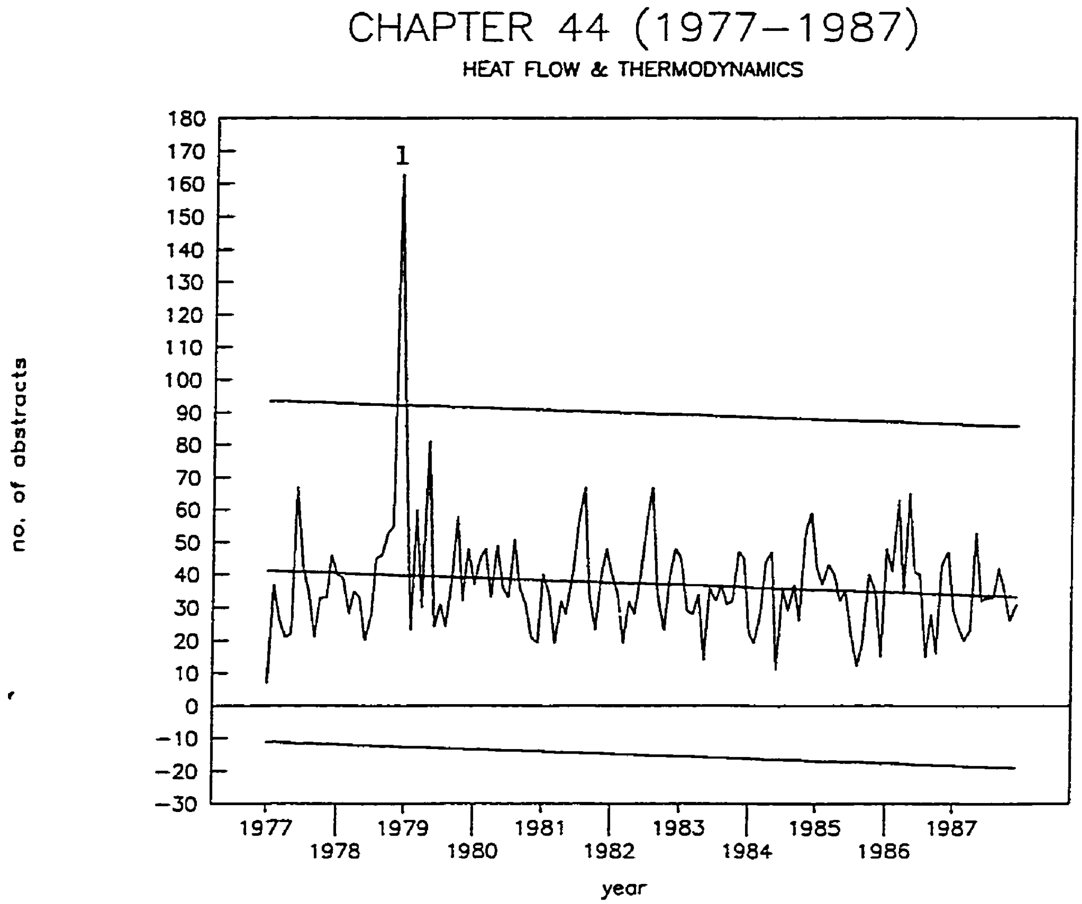


Figure A27. Chapter 44.

Outlier 1 of December 1978 contains 163 abstracts, including the *International Heat Transfer Conference*, 6th, Toronto, August 1978 (104 notices). Papers from all conferences account for 63.8% of the contents.

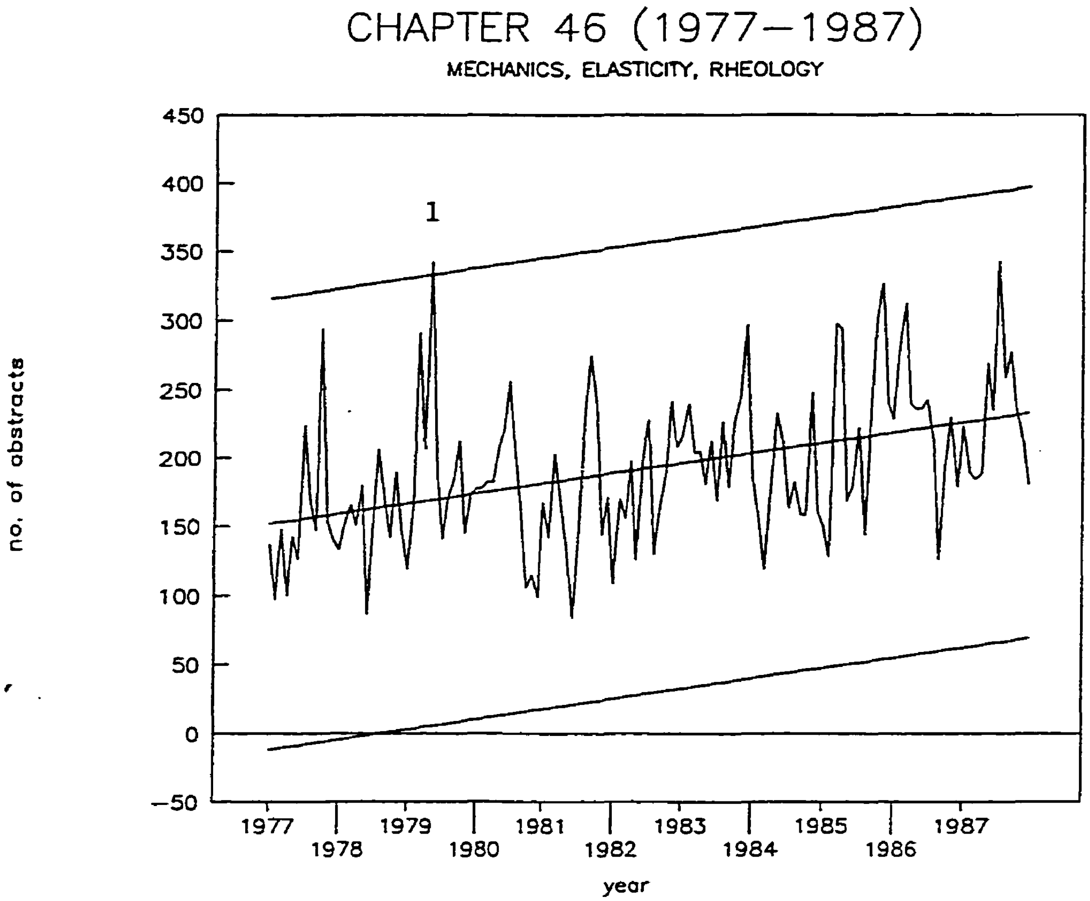


Figure A28. Chapter 46.

Outlier 1 of May 1979 contains 342 abstracts, including the *International Conference on Experimental Stress Analysis*, 6th, Munich, September 1978 (33 notices). Papers from all conferences account for 9.6% of the contents.

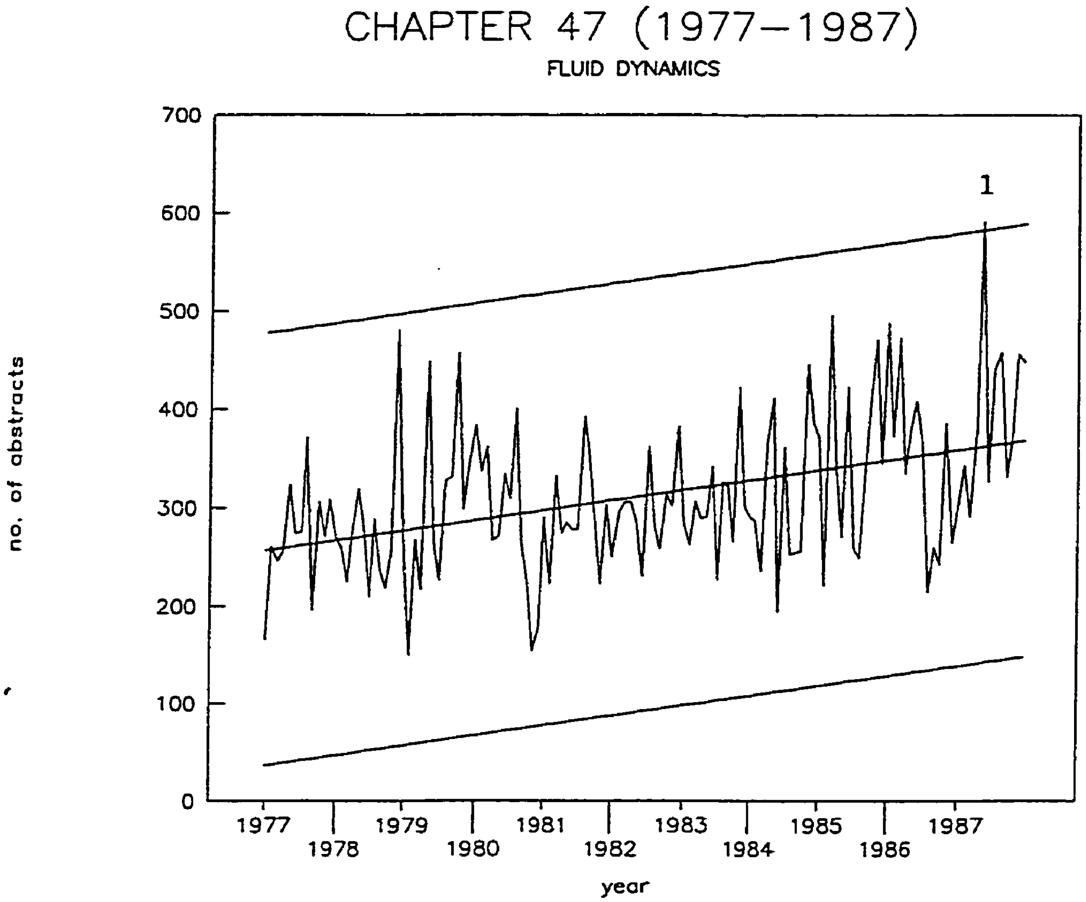


Figure A29. Chapter 47.

Outlier 1 of May 1987 contains 592 abstracts, including the following:

- *International Symposium on Finite Element Methods in Flow Problems*, 6th, Antibes, France, June 1986 (26 notices),
- *IUTAM Symposium on Fluid Mechanics in the Spirit of G.I. Taylor*, Cambridge, England, March 1986 (21 notices), and
- *Spatio-Temporal Coherence and Chaos in Physical Systems*, Los Alamos, NM, January 1986 (12 notices).

Papers from all conferences account for 10% of the contents.

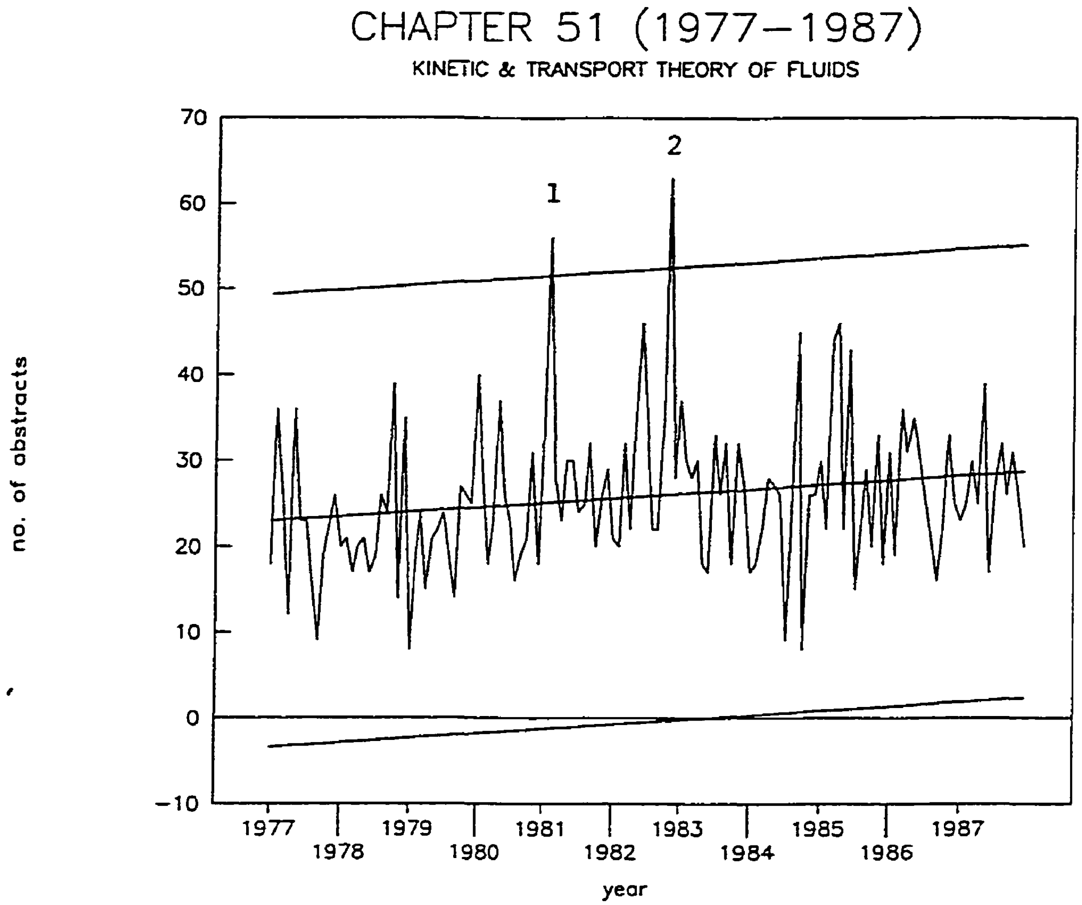


Figure A30. Chapter 51.

Outlier 1 of February 1981 contains 56 abstracts, including the *International Conference on Gas Discharges and Their Applications*, 6th, Edinburgh, September 1980 (29 notices). Papers from all conferences account for 51.8% of the contents.

Outlier 2 of November 1982 contains 63 abstracts, including the *International Conference on Gas Discharges and Their Applications*, 7th, London, September 1982 (24 notices). Papers from all conferences account for 38.1% of the contents.

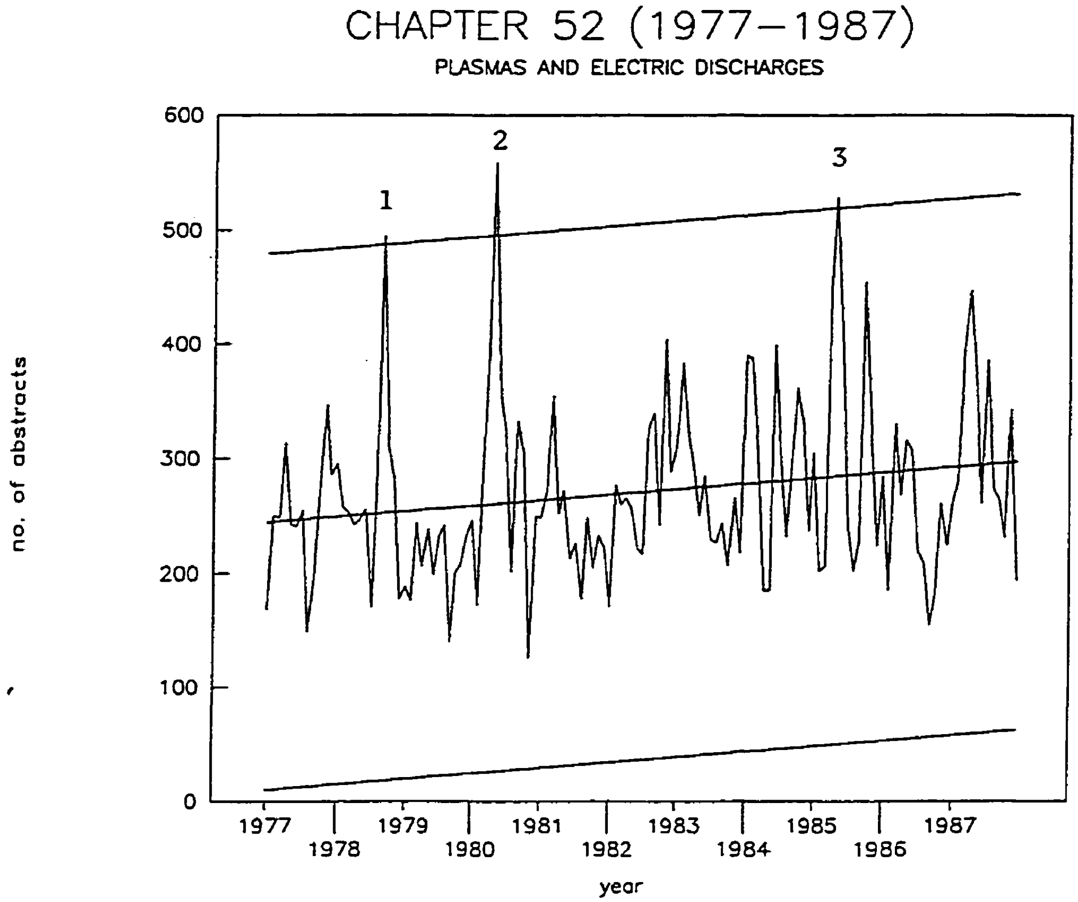


Figure A31. Chapter 52.

Outlier 1 of September 1978 contains 495 abstracts, including the *International Conference on Phenomena in Ionized Gases*, 13th, Berlin, September 1977 (257 notices). Papers from all conferences account for 51.9% of the contents.

Outlier 2 of May 1980 contains 559 abstracts, including the *International Conference on Phenomena in Ionized Gases*, 14th, Grenoble, July 1979 (331 notices). Papers from all conferences account for 59.2% of the contents.

Figure A31.--Continued.

Outlier 3 of May 1985 contains 529 abstracts, including the following:
- *1984 IEEE International Conference on Plasma Science*, St. Louis, MO, May 1984 (138 notices),
- *International Conference on Plasma Surface Interactions in Controlled Fusion Devices*, 6th, Nagoya, May 1984 (56 notices),
- *International Symposium on Heating in Toroidal Plasma*, 4th, Rome, March 1984 (118 notices), and
- *Symposium on Plasma Double Layers and Related Topics*, 2nd, Innsbruck, July 1984 (17 notices).

Papers from all conferences account for 62.2% of the contents.

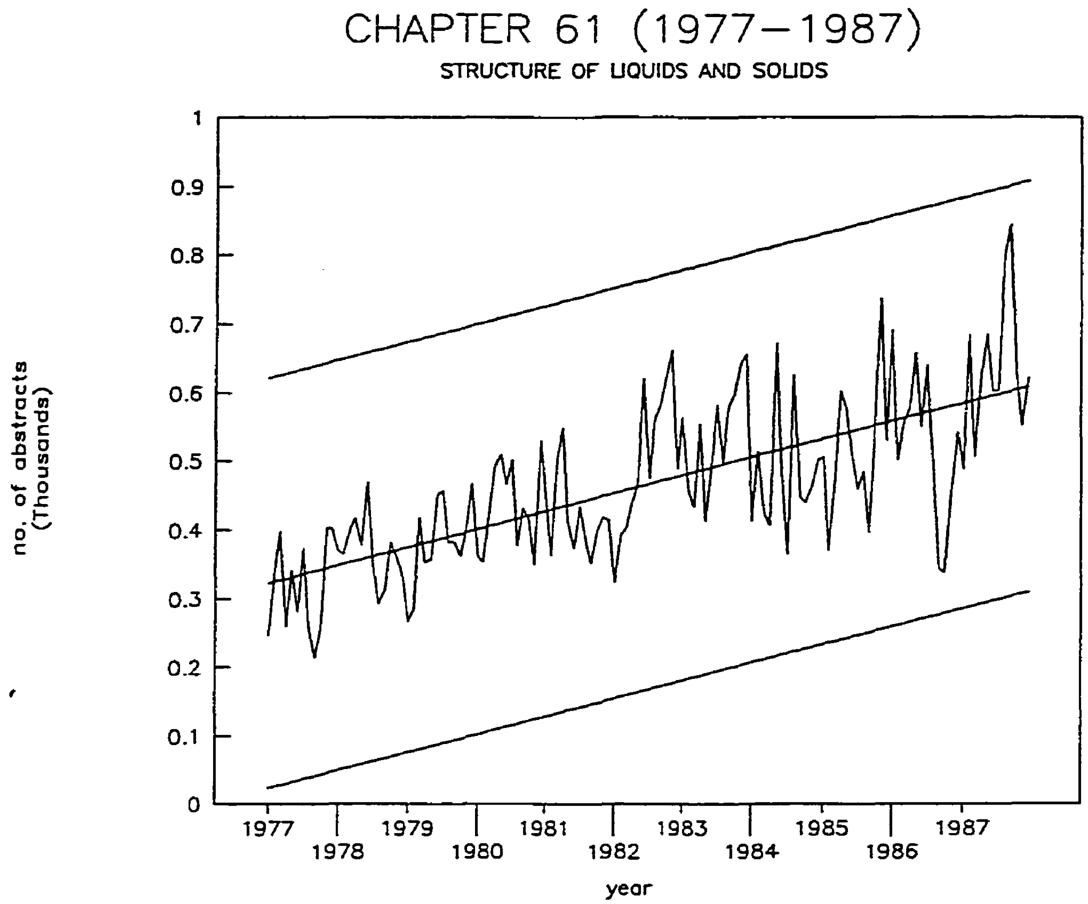


Figure A32. Chapter 61.

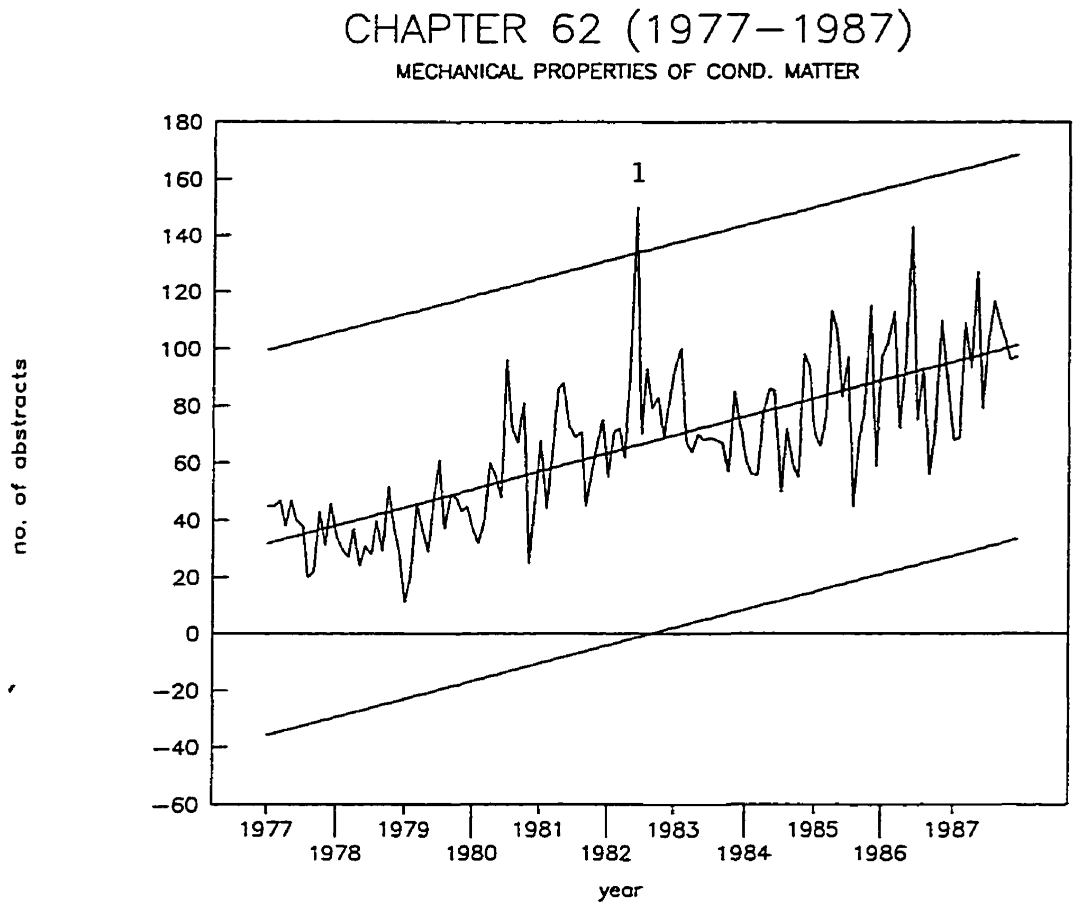


Figure A33. Chapter 62.

Outlier 1 of June 1982 contains 150 abstracts, including the *International Conference on Internal Friction and Ultrasonic Attenuation in Solids*, 7th, Lausanne, July 1981 (48 notices) and the *International Conference on Phonon Physics*, Bloomington, IN, August-September 1981 (13 notices). Papers from all conferences account for 40.7% of the contents.

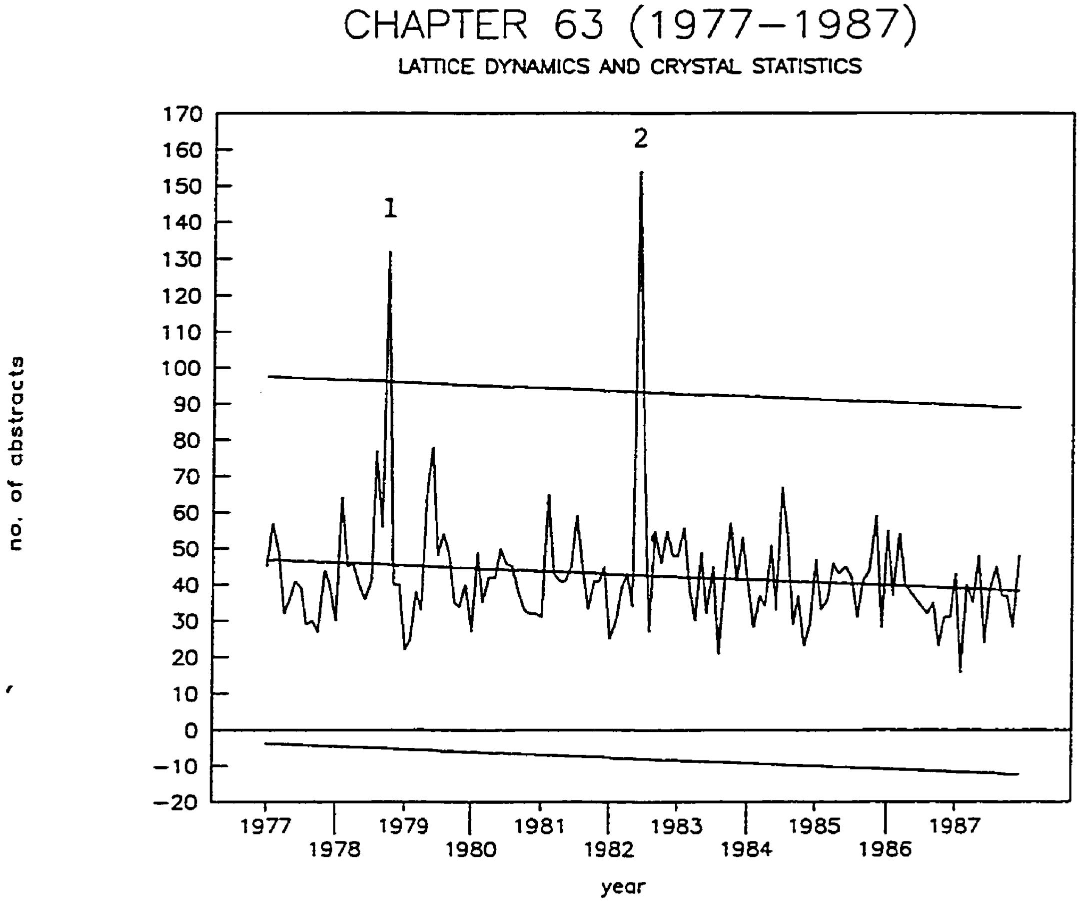


Figure A34. Chapter 63.

Outlier 1 of October 1978 contains 132 abstracts, including the *International Conference on Lattice Dynamics*, Paris, September 1977 (79 notices). Papers from all conferences account for 59.8% of the contents.

Outlier 2 of June 1982 contains 154 abstracts, including the *International Conference on Phonon Physics*, Bloomington, IN, August-September 1981 (87 notices). Papers from all conferences account for 56.5% of the contents.

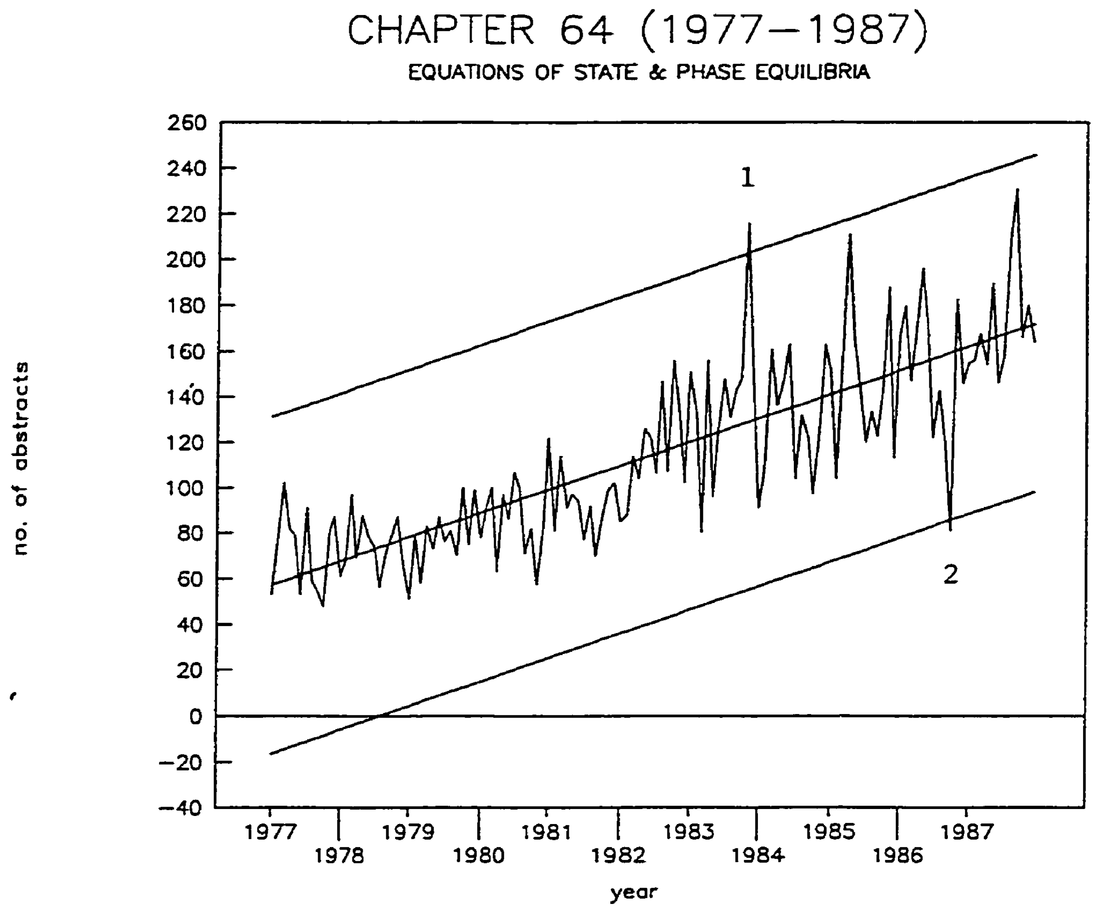


Figure A35. Chapter 64.

Outlier 1 of November 1983 contains 216 abstracts consisting of articles from various journals and conference proceedings.

Outlier 2 of October 1986 contains 81 notices (average 163 over two years) and is a NEGATIVE outlier. There are no content related characteristics to account for the sudden low volume of inclusion. Possible reasons are discussed in the text (see Section 6.2, page 110).

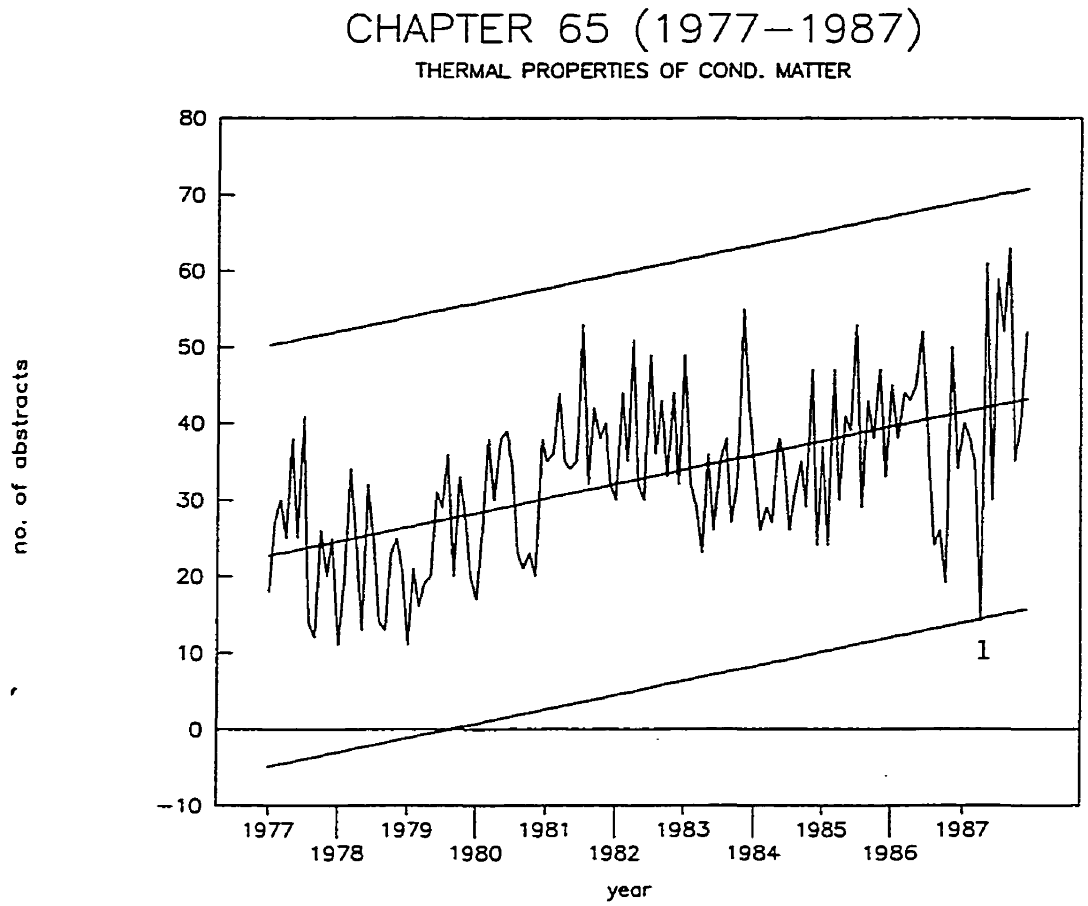


Figure A36. Chapter 65.

Outlier 1 of April 1987 contains 14 notices (average 37 over two years) and is a NEGATIVE outlier. There are no content related characteristics to account for the sudden low volume of inclusion. Possible reasons are discussed in the text (see Section 6.2, page 110).

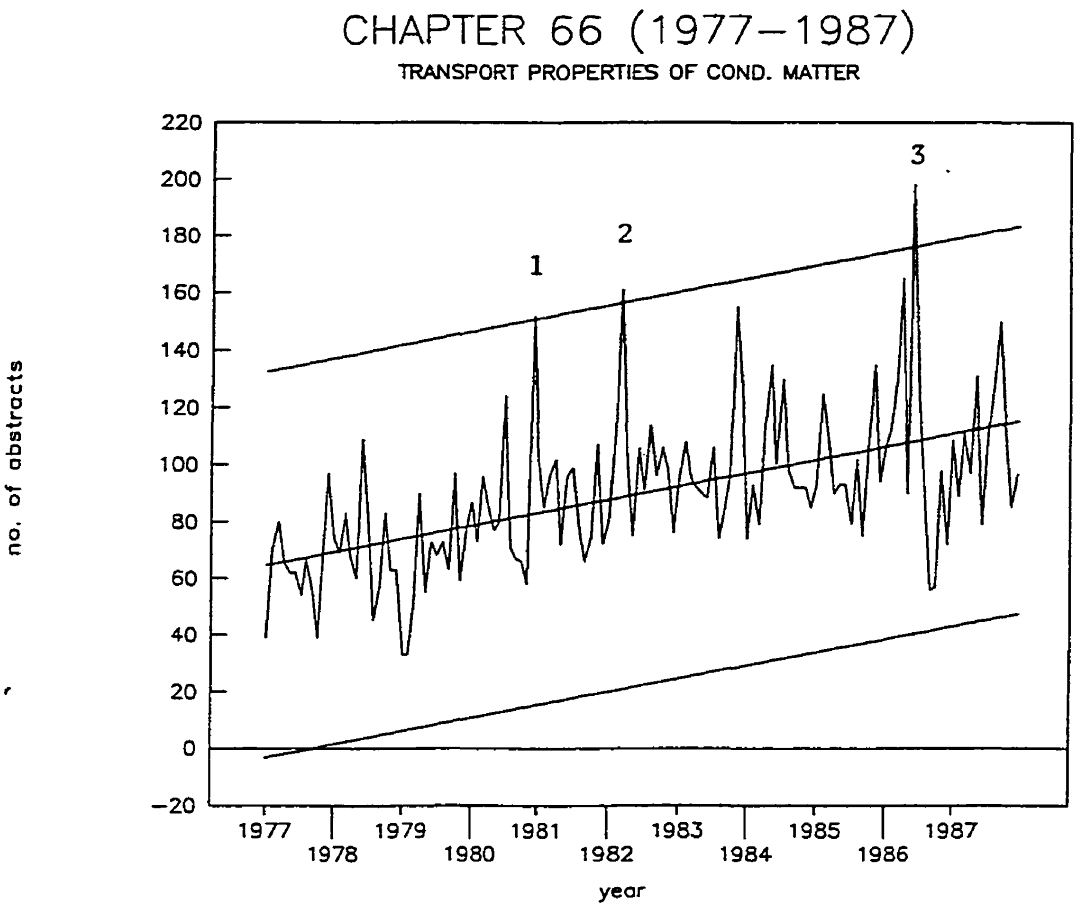


Figure A37. Chapter 66.

Outlier 1 of December 1980 contains 152 abstracts, including the *Europhysics Topical Conference: Lattice Defects in Ionic Crystals*, 3rd, Canterbury, England, September, 1979 (16 notices) and *Fast Ion Transport in Solids, Electrodes and Electrolytes*, Lake Geneva, WI, May 1979 (65 notices). Papers from all conferences account for 53.3% of the contents.

Figure A37.--Continued.

Outlier 2 of March 1982 contains 161 abstracts, including the *International Meeting on Solid Electrolytes - Solid State Ionics and Galvanic Cells*, 3rd, Tokyo, September 1980 (31 notices) and the *International Conference on Fast Ionic Transport in Solids*, Gatlinburg, TN, May 1981 (57 notices). Papers from all conferences account for 54.7% of the contents.

Outlier 3 of June 1986 contains 198 abstracts, including the *International Conference on Solid State Ionics*, 5th, Lake Tahoe, NV, August 1985 (93 notices). Papers from all conferences account for 47% of the contents.

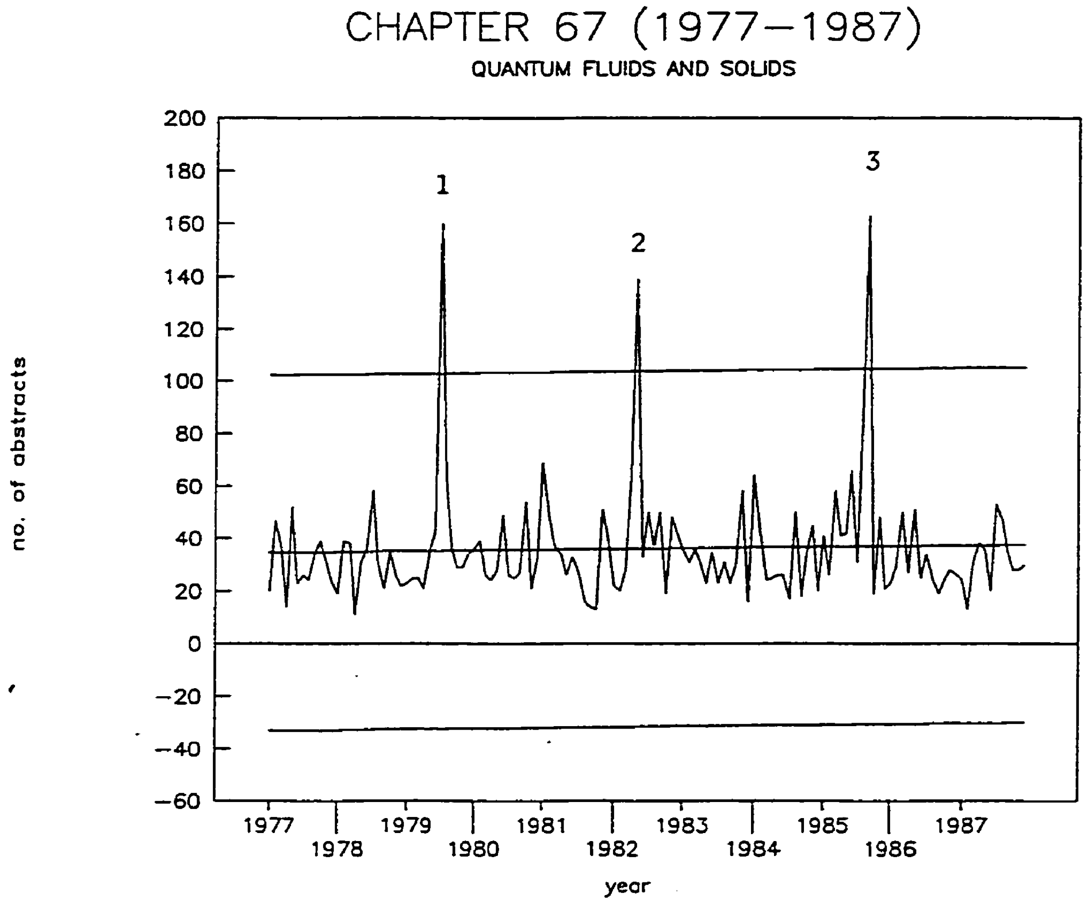


Figure A38. Chapter 67.

Outlier 1 of July 1979 contains 160 abstracts, including the *International Conference on Low Temperature Physics*, 15th, Grenoble, August 1978 (130 notices). Papers from all conferences account for 81.3% of the contents.

Outlier 2 of May 1982 contains 139 abstracts, including the *International Conference on Low Temperature Physics*, 16th, Los Angeles, August 1981 (124 notices). Papers from all conferences account for 89.2% of the contents.

Figure A38.--Continued.

Outlier 3 of September 1985 contains 163 abstracts, including the *International Conference on Low Temperature Physics*, 17th, Karlsruhe, August 1984 (140 notices). Papers from all conferences account for 85.9% of the contents.

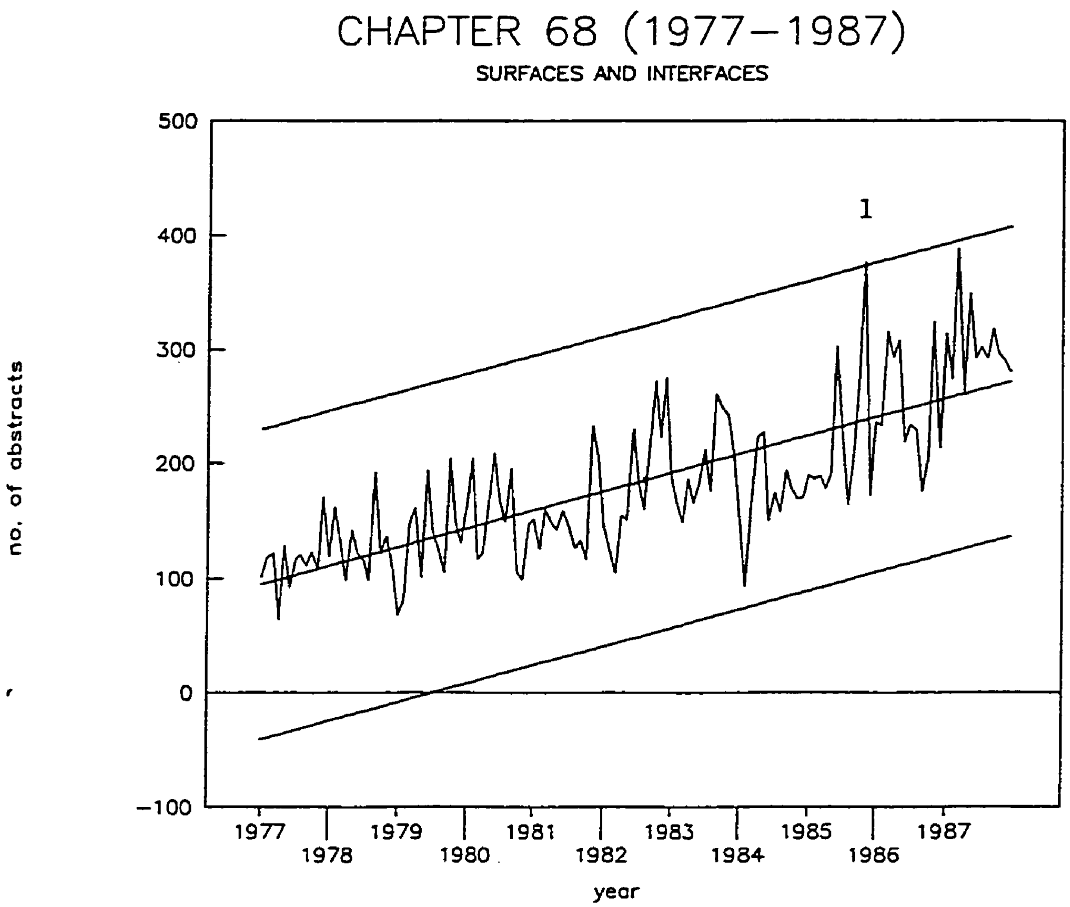


Figure A39. Chapter 68.

Outlier 1 of November 1985 contains 377 abstracts, including the *International Conference on Solid Films and Surfaces*, Sydney, Australia, August 1984 (29 notices). Papers from all conferences account for 7.7% of the contents.

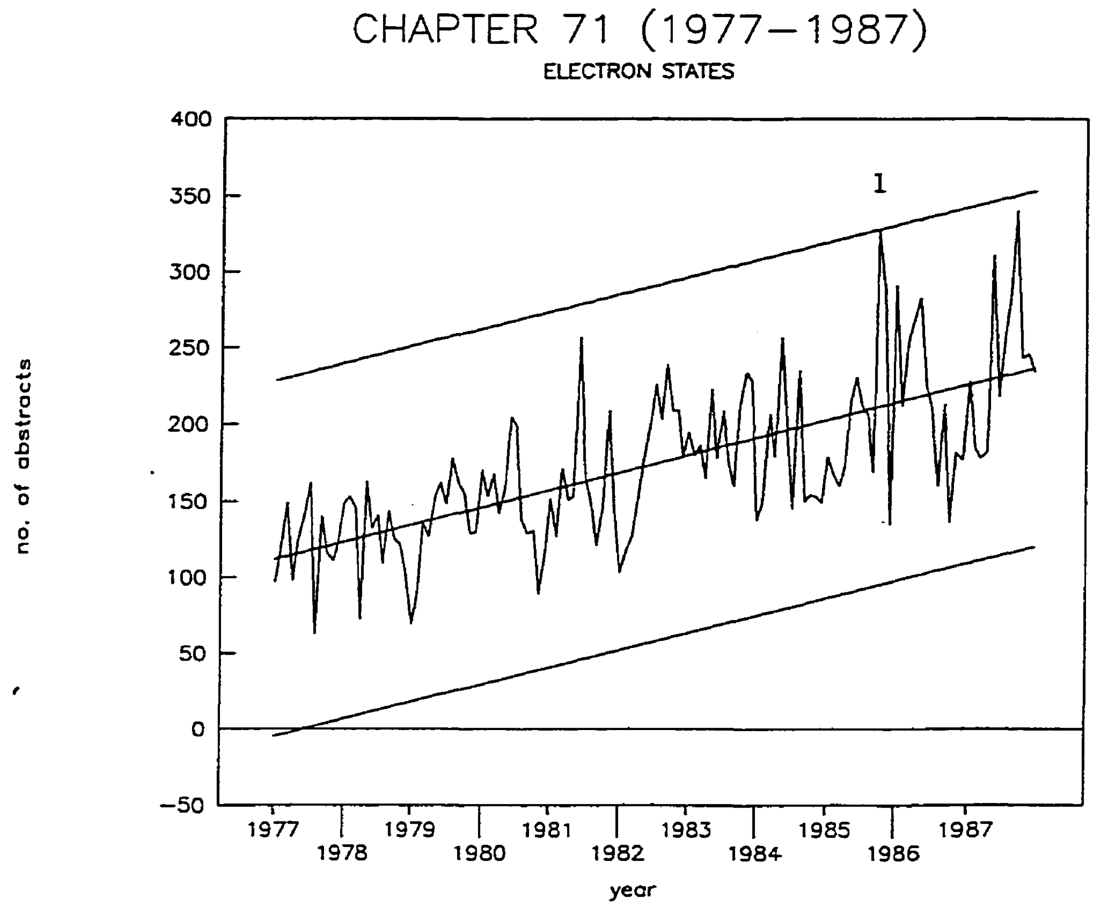


Figure A40. Chapter 71.

Outlier 1 of October 1985 contains 328 abstracts, including the *International Conference on Valence Fluctuations*, 4th, Cologne, August 1984 (73 notices). Papers from all conferences account for 22.3% of the contents.

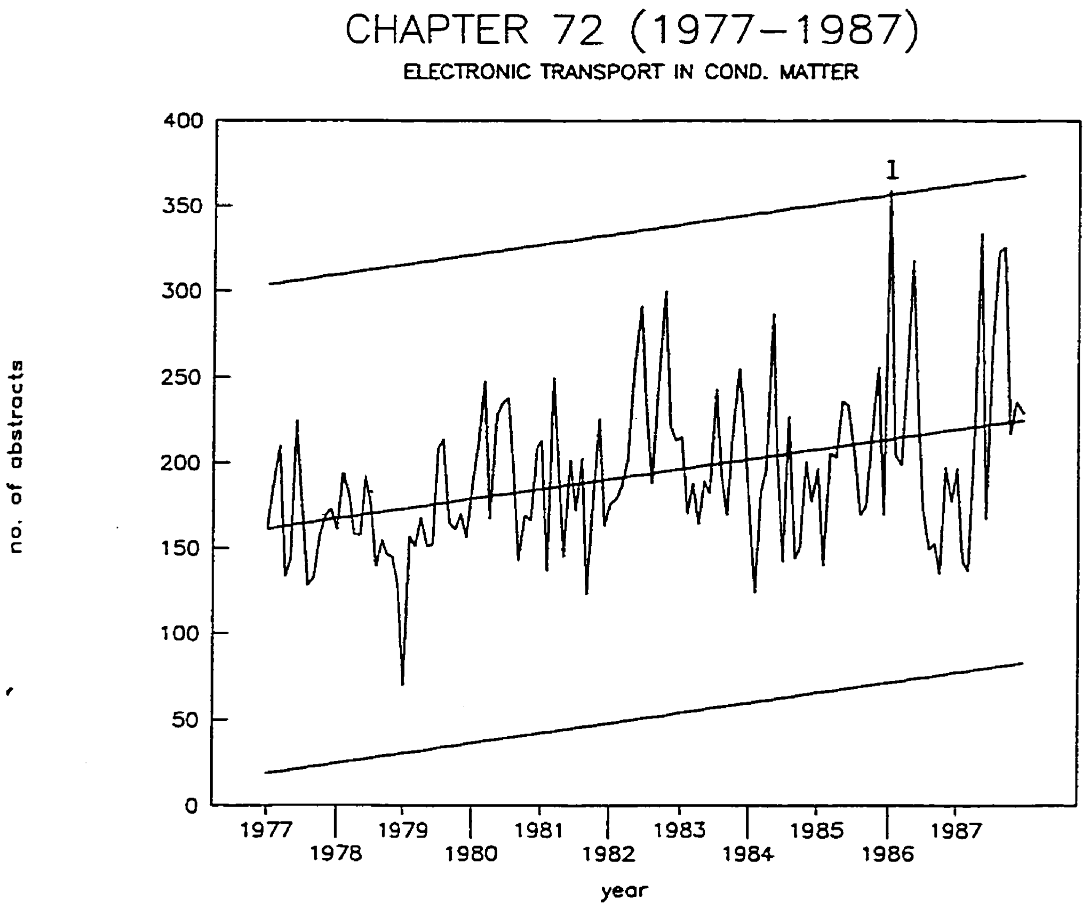


Figure A41. Chapter 72.

Outlier 1 of January 1986 contains 359 abstracts, including the *International Conference on the Physics and Chemistry of Low-Dimensional Synthetic Metals (ICSM '84)*, Abano Terme, Italy, June 1984 (96 notices). Papers from all conferences account for 26.7% of the contents.

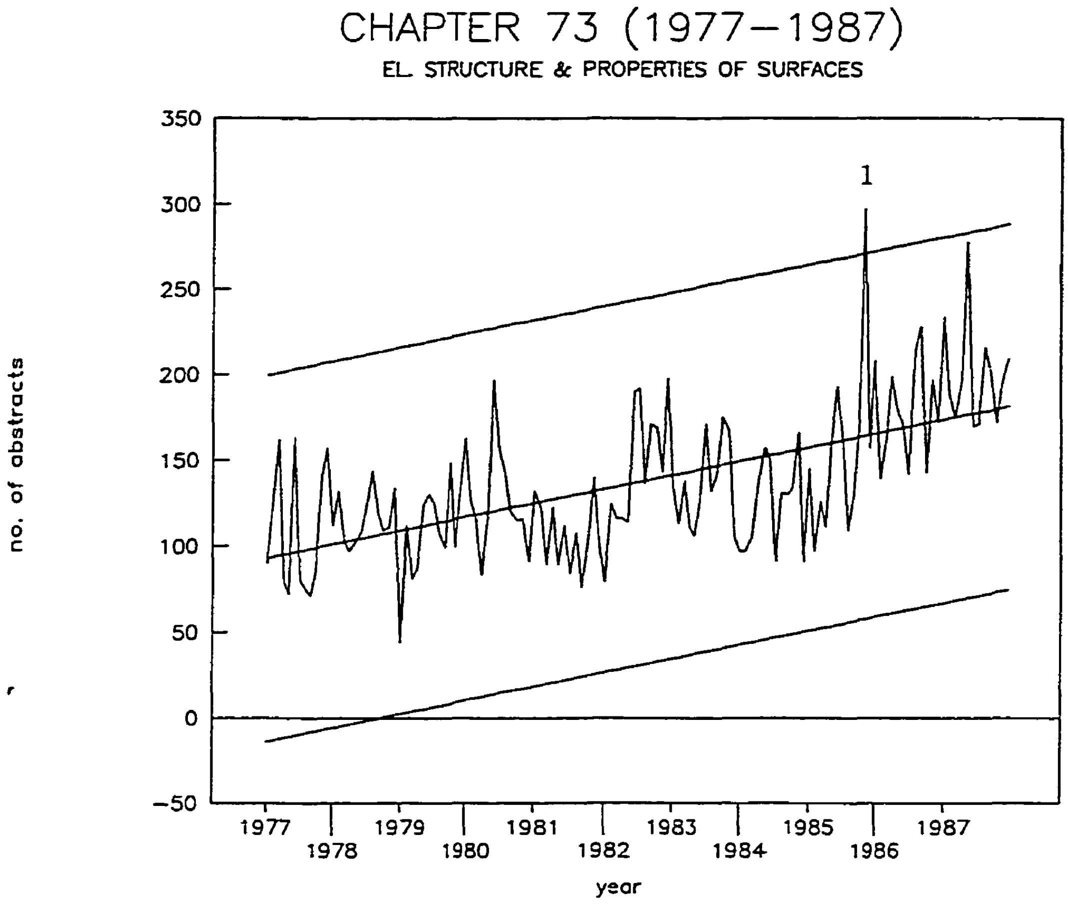


Figure A42. Chapter 73.

Outlier 1 of November 1985 contains 297 abstracts, including the *International Conference on Solid Films and Surfaces*, Sydney, Australia, August 1984 (20 notices) and the *National Symposium on the American Vacuum Society*, 31st, Reno, NV, December 1984 (20 notices). Papers from all conferences account for 13.5% of the contents.

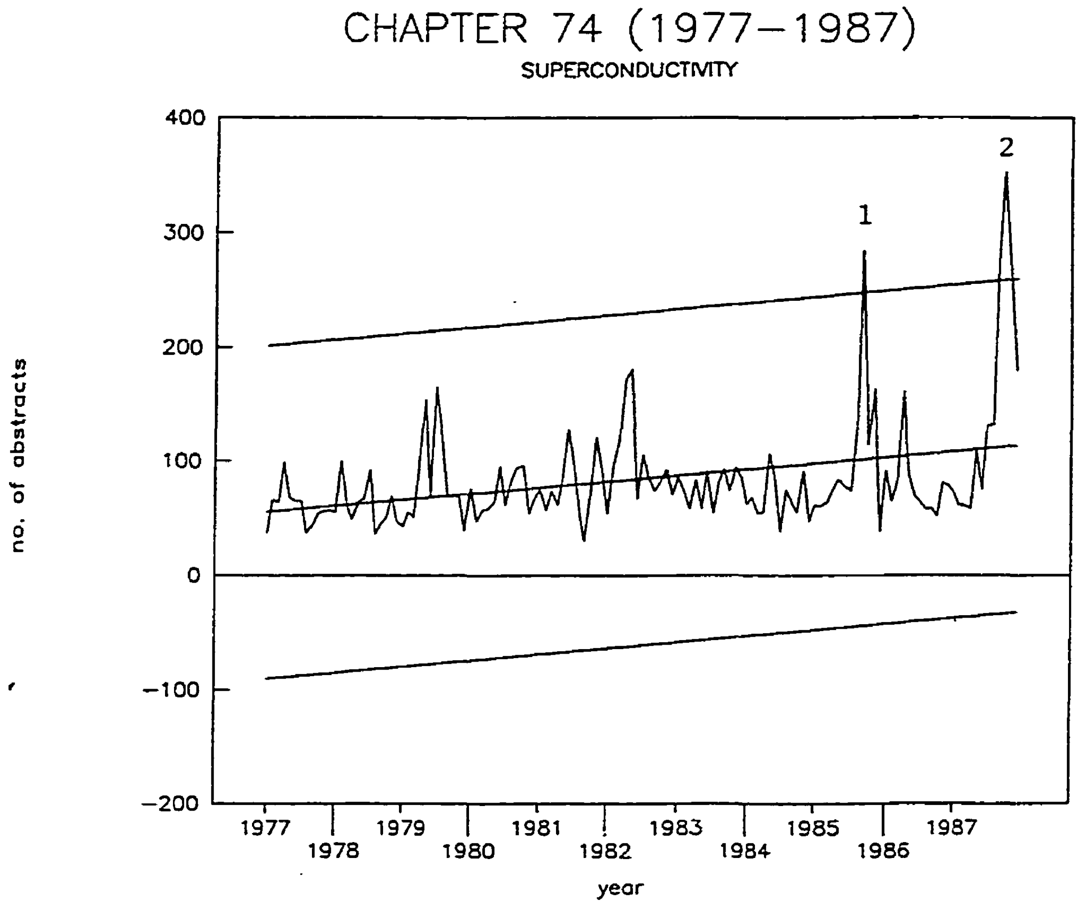


Figure 43. Chapter 74.

Outlier 1 of September 1985 contains 284 abstracts, including the *International Conference on Low Temperature Physics*, 17th, Karlsruhe, August 1984 (175 notices) the *International Cryogenic Engineering Conference*, 10th, Helsinki, Aug., 1984 (12 notices) and the *International Cryogenic Materials Conference*, 5th, Colorado Springs, CO, August 1983 (18 notices). Papers from all conferences account for 72.2% of the contents.

Outlier 2 of September-November 1987 contains 903 abstracts, almost all being articles from various journals. There are no discernible groupings of conference proceedings or special journal issues that account for it.

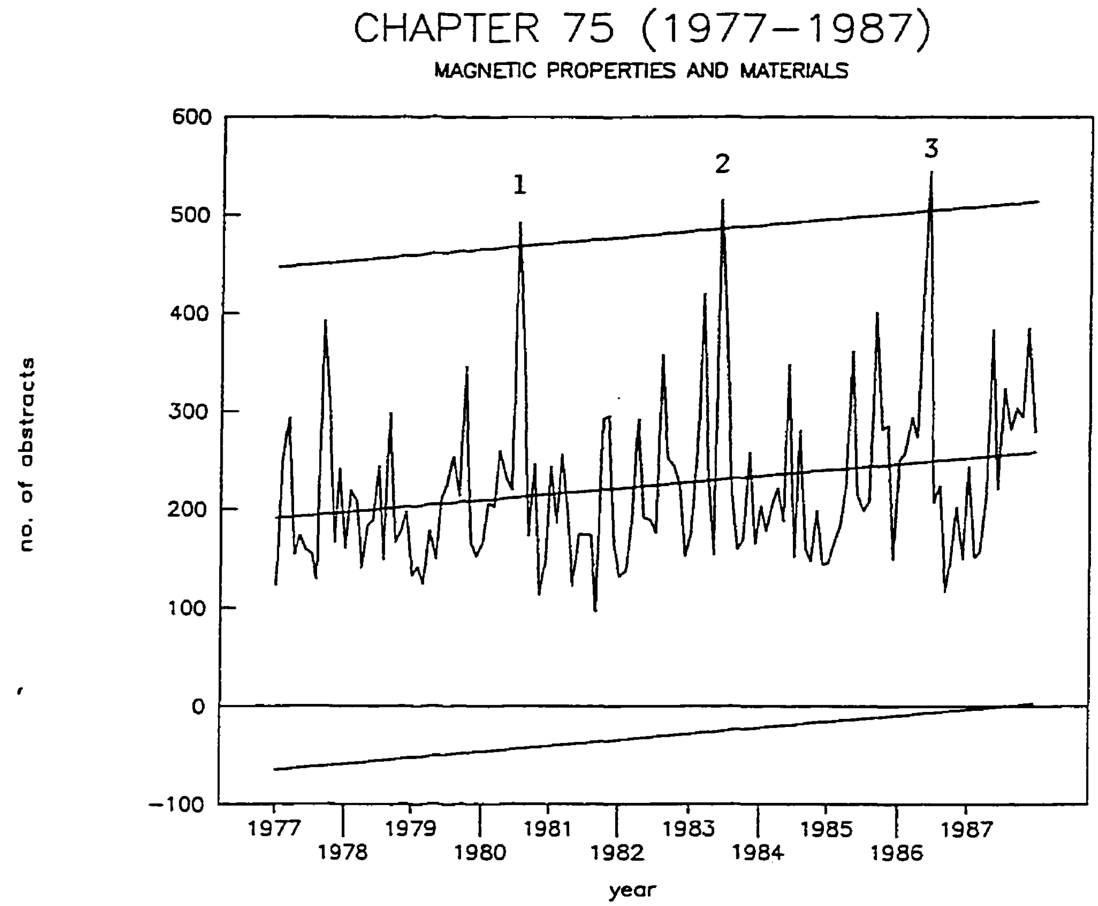


Figure A44. Chapter 75.

Outlier 1 of July 1980 contains 493 abstracts, including the *International Conference on Magnetic Fluids*, 2nd, Orlando, March 1980 (13 notices) and the *International Conference on Magnetism*, Munich, September 1979 (275 notices). Papers from all conferences account for 58.4% of the contents.

Outlier 2 of June 1983 contains 516 abstracts, including the *International Conference on Magnetism*, Kyoto, September 1982 (321 notices) and *Impact of Polarized Neutrons on Solid-State Chemistry and Physics*, Grenoble, October 1982 (37 notices). Papers from all conferences account for 69.4% of the contents.

Figure A44.--Continued.

Outlier 3 of June 1986 contains 545 abstracts, including the *Annual Meeting of the Magnetics Society of Japan*, 8th, Hiroshima, November 1984 (47 notices), the *Conference on Electronic Structure and Properties of Rare Earth and Actinide Intermetallics (REACIM 84)*, St. Polten, Austria, September 1984 (15 notices) and the *International Conference on Magnetism*, San Francisco, August 1985 (296 notices). Papers from all conferences account for 65.7% of the contents.

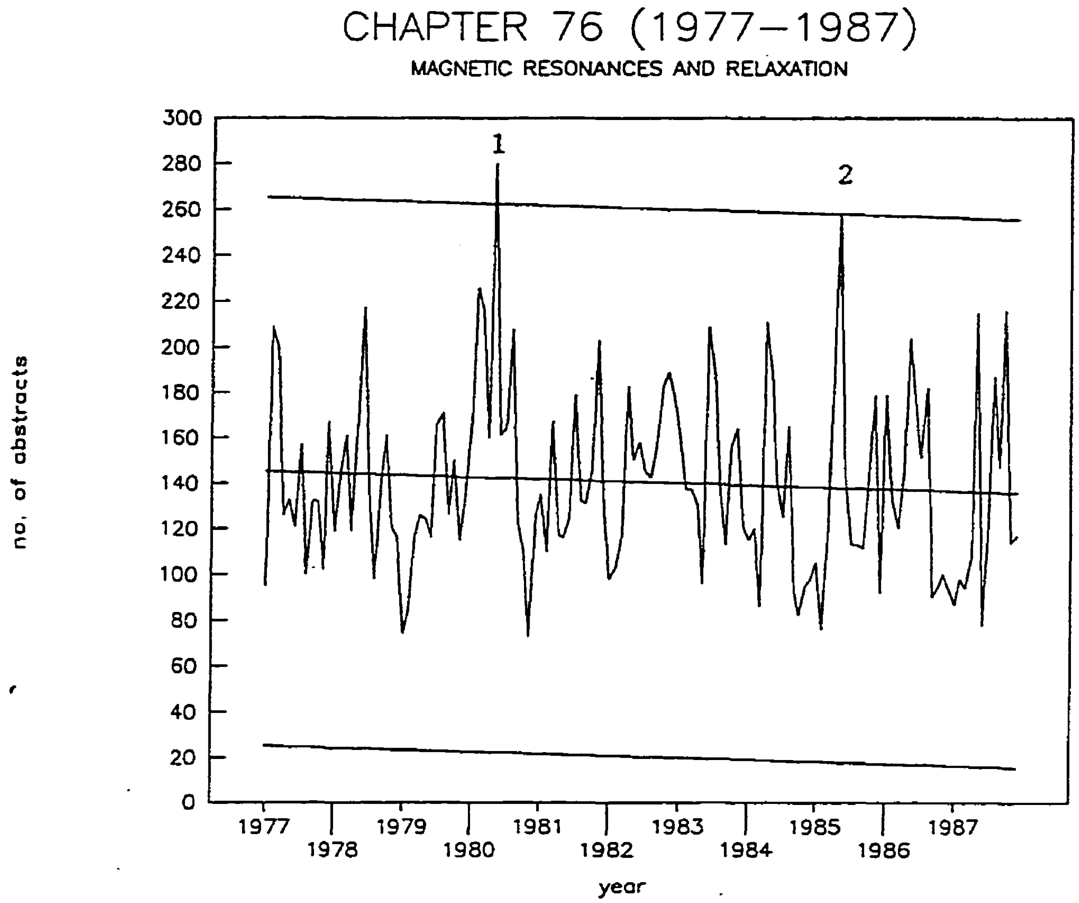


Figure A45. Chapter 76.

Outlier 1 of May 1980 contains 280 abstracts, including the *International Conference on Mossbauer Spectroscopy*, Portorez, Yugoslavia, September 1979 (127 notices) and the *Joint INTERMAG-MMM Conference*, New York, July 1989 (14 notices). Papers from all conferences account for 50.4% of the contents.

Outlier 2 of May 1985 contains 257 abstracts, including *Congrès Ampère on Magnetic Resonance and Related Phenomena*, 22nd, Zurich, September 1984 (102 notices). Papers from all conferences account for 39.7% of the contents.

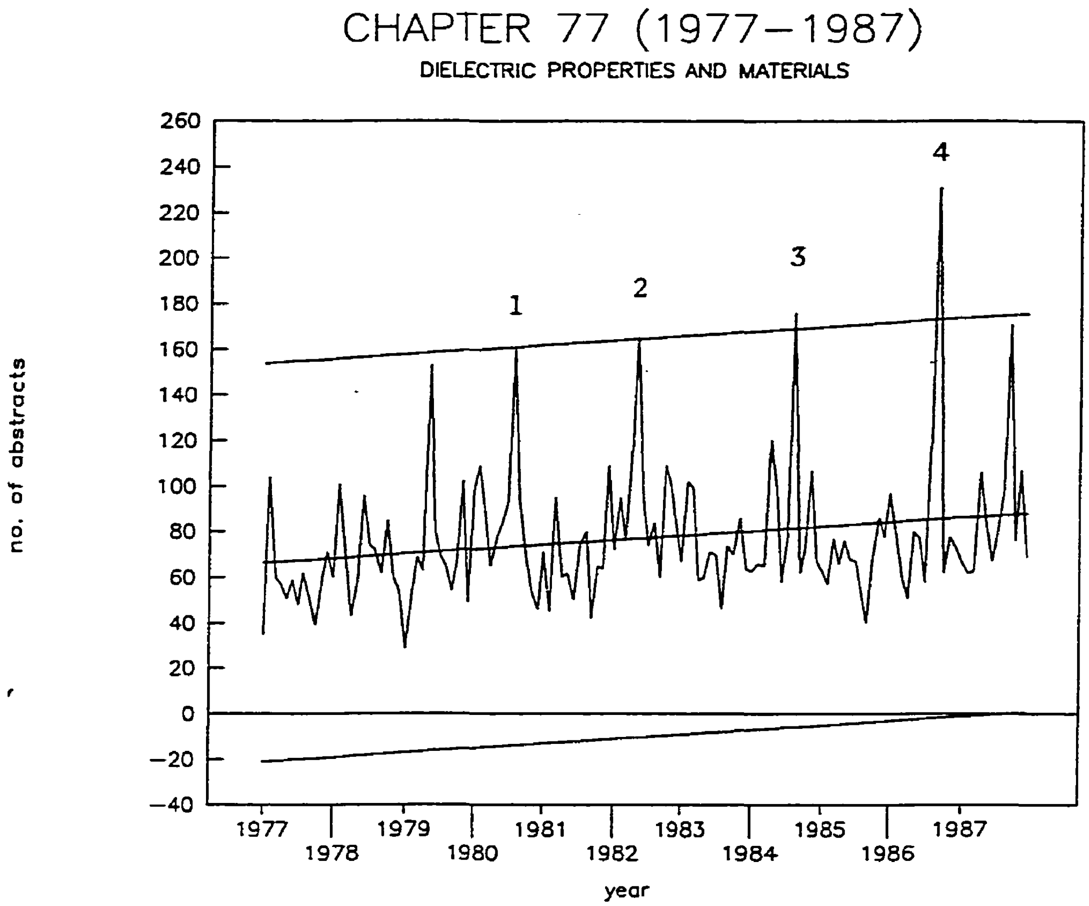


Figure A46. Chapter 77.

Outlier 1 of August 1980 contains 160 abstracts, including the *European Meeting on Ferroelectricity*, 4th, Portoroz, Yugoslavia, September 1979 (77 notices), accounting for 48.1% of the contents.

Outlier 2 of May 1982 contains 165 abstracts, including the *International Meeting on Ferroelectricity*, 5th, University Park, PA, August 1981 (93 notices). Papers from all conferences account for 56.4% of the contents.

Figure A46.--Continued.

Outlier 3 of August 1984 contains 176 abstracts, including the *European Meeting on Ferroelectricity*, 5th, Malaga, Spain, September 1983 (75 notices). Papers from all conferences account for 42.6% of the contents.

Outlier 4 of September 1986 contains 231 abstracts, including the *International Meeting on Ferroelectricity*, 6th, Kobe, Japan, August 1985 (134 notices) and the *International Symposium on Electrets*, 5th, Heidelberg, September 1985 (35 notices). Papers from all conferences account for 73.2% of the contents.

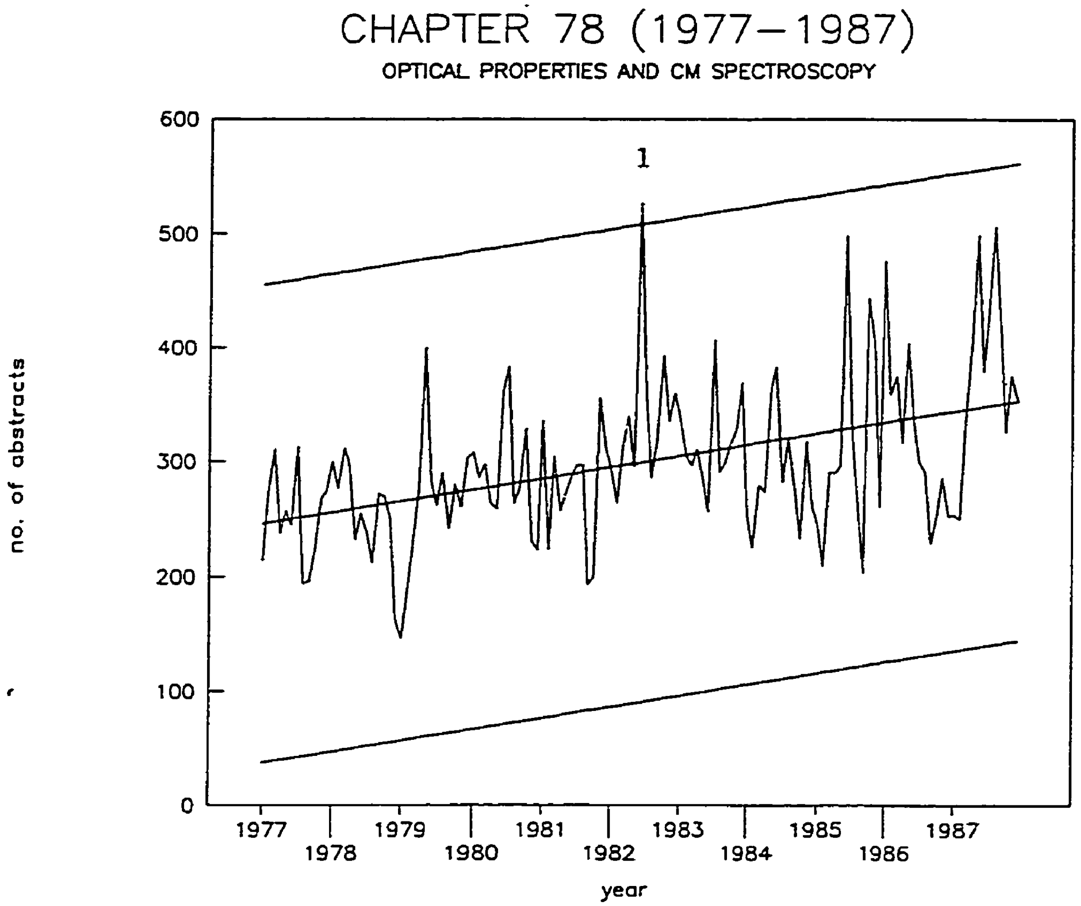


Figure A47. Chapter 78.

Outlier 1 of June 1982 contains 527 abstracts, including the following:
- *International Conference on Amorphous and Liquid Semiconductors*, 9th, Grenoble, July 1981 (18 notices),
- *International Conference on Luminescence*, Berlin, July 1981 (127 notices),
- *International Conference on Phonon Physics*, Bloomington, IN, August-September 1981 (47 notices), and
- *International Meeting on Ferroelectricity*, 5th, University Park, PA, August 1981 (10 notices).

Papers from all conferences account for 38.3% of the contents.

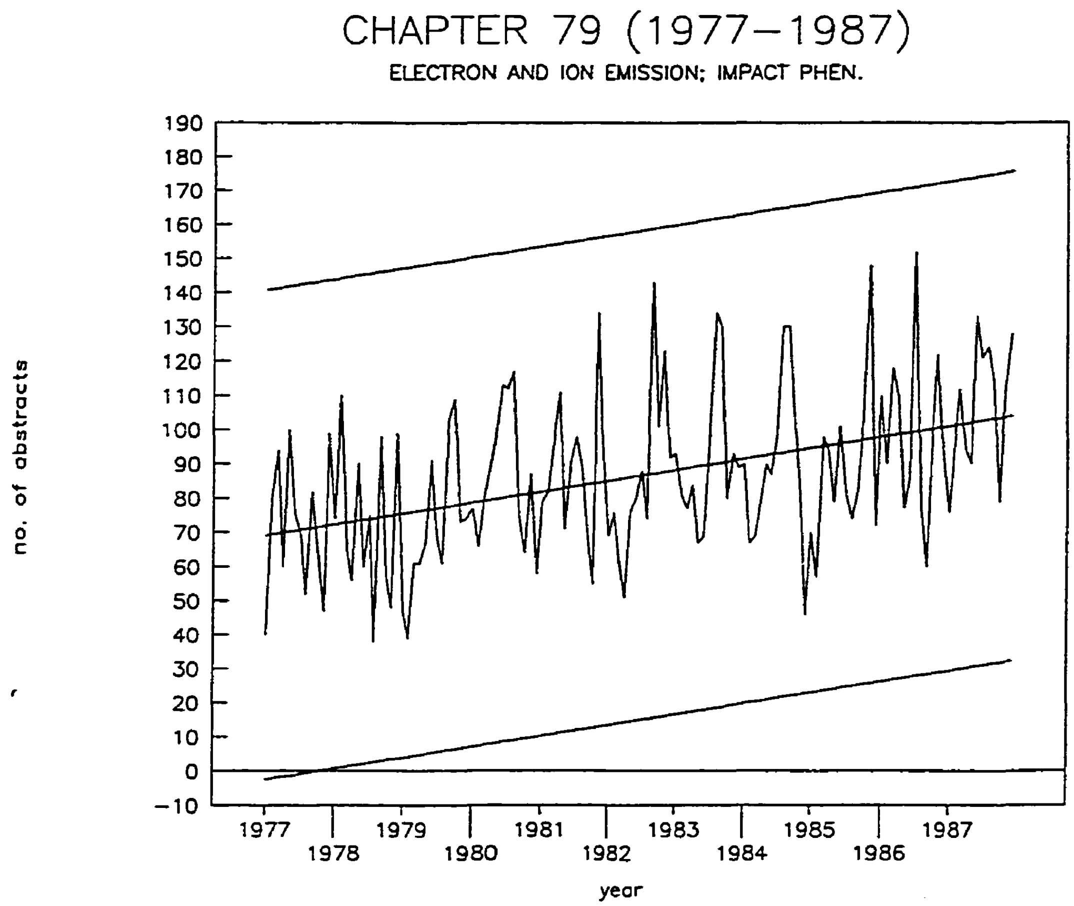


Figure A48. Chapter 79.

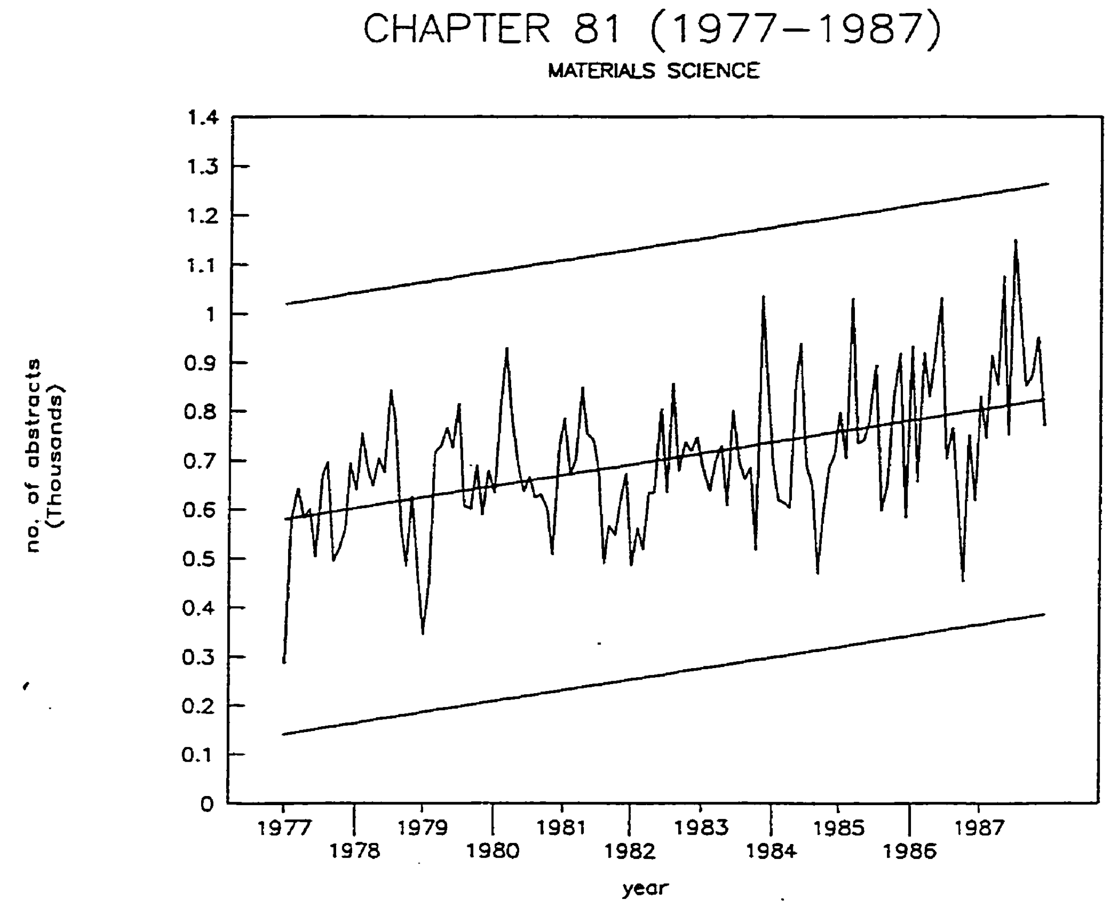


Figure A49. Chapter 81.

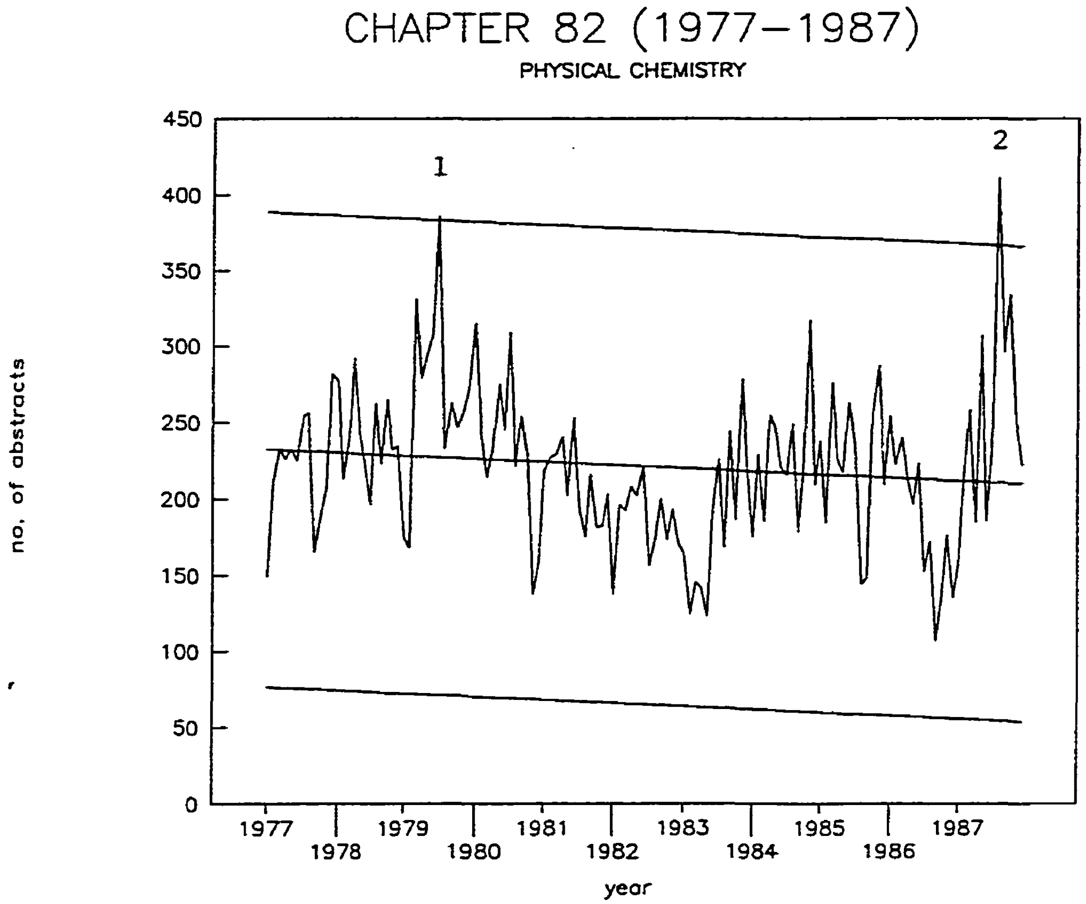


Figure A50. Chapter 82.

Outlier 1 of September 1979 contains 386 abstracts, including the *Conference on the Applications of Small Accelerators in Research and Industry*, Denton, TX, November 1978 (14 notices), the *Informal Conference on Photochemistry*, 12th, Gaithersburg, MD, June-July 1976 (29 notices), and the *International Topical Meeting on Muon Spin Rotation*, 1st, Rorschach, Switzerland, September 1978 (10 notices).

Papers from all conferences account for 13.7% of the contents.

Figure A50.--Continued.

Outlier 2 of August 1987 contains 411 abstracts, including the following:
- *Colloquium Spectroscopicum Internationale XXIV*, Garmisch-Partenkirchen, September 1985 (33 notices),
- *International Conference on Particle Induced X-ray Emission and Its Analytical Applications*, 4th, Tallahassee, FL, June 1986 (26 notices),
- *International Conference on Resonance Ionization Spectroscopy*, 3rd, Swansea, Wales, September 1986 (17 notices), and
- *International Conference on the Application of Accelerators in Research and Industry, Denton*, TX, November 1986 (17 notices).

Papers from all conferences account for 22.6% of the contents.

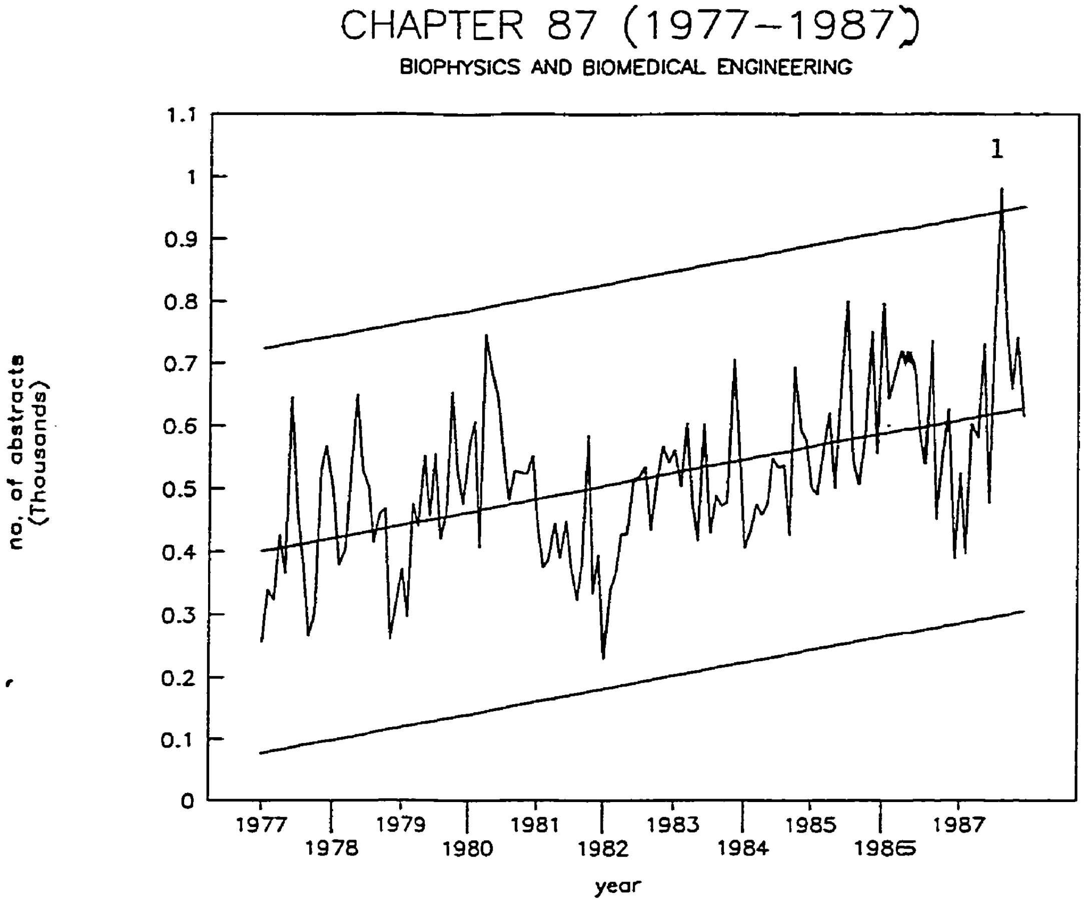


Figure A51. Chapter 87.

The outlier of August 1987 contains 982 abstracts, including the *Annual Conference of the IEEE Engineering in Biology and Medicine Society*, 8th, Fort Worth, TX, November 1986 (149 notices) and the *International Conference on Solid State Dosimetry*, 8th, Oxford, Eng., August 1986 (79 notices). Papers from all conferences account for 23.2% of the contents.

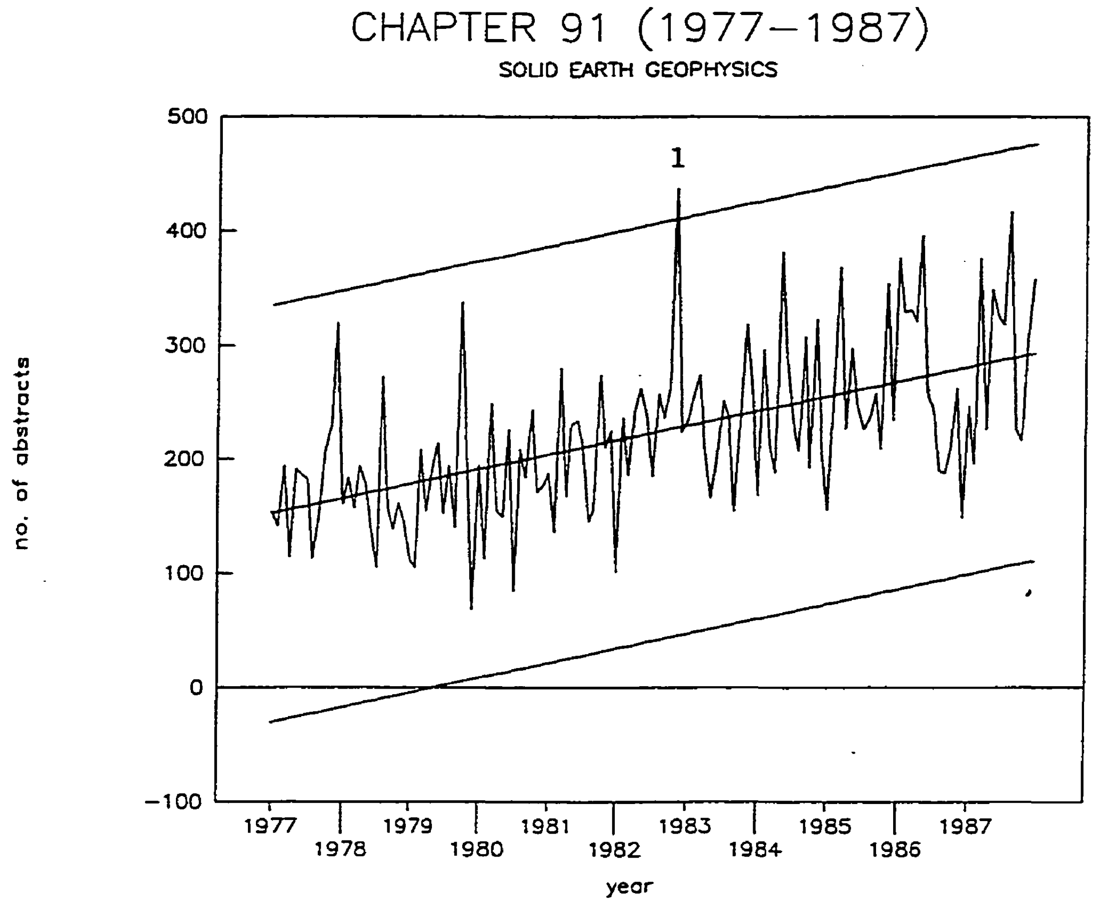


Figure A52. Chapter 91.

Outlier 1 of November 1982 contains 437 abstracts, including papers on MAGSAT published in *Geophysical Research Letters*, v.9, no.4, April 1982 (23 notices), papers on the eruption of Soufriere Volcano in St. Vincent, 1979 published in *Science*, v.216, no.4550, 4 June 1982 (11 notices) and the *Symposium on Properties of Materials at High Pressures and High Temperatures*, London, Ontario, July 1981 (12 notices). Papers from these three publications account for 10.5% of the contents.

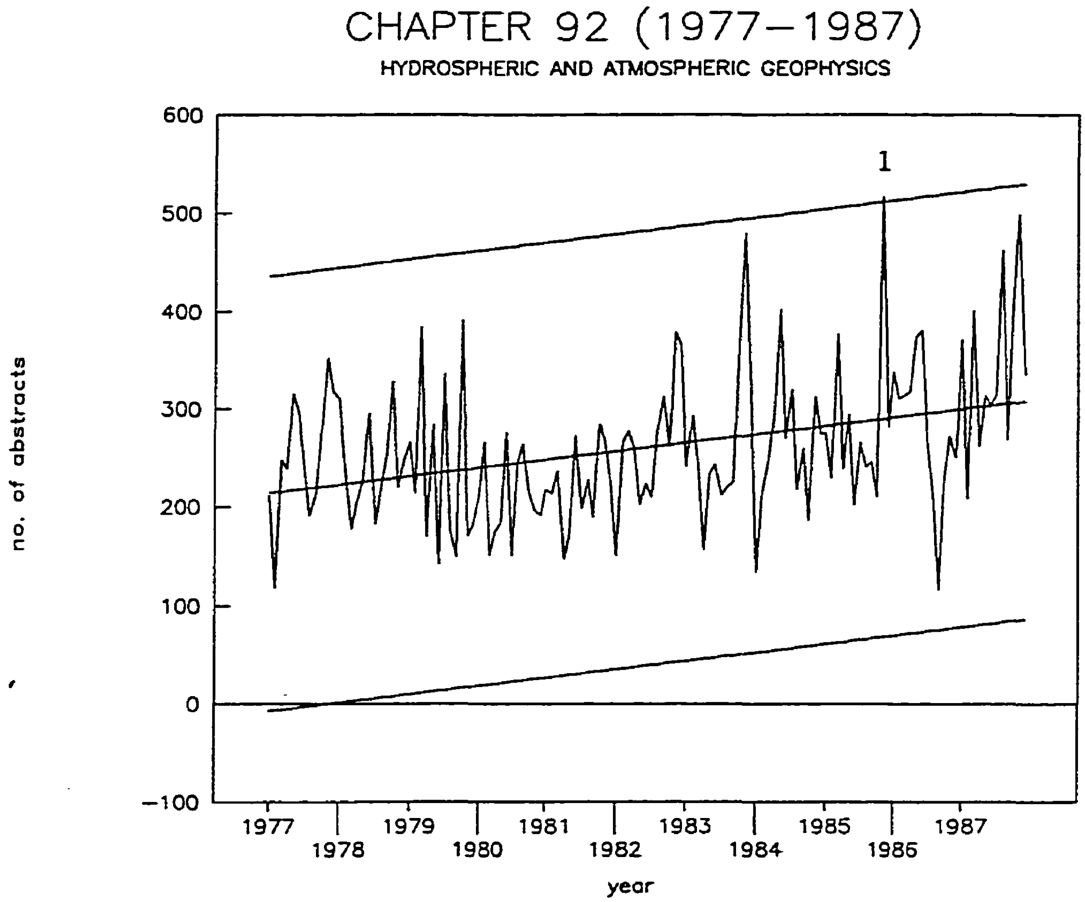


Figure A53. Chapter 92.

Outlier 1 of November 1985 contains 517 abstracts, including the *International Atmospheric Electricity Conference*, 7th, Albany, NY, June 1984 (39 notices) and papers from two issues of the *Journal of Geophysical Research*, v.90, no.C4, (32 notices) and no.D3, 1985 (20 notices).

Papers from these three publications account for 17.6% of the contents.

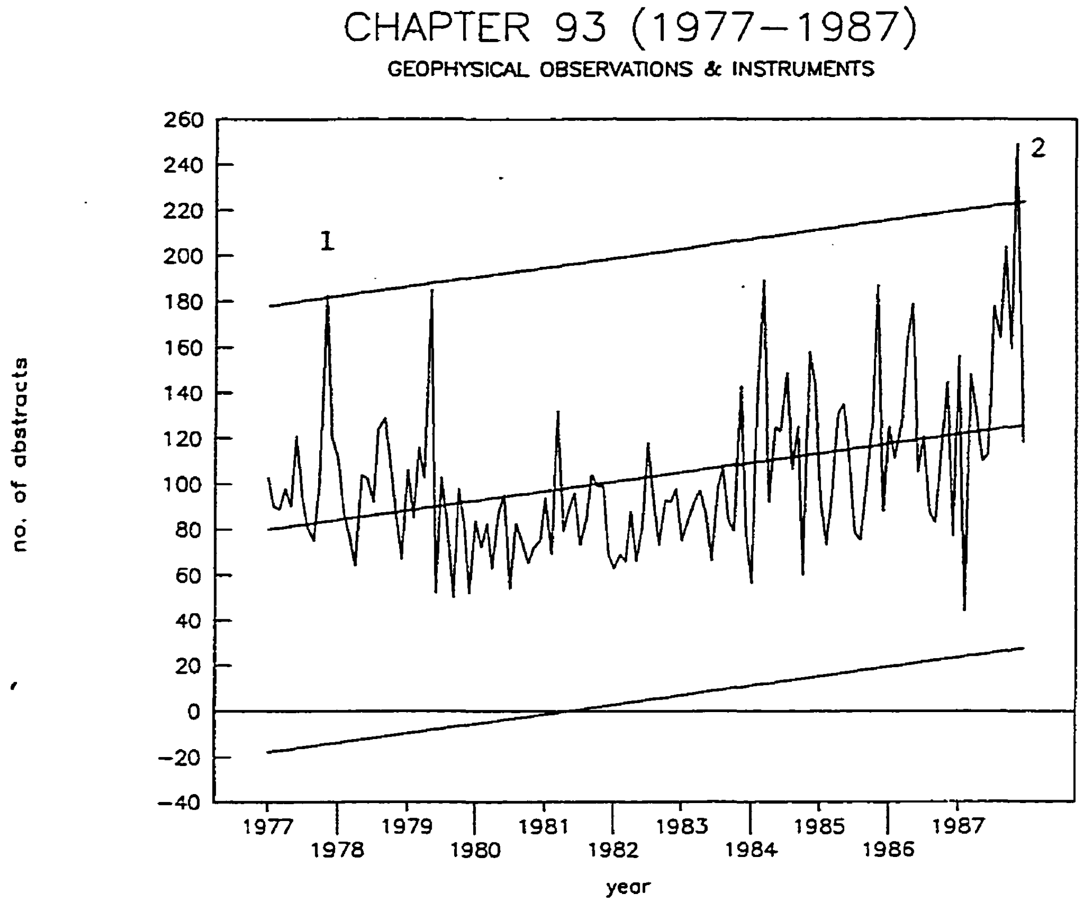


Figure A54. Chapter 93.

Outlier 1 of November 1977 contains 183 abstracts, including the *Annual Symposium on Machine Processing of Remotely Sensed Data*, 4th, West Lafayette, IN, June 1977 (58 notices) and the *International Congress on Electronics*, 24th, Rome, March 1977 (16 notices). Papers from all conferences account for 40.4% of the contents.

Outlier 2 of November 1987 contains 249 abstracts, including papers from the *Journal of Atmospheric and Oceanic Technology*, vol.2 nos. 1, 2, 3, and 4 and vol.3, nos.1 and 2, 1985 (81 notices). Papers from all issues account for 32.5% of the contents. For possible reasons see Section 6.2 of the text.

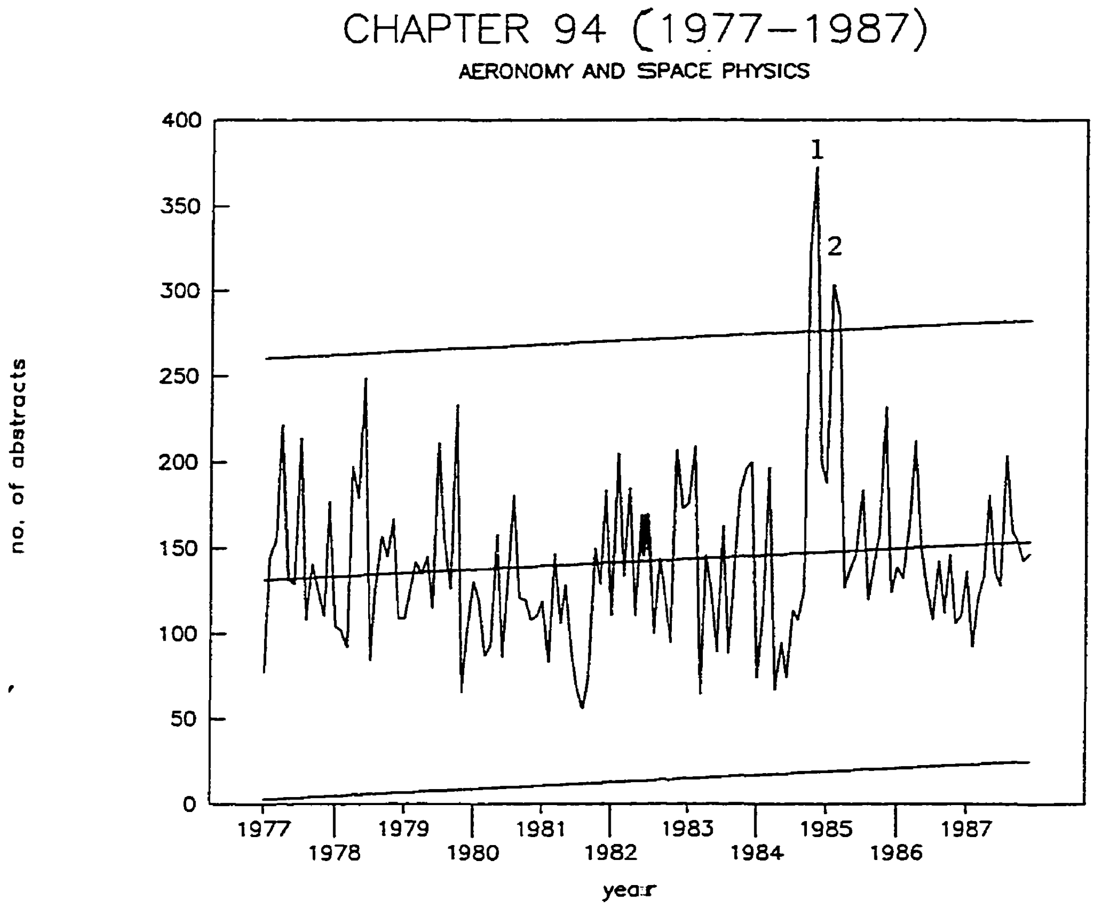


Figure A55. Chapter 94.

Outlier 1 of October 1984 contains 321 abstracts, including the *International Cosmic Ray Conference*, 18th, Bangalore, India, August-September 1983 (194 notices) and the *NATO Advanced Study Institute on Composition and Origin of Cosmic Rays*, Erice, Italy, June 1982 (14 notices). Papers from all conferences account for 64.8% of the contents.

Figure A55.--Continued.

Outlier 1 also covers November 1984 and contains 373 abstracts, including the *International Cosmic Ray Conference*, 18th, Bangalore, India, August-September 1983 (244 notices) and *EISCAT Science: Results from the First Year's Operation of the European Incoherent Scatter Radar*: Papers from the EISCAT Workshop, Assois, France, September 1983 (14 notices). Papers from these two conferences account for 69.2% of the contents.

Outlier 2 covers February and March 1985, containing 304 and 225 abstracts, respectively, all from the *International Cosmic Ray Conference*, 18th, Bangalore, India, August-September 1983 with 207 and 134 notices, respectively. These papers account for 68.1% of the abstract contents for February 1985 and 59.6% of the abstract contents for March 1985.

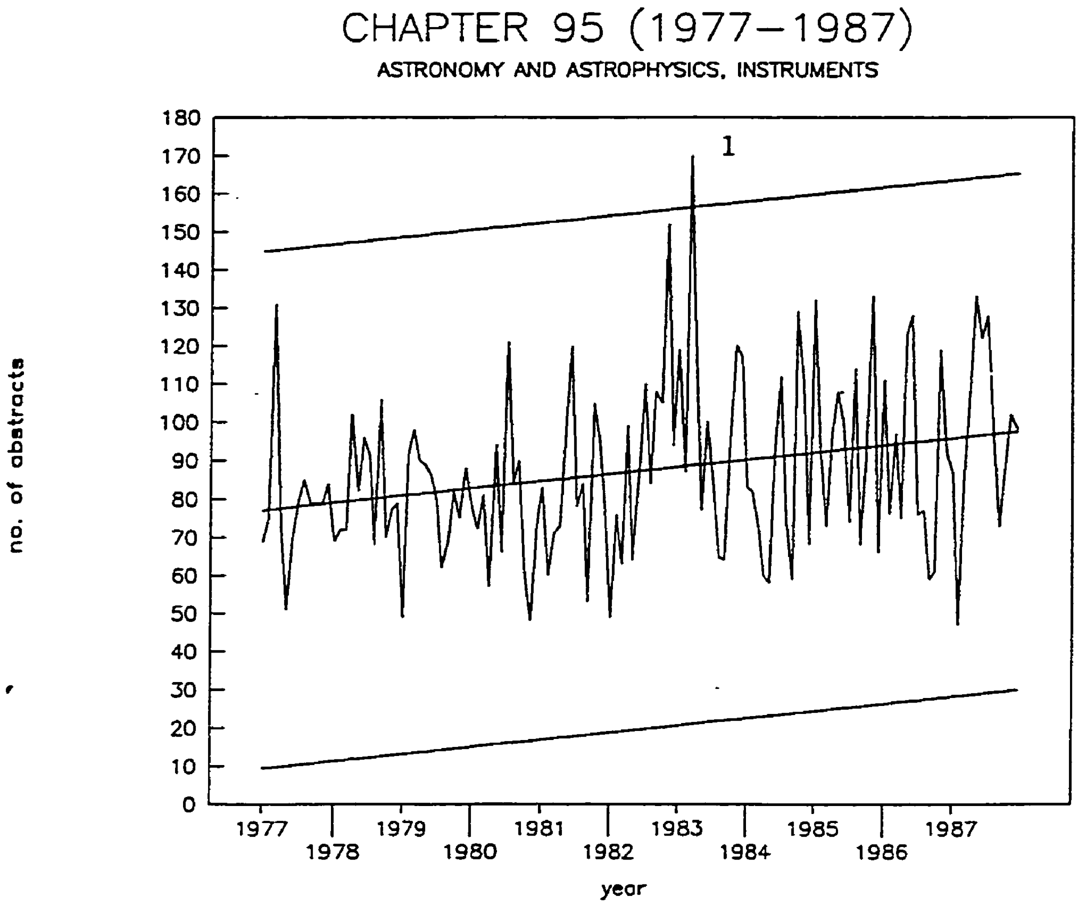


Figure A56. Chapter 95.

Outlier 1 of March 1983 contains 170 abstracts, including the *IAU Colloquium on Instrumentation for Astronomy with Large Optical Telescopes*, no.67, Zelenchukskaya, USSR, September 1981 (26 notices) and the *Oberwolfach Conference on Mathematical Methods in Celestial Mechanics*, 7th, August 1981 (20 notices). Papers from all conferences account for 27.1% of the contents.

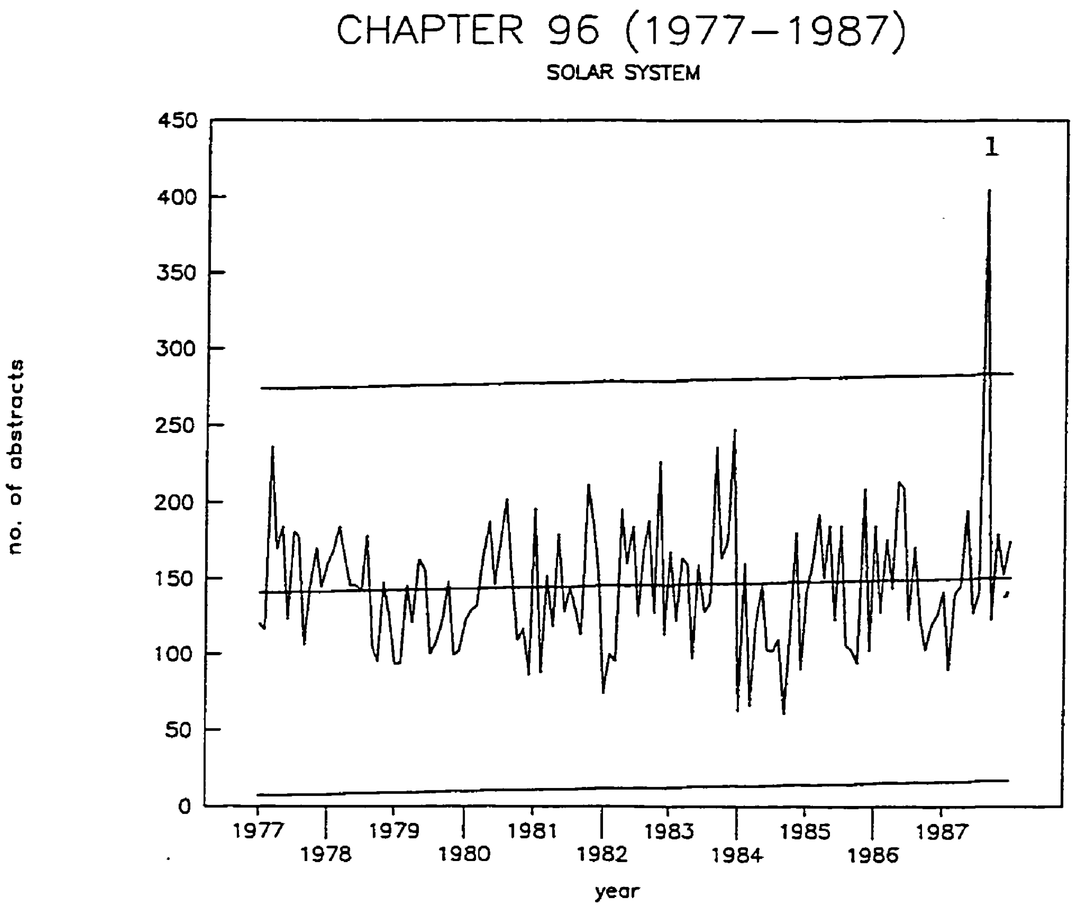


Figure A57. Chapter 96.

Outlier 1 of August 1987 contains 405 abstracts, including the 20th *ESLAB Symposium on the Exploration of Halley's Comet*, Heidelberg, October 1986 (232 notices). Papers from all conferences account for 57.3% of the contents.

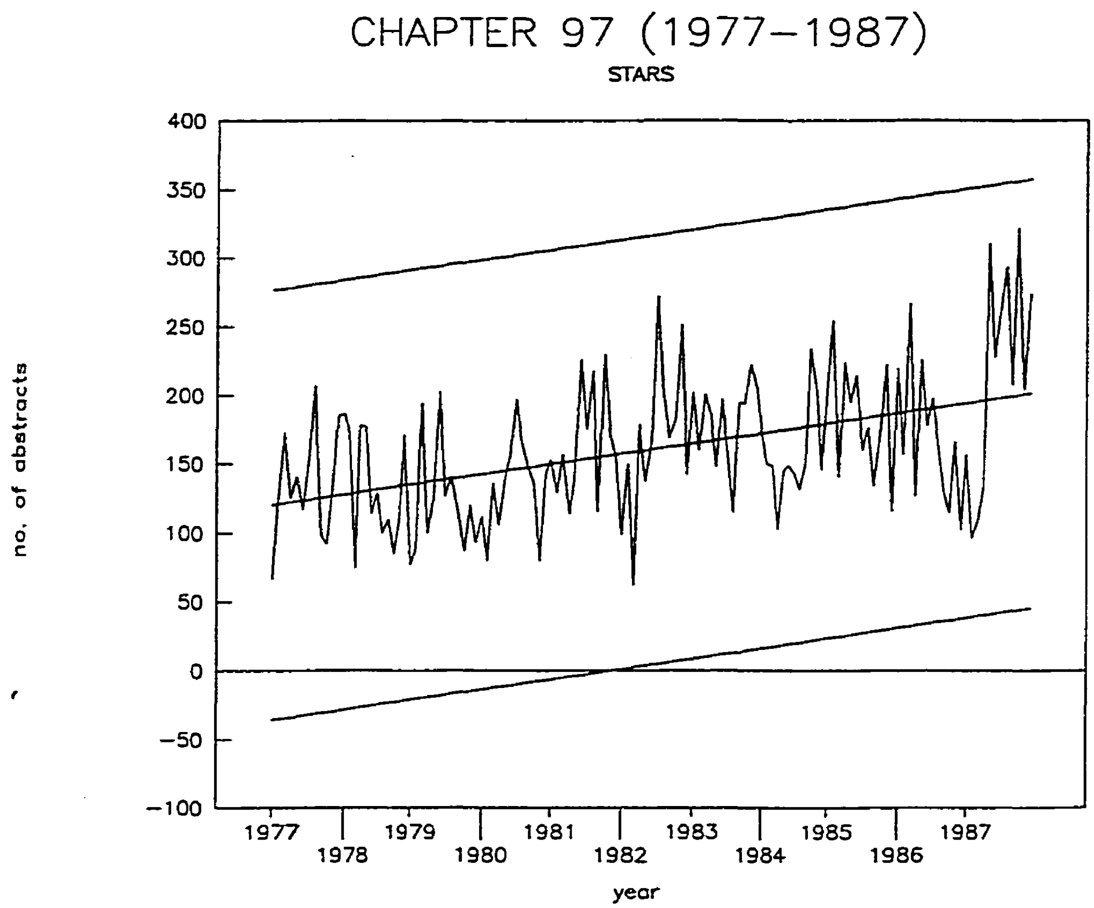


Figure A58. Chapter 97.

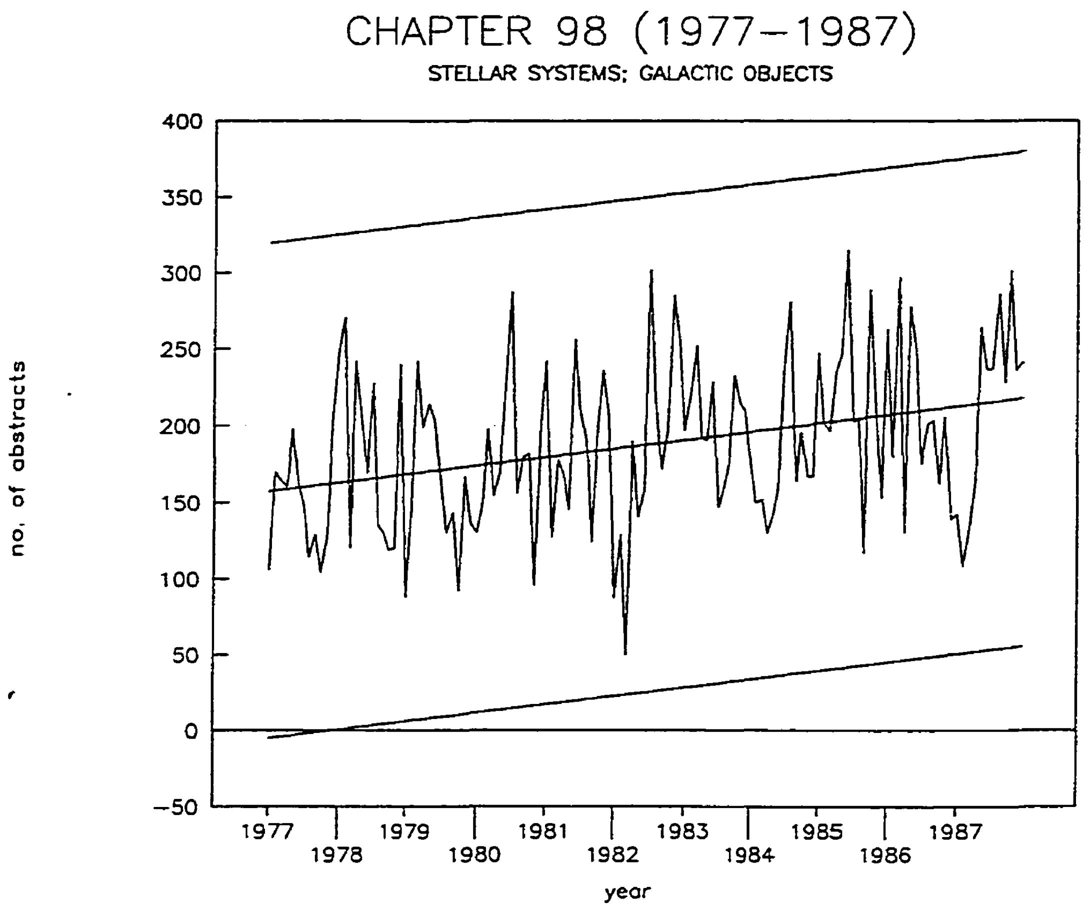


Figure A59. Chapter 98.

APPENDIX B

SUMMARY STATISTICS OF *PHYSICS ABSTRACTS* CHAPTERS 1977-1987

Table B1
Regression Statistics on *Physics Abstracts* Chapters (1977-1987)

Chapter no.	R^2	Significance test F	t	Regression coeff. b0	b1	Stand. Dev.
TOTAL	0.49	124.66	0.0000	7454.87	32.03	1751.07
01	0.49	127.34	0.0000	61.29	0.86	47.00
02	0.09	13.29	0.0004	16.69	0.35	43.77
03	0.35	71.28	0.0000	102.97	0.57	36.92
04	0.30	55.64	0.0000	48.57	0.40	27.67
05	0.62	209.99	0.0000	58.50	0.87	42.11
06	0.01	1.69	0.1957	60.49	0.06	19.73
07	0.01	0.87	0.3524	263.58	0.12	54.72
11	0.37	75.11	0.0000	88.34	0.59	37.30
12	0.28	51.74	0.0000	89.59	0.49	35.46
13	0.00	0.50	0.4823	120.79	-0.05	30.85
21	0.02	2.67	0.1049	94.66	0.12	33.52
23	0.01	1.86	0.1753	38.81	-0.04	13.79
24	0.00	0.06	0.8126	34.67	0.01	15.27
25	0.01	0.94	0.3331	131.45	0.11	49.98
28	0.26	44.55	0.0000	241.47	2.06	156.27
29	0.10	15.16	0.0002	123.74	0.79	93.64
31	0.42	93.28	0.0000	75.50	0.63	37.44
32	0.08	10.93	0.0012	64.28	0.14	19.27
33	0.40	85.32	0.0000	149.76	1.08	65.73
34	0.07	9.80	0.0022	107.86	0.31	44.70
35	0.16	24.43	0.0000	31.94	-0.10	9.82
36	0.30	55.35	0.0000	16.14	0.18	12.41
41	0.15	22.80	0.0000	45.10	0.18	17.57
42	0.32	61.99	0.0000	384.67	1.92	129.17
43	0.07	9.56	0.0024	108.82	0.30	44.37
44	0.02	2.31	0.1308	41.27	-0.06	17.44
46	0.19	30.07	0.0000	151.26	0.62	54.60
47	0.17	26.93	0.0000	256.01	0.85	78.68
51	0.04	4.92	0.0283	22.93	0.04	8.79
52	0.04	5.18	0.0245	244.35	0.40	78.16

Table B1--Continued

Chapter no.	R^2	Significance test F	t	Regression coeff. b0	b1	Stand. Dev.
61	0.49	123.91	0.0000	320.06	2.18	119.55
62	0.56	163.18	0.0000	31.35	0.53	27.03
63	0.02	2.93	0.0894	46.95	-0.07	16.86
64	0.66	255.88	0.0000	56.40	0.87	41.04
65	0.29	54.28	0.0000	22.52	0.16	11.03
66	0.30	54.56	0.0000	64.26	0.39	27.13
67	0.00	0.20	0.6515	34.55	0.02	22.54
68	0.58	182.64	0.0000	93.37	1.35	67.68
71	0.47	115.45	0.0000	110.91	0.95	52.99
72	0.15	23.58	0.0000	160.83	0.49	47.50
73	0.36	74.72	0.0000	92.15	0.67	42.72
74	0.12	17.98	0.0000	54.97	0.44	48.60
75	0.05	7.28	0.0079	190.45	0.51	85.27
76	0.00	0.64	0.4243	145.22	-0.07	39.99
77	0.05	6.51	0.0119	66.17	0.17	29.14
78	0.20	32.55	0.0000	245.71	0.81	69.59
79	0.18	28.84	0.0000	68.90	0.27	23.86
81	0.24	40.23	0.0000	578.84	1.86	146.27
82	0.02	2.16	0.1438	232.62	-0.17	52.00
87	0.27	46.91	0.0000	398.81	1.74	129.22
91	0.32	60.40	0.0000	151.35	1.08	73.06
92	0.13	18.91	0.0000	213.66	0.71	76.39
93	0.14	21.86	0.0000	79.66	0.35	35.00
94	0.02	2.07	0.1524	131.34	0.17	51.54
95	0.07	9.76	0.0022	77.00	0.16	22.55
96	0.00	0.58	0.4458	140.32	0.08	44.40
97	0.20	33.49	0.0000	119.78	0.62	52.05
98	0.11	15.59	0.0001	156.84	0.46	54.07

Table B2
Positive Spikes and Outliers in *Physics Abstracts* (1977-1987)[a]

Chapter	2.0-2.49	2.5-2.99	3.0-3.99	4.0-4.99	5.0+
TOTAL	1 83	3 107,125,128			
A01	1 43	2 41,89			
A02	2 10,124	1 129		2 126,128	1 127
A03	1 40	1 107	1 71		
A04	1 42			2 102,127	
A05	2 102,127	1 114	2 83,125		
A06	3 51,68,131	1 29	1 107		
A07	2 46,125				
A11	1 74		1 71		
A12	2 58,71				
A13	3 18,64,88	2 101,126			
A21		1 26	1 131	1 130	
A23	3 45,64,85	1 26	2 4,125		
A24	1 80	1 58	1 55		
A25	2 89,125	3 26,120,13	2 4,131		
A28	1 39		2 78,96		
A29	2 107,112		4 12,33,58,113	1 83	
A31	4 26,29,49,130	1 128			
A32	4 3,75,125,131		1 114		
A33	3 69,79,106	1 116	1 128		
A34	3 87,107,117	1 85	1 115		1 86
A35	3 21,31,57	2 58,59			
A36	2 92,130		2 107,125	1 122	

[a]Due to space considerations all spikes (and outliers) are indicated as points rather than months and years. There are a total of 132 points in the data, for 11 years of 12 months each. They are given for illustrative purposes only. All outliers analyzed in the text are indicated by their month and year of occurrence.

Table B2--Continued

Chapter	2.0-2.49	2.5-2.99	3.0-3.99	4.0-4.99	5.0+
A41	3 30,90,125	1 132			
A42	3 43,107,125	1 131			
A43	4 62,66,74,131	1 125	1 109		
A44	2 25,29				1 24
A46	3 27,107,127	1 10	1 29		
A47	3 29,34,99	1 24	1 125		
A51	3 66,93,100		1 50	1 71	
A52	2 100,106		3 21,41,101		
A61	3 71,107,128	1 129			
A62	1 43	1 114		1 66	
A63					2 22,66
A64		2 100,129	1 83		
A65	4 55,83,125,129				
A66		2 83,112	3 48,63,114		
A67	1 104			1 65	2 31,105
A68		1 123	1 107		
A71	2 54,125	1 129	1 106		
A72	5 66,70,113,128,129	1 125	1 109		
A73	1 42	1 125	1 107		
A74	2 31,65		3 105,129,131		1 130
A75	3 9,75,113		2 78,114		
A76	2 38,130	1 101	1 41		
A77		2 29,129	3 44,65,92		1 117
A78	3 29,109,125	2 102,128	1 66		
A79	4 59,80,107,115	1 69			
A81	4 39,83,99,125	1 127			
A82	1 130		2 31,128		
A87	2 6,103	1 40	1 128		

Table B2--Continued

Chapter	2.0-2.49	2.5-2.99	3.0-3.99	4.0-4.99	5.0+
A91	4 34,89,113,128	1 12	1 71		
A92	3 27,34,128	2 83,131	1 107		
A93	3 87,107,129	1 29	2 11,131		
A94	1 18	1 99	2 94,98	1 95	
A95	1 3	1 71	1 75		
A96	3 3,81,84				1 128
A97	3 67,125,128	1 130			
A98	4 14,43,67,102				
SUM	129	49	57	11	10

Table B3
Negative Spikes and Outliers in *Physics Abstracts* (1977-1987)[b]

Chapter	2.0-2.49		2.5-2.99		3.0-3.99	
TOTAL	2	61,118	1	117		
A01	1	118				
A02						
A03	1	68	1	120		
A04	1	86				
A05	1	47	1	117		
A06						
A07	2	1,9				
A11						
A12	4	53,61,97,117				
A13	2	47,59				
A21						
A23						
A24						
A25						
A28						
A29						
A31	1	96	1	94		
A32	1	88				
A33	1	117			1	121
A34						
A35						
A36						
A41	1	117				
A42	1	118	1	117		
A43						
A44						
A46	1	54				
A47						
A51	2	91,94				
A52						

[b]There were no negative spikes that went beyond -3.99 standard deviations.

Table B3--Continued

Chapter	2.0-2.49		2.5-2.99		3.0-3.99	
A61	2	117,118				
A62	2	104,117				
A63						
A64					1	118
A65	1	118			1	124
A66	2	117,118				
A67						
A68			1	86		
A71	2	108,118				
A72	1	25				
A73						
A74						
A75						
A76						
A77						
A78	1	105				
A79	1	96				
A81	3	1,25,93	1	118		
A82	1	117				
A87	1	61				
A91	2	36,120				
A92			1	117		
A93	1	122				
A94						
A95	1	122				
A96						
A97	2	63,122				
A98	1	122	1	63		
SUM	46		9		3	

APPENDIX C

COPIES OF BLUESHEETS OF THE *INSPEC* AND *SCIENCE CITATION INDEX* DATABASES ON DIALOG

INSPEC

ONTAP® INSPEC (213)

DIALOG Information Retrieval Service

FILE DESCRIPTION

INSPEC (The Database for Physics, Electronics and Computing) corresponds to the three Science Abstracts print publications: *Physics Abstracts*, *Electrical and Electronics Abstracts*, and *Computer and Control Abstracts*. The Science Abstracts family of abstract journals began publication in 1898. Approximately 16 percent of the database's source publications are in languages other than English, but all articles are abstracted and indexed in English. Author-prepared abstracts are used when available.

The INSPEC database utilizes controlled vocabulary from the *INSPEC Thesaurus*. A single classification scheme is used for all records from 1969 to date. The special DIALOG online thesaurus feature is available to assist searchers in determining appropriate subject terms and codes. Beginning in January 1987, INSPEC records also include chemical substance indexing and numerical index terms.

All INSPEC records from 1969 to the present are contained in File 2. File 3 contains records from 1969 to 1982; File 4 contains records from 1983 to the present. File 213 is available for **ON**line **T**raining **A**nd **P**ractice and contains INSPEC records from January and February 1989.

SUBJECT COVERAGE

Physics (Subfile A) includes:

- Acoustics
- Astronomy and Astrophysics
- Atomic and Molecular Physics
- Biophysics and Medical Physics
- Elementary Particle Physics
- Energy Research
- Environmental Science
- Gases, Fluid Dynamics, and Plasmas
- Geophysics
- Instrumentation and Measurement
- Materials Science
- Mathematics and Mathematical Physics
- Nuclear Physics
- Optics (including Lasers)
- Physical Chemistry
- Properties of Matter
- Quantum Mechanics
- Thermodynamics

Electrical Engineering & Electronics (Subfile B) includes:

- Circuits and Components
- Electricity Generation and Supply
- Electromagnetic Fields and Waves
- Electronic Devices and Materials
- Electronic Instrumentation
- Optics and Electro-optics
- Power Systems and Applications
- Radar and Radionavigation
- Telecommunications

Computers & Control (Subfile C) includes:

- Computational Mathematics
- Computer Applications
- Computer Hardware
- Computer Software
- Control Applications
- Control Systems
- Information Science
- Systems and Control Theory

Information Technology (Subfile D) includes:

- Business and Financial Applications
- Communications, Computing and Systems
- Engineering and Industry Applications
- Management
- Office Automation

SOURCES

As of November 1990, over 4,100 journals and serials are scanned, of which 750 are abstracted cover-to-cover. These constitute 82% of the records in the database, including 6% that represent conference papers published in journals. A further 16% comes from conference proceedings. Other source materials include books, reports, and dissertations. Prior to 1976, a small number of patents were covered.

DIALOG FILE DATA

	File 2	File 3	File 4	File 213
Inclusive Dates:	1969 to present	1969 to 1982	1983 to present	Jan.–Feb. 1989
Update Frequency:	Twice a month (approximately 11,000 records per update)	Closed file	Twice a month (approximately 11,000 records per update)	Not applicable — special file
File Size:	Approximately 3,750,000 records as of November 1990	1,959,518 records	Approximately 1,750,000 records as of November 1990	37,667 records

ORIGIN

INSPEC is produced by the Institution of Electrical Engineers (IEE). Questions concerning file content should be directed to either of the following offices:

North, Central & South America
INSPEC Marketing Department, IEEE Service Center
445 Hoes Lane, P.O. Box 1331
Piscataway, NJ 08855-1331, U.S.A.
Telephone: 908/562-5549
Fax: 908/981-0027
DIALMAIL: 11450

U.K. & Rest of World
INSPEC Marketing Department, IEE
Michael Faraday House, Six Hills Way
Stevenage, Herts SG1 2AY, United Kingdom
Telephone: 0438 313311 Help Desk: 0438 742857
Fax: 0438 742840
DIALMAIL: 11472

INSPEC

BLUESHEET Summary Box

Five major search fields:

FIELD NAME	SUFFIX/PREFIX
Title	/TI
Descriptor	/DE
Classification Code	CC=
Journal Name	JN=
Author	AU=

LIMITS:
/ENG, /NONENG, /ART, /NAR, /PHYS, /TECH (see page 2-7)

SPECIAL FEATURES:
Chemical Indexing (see page 2-5)
Numerical Indexing (see page 2-5)

OUTPUT FORMATS displaying full record:
Formats 5 and 9
Format 4 (Tagged)

INSPEC

SAMPLE RECORD

DIALOG Accession Number

```
        3603153  A90052690, B90029589                                        AZ=
          Title: A 970 nm strained-layer InGaAs/GaAlAs quantum well laser for   /TI
        pumping an erbium-doped optical fiber amplifier
AU=       Author(s): Ming C. Wu; Olsson, N.A.; Sivco, D.; Cho, A.Y.
CS=       Author Affil: AT&T Bell Labs., Murray Hill, NJ, USA
JN=, SO=  Journal: Applied Physics Letters   vol.56, no.3,   p.: 221-3
PY=       Publication Date: 15 Jan. 1990   Country of Publication: USA       CP=
CO=       CODEN: APPLAB   ISSN: 0003-6951                                      SN=
          U.S. Copyright Clearance Center Code: 0003-6951/90/030221-03$02.00
TC=       Treatment: Experimental (X)
LA=       Language: English   Document Type: Journal Paper (JP)            DT=
          (18 Refs)
          Abstract: The authors report the performance of a 970 nm strained-layer   /AB
        InGaAs/GaAlAs quantum well laser and its application for pumping Er-doped
        optical fiber amplifiers. The laser was grown by molecular beam epitaxy and
        has three In/sub 0.2/Ga/sub 0.8/As/GaAs quantum wells. For a 5- mu m-wide
        and 400- mu m-long ridge-waveguide laser, a CW threshold current of 20 mA
        and an external quantum efficiency of 0.28 mW/mA per facet were obtained.
        Maximum output power exceeds 32 mW/facet. With antireflection coating, even
        higher external quantum efficiency (0.40 mW/mA) was achieved, and more than
        20 mW of power was coupled into a single mode fiber. Preliminary
        experiments of pumping the Er-doped fiber amplifier gave 15 dB of gain at
        1.555 mu m for a pump power of 14 mW into the Er fiber.
          Descriptors: erbium; fibre optics; gallium compounds; gradient index   /DE
        optics; indium compounds; optical fibres; optical pumping; optical
        waveguides; semiconductor junction lasers; solid lasers
          Identifiers: GRIN-SCH laser; strained-layer; quantum well laser; pumping;   /ID
        optical fiber amplifier; Er-doped optical fiber; molecular beam epitaxy;
        ridge-waveguide laser; CW threshold current; external quantum efficiency;
        output power; antireflection coating; single mode fiber; gain; 970 nm;
        1.555 micron; 5 micron; 14 mW; 400 micron; 20 mA; In/sub 0.2/Ga/sub
        0.8/As-GaAs; InGaAs-GaAlAs
          Chemical Indexing:
CI=       In0.2Ga0.8As-GaAs int - In0.2Ga0.8As int - Ga0.8 int - In0.2 int - GaAs
        int - As int - Ga int - In int - In0.2Ga0.8As ss - Ga0.8 ss - In0.2 ss - As
        ss - Ga ss - In ss - GaAs bin - As bin - Ga bin (Elements - 3,2,3)
          Er ss - Er el - Er dop (Elements - 1)                               NE=
          InGaAs-GaAlAs int - GaAlAs int - InGaAs int - Al int - As int - Ga int -
        In int - GaAlAs ss - InGaAs ss - Al ss - As ss - Ga ss - In ss (Elements - 3,3,4)
NI=       Numerical Indexing: wavelength 9.7E-07 m; wavelength 1.555E-06 m; size   WA=
        5.0E-06 m; power 1.4E-02 W; size 4.0E-04 m; current 2.0E-02 A            CU=
PO=       Class Codes: A4260B (Design of specific laser systems); A4255P (Lasing
CC=     action in semiconductors with junctions); A4280L (Optical waveguides and   CN=
SF=     couplers); A4281F (Other optical properties); A4255R (Lasing action in
        other solids); B4320J (Semiconductor junction lasers); B4125 (Fibre
        optics); B4130 (Optical waveguides); B4320G (Solid lasers)
```

INSPEC

SEARCH OPTIONS

BASIC INDEX

SEARCH SUFFIX+	DISPLAY CODE	FIELD NAME	INDEXING	SELECT EXAMPLES
/AB	AB	Abstract	Word	S THRESHOLD(W)CURRENT/AB
/DE	DE	Descriptor[1]	Word & Phrase	S FIBRE?(N)OPTIC?/DE S OPTICAL FIBRES/DE
/ID	ID	Identifier[2]	Word & Phrase	S GAALAS/ID S INGAAS-GAALAS/ID
/TI	TI	Title	Word	S QUANTUM(W)WELL(S)LASER?/TI

+If no suffix is specified *all* Basic Index fields are searched.
[1]Also /DF.
[2]Also /IF.

ADDITIONAL INDEXES

BIBLIOGRAPHIC FIELDS

SEARCH PREFIX	DISPLAY CODE	FIELD NAME	INDEXING	SELECT EXAMPLES
AC=	AC	Patent Application Country[3]	Word	S AC=JP
AD=	AD	Patent Application Date[3]	Phrase	S AD=710402
AN=	AN	Patent Application Number[3]	Phrase	S AN=20162
AZ=	AZ	INSPEC Abstract Number	Phrase	S AZ=A90052690
—	AZ	DIALOG Accession Number		
AU=	AU	Author	Phrase	S AU=OLSSON, N.A.
BN=	BN	International Standard Book Number (ISBN)	Phrase	S BN=0 8186 1986 4 S BN=0818619864
CC=	CC	Classification Code	Phrase	S CC=A4260B S CC=A42
—	CF	Conference Information		
CL=	CL	Conference Location	Word	S CL=(SANTA(W)CLARA)
CN=	CN	Classification Name	Word & Phrase	S CN=(LASING(W)ACTION) S CN=OPTICAL WAVEGUIDES
CO=	CO	CODEN	Phrase	S CO=APPLAB
CP=	CP	Country of Publication	Word & Phrase	S CP=USA S CP=WEST GERMANY
CS=	CS	Corporate Source	Word	S CS=(AT(W)T(S)MURRAY(W)HILL)
CT=	CT	Conference Title	Word	S CT=(COMPUTER(W)AIDED(W)DESIGN)
CY=	CY	Conference Year	Phrase	S CY=1990
DT=	DT	Document Type	Phrase	S DT=JOURNAL PAPER S DT=JP
—	FN	File Name		
JN=	JN	Journal Name	Phrase	S JN=APPLIED PHYSICS LETTERS S JN=APPL. PHYS. LETT. (USA)
LA=	LA	Language	Phrase	S LA=FRENCH
PA=	PA	Patent Assignee[3]	Word & Phrase	S PA=PIONEER S PA=PIONEER ELECTRONIC?
PC=	PC	Patent Country[3]	Word	S PC=GB
PD=	PD	Patent Date[3]	Phrase	S PD=720329
—	PI	Patent Information[3]		
PN=	PN	Patent Number[3]	Phrase	S PN=GB 1379306
PU=	PU	Publisher[4]	Word	S PU=(IEEE AND WASHINGTON)
PY=	PY	Publication Year	Phrase	S PY=1989:1990
RN=	RN	Report or Contract Number	Word & Phrase	S RN=CERN S RN="CERN/SPS/ACC/79-13" S RN=CERNSPSACC7913
SF=	—	Subfile	Phrase	S SF=A
SN=	SN	International Standard Serial Number (ISSN)	Phrase	S SN=0003-6951 S SN=00036951
SO=	SO	Source Information[5]	Word	S SO=(APPL?(W)PHYS?)
SP=	SP	Conference Sponsor	Word	S SP=(ACM AND IEEE)
TC=	TC	Treatment Code	Phrase	S TC=EXPERIMENTAL S TC=X
UD=	—	Update[6]	Phrase	S UD=9001B1:9999

[3]Files 2 and 3 only; dates of patent coverage are 1969-1976.
[4]Available for conference proceedings and books only.
[5]Search field includes journal title words and volume and issue numbers. Display varies depending on document type.
[6]Not available in File 213.

INSPEC

CHEMICAL INDEXING FIELDS (available since January 1987)

SEARCH PREFIX	DISPLAY CODE	FIELD NAME	INDEXING	SELECT EXAMPLES
CI=	CI	Substance, component, or material system (including role modifier)[7]	Word & Phrase	S CI=GAAS S CI=(GA(S)AS(S)INT) S CI=AS S CI=AS BIN S CI=SS
NE=	NE	Number of elements in substance, component, or material system	Phrase	S NE=3(S)CI=(GA(S)AL(S)AS)

[7]Role modifiers include:

EL	(element)	INT	(interface system)
DOP	(dopant)	SUR	(surface or substrate)
BIN	(binary system)	ADS	(adsorbate, or any sorbate, i.e., species being (ad)sorbed onto a substrate)
SS	(system with 3 or more components)		

NUMERICAL INDEXING FIELDS (available since January 1987)

Numeric data for each physical quantity shown on the following page (temperature, pressure, frequency, etc.) are indexed into a separate numeric field (TE=, PR=, FR=, etc.). In the record display, numeric values appear in an exponential floating point format. For example, a frequency of 25 kHz (25000 Hz) is converted to 2.5×10^4 and indexed as FR=2.5E+04.

For searching, the numeric values can be entered in several ways: either directly, without converting to exponential form, or in exponential form. Note that the plus sign ('+') is not required when entering values with positive exponents, and if used, requires masking with quotation marks:

```
      SELECT FR=25000
or    SELECT FR=2.5E4
or    SELECT FR=2.5E04
or    SELECT FR="2.5E+04"
```

Truncation is not allowed when searching numeric data. Range searching is recommended for best results, e.g., S FR=25000:30000. The smallest and largest numbers that may be searched are 5.4E-79 and 7.2E+75.

For specifying precise minimum or maximum numeric values, the LO= and HI= search prefixes may be used. LO= and HI= are generic prefixes not specific to any physical quantity. Searches using LO= and HI= should be qualified with the addition of the desired physical quantity using the NI= prefix. Refer to the table on the following page for the physical quantities available for searching.

SEARCH PREFIX	DISPLAY CODE	FIELD NAME	INDEXING	SELECT EXAMPLES
LO=	NI	Lowest value	Numeric	S LO=100(S)NI=TEMPERATURE S LO>=3.16E7(S)NI=AGE
HI=	NI	Highest value	Numeric	S HI=2.5E4(S)NI=FREQUENCY S HI<=9.7E-7(S)NI=WAVELENGTH

INSPEC

NUMERICAL INDEXING FIELDS (available since January 1987)

SEARCH PREFIX	DISPLAY CODE	PHYSICAL QUANTITY (UNIT OF MEASURE)**	SELECT EXAMPLES
AG=	NI	Age (yr; Year)	S AG>=1E9
AL=	NI	Altitude (m; Meter)	S AL=2E4:9E5
AP=	NI	Apparent Power (VA; Volt-amp)	S AP=3E6 S AP=3000000
BI=	NI	Bit Rate (Bit/s; Bits per Second)	S BI=64000
BW=	NI	Bandwidth (Hz; Hertz)	S BW=5E7
BY=	NI	Byte Rate (Byte/s; Bytes per Second)	S BY=2.5E6
CA=	NI	Capacitance (F; Farad)	S CA=2E-13
CD=	NI	Conductance (S; Siemen)	S CD=2:5
CE=	NI	Computer Execution Rate (IPS; Instructions per Second)	S CE>=1E6 S CE>=1000000
CM=	NI	Computer Speed (FLOPS; Floating Point Operations per Second)	S CM>=3.5E6 S CM>=3500000
CU=	NI	Current (A; Ampere)	S CU=0.051
DI=	NI	Distance (m; Meter)	S DI=0.002
DP=	NI	Depth (m; Meter)	S DP=2E4:9E5
EF=	NI	Efficiency (Percent)	S EF=60
EL=	NI	Electrical Conductivity (S/m; Siemen per Meter)	S EL=7.0E4 S EL=70000
EN=	NI	Energy (J; Joule)	S EN=0.5
ER=	NI	Electrical Resistivity (ohm; Ohm Meter)	S ER=1.7E-4 S ER=0.00017
EV=	NI	Electron Volt Energy (eV; Electron Volt)	S EV=-0.5:0
FR=	NI	Frequency (Hz; Hertz)	S FR=0:1
GA=	NI	Gain (dB; Decibel)	S GA=14
GD=	NI	Galactic Distance (pc; Parsec)	S GD>=1E7
GE=	NI	Geocentric Distance (m; Meter)	S GE=>3.7E10
HD=	NI	Heliocentric Distance (AU; Astronomical Unit)	S HD=5E4 S HD=50000
LS=	NI	Loss (dB; Decibel)	S LS=-60:0
MA=	NI	Mass (kg; Kilogram)	S MA=6E14
MD=	NI	Magnetic Flux Density (T; Tesla)	S MD=1E-2
MS=	NI	Memory Size (Byte)	S MS>=3E7
NF=	NI	Noise Figure (dB; Decibel)	S NF=1:2
PO=	NI	Power (W; Watt)	S PO=4E-5:2E-4
PR=	NI	Pressure (Pa; Pascal)	S PR=1.3E-3
PS=	NI	Printer Speed (cps; Characters per Second)	S PS>=2E2 S PS>=200
PX=	NI	Picture Size (pixel; Picture Element)	S PX=512
RA=	NI	Radiation Absorbed Dose (Gy; Gray)	S RA=2
RD=	NI	Radiation Dose Equivalent (Sv; Sievert)	S RD=1E-6:1E-2 S RD=0.000001:0.01
RE=	NI	Resistance (ohm)	S RE=7E-5:0.1
RP=	NI	Reactive Power (VAr; Volt-Amp Reactive)	S RP=1E5 S RP=100000
RX=	NI	Radiation Exposure (C/kg; Coulomb per Kilogram)	S RX<=0.1 S RX<=1E-1
RY=	NI	Radioactivity (Bq; Becquerel)	S RY=1E8:1E12
SI=	NI	Size (m; Meter)	S SI=0.7:15
SM=	NI	Stellar Mass (Msol; Solar Mass)	S SM=1E-2:3000
SR=	NI	Storage Capacity (Bit)	S SR=4.2E6
TE=	NI	Temperature (K; Kelvin)	S TE=3.26E2 S TE=326
TM=	NI	Time (s; Second)	S TM=2E-11:4E-11
VE=	NI	Velocity (m/s; Meters per Second)	S VE=-5E4:-2E2
VO=	NI	Voltage (V; Volt)	S VO>=1000
WA=	NI	Wavelength (m; Meter)	S WA=8.8E-7:1E-1
WL=	NI	Word Length (Bit)	S WL=32

**Each physical quantity and its corresponding abbreviated unit of measure are optionally searchable using NI=.

INSPEC

LIMITING

Sets and terms may be limited by Basic Index suffixes, i.e., /AB, /DE, /DF, /ID, /IF, /TI (e.g., S S6/TI), as well as by the following features:

SUFFIX	FIELD NAME	EXAMPLES
None	DIALOG Accession Number	S S1/3259712-9999999
None	Publication Year	S SUPERCONDUCTOR?/1989:1990
/ENG	English Language	S S9/ENG
/NONENG	Non-English Language	S LASERS/NONENG
/ART	Journal Article	S S2/ART
/NAR	Non-Journal Article	S AMPLIFIER?/NAR
/PHYS	Physics Subfile	S SEMICONDUCTOR?/PHYS
/TECH	Electronics, Computing, and Information Technology Subfiles	S HOLOGRAPHY/TECH

SORTING

SORTABLE FIELDS	EXAMPLES
Online (SORT) and offline (PRINT[6]) AU, AZ, CC, CS, JN, PY, TI	SORT S6/ALL/JN/PY,D PRINT S3/5/1-24/AU

OUTPUT OPTIONS†

USER-DEFINED FORMAT OPTIONS

User-defined formats may be specified using the display codes indicated in the Search Options tables, e.g., TYPE S3/AU,TI,SO/1-5.

PREDEFINED FORMAT OPTIONS

NUMBER	RECORD CONTENT
Format 1	DIALOG Accession Number
Format 2	Full Record except Abstract
Format 3	Bibliographic Citation
Format 4	Full Record with Tagged Fields
Format 5, 9	Full Record
Format 6	Title and INSPEC Abstract Number
Format 7	Full Record except Indexing
Format 8	Title and Indexing
Format K	KWIC (Key Word In Context) displays a window of text; may be used by itself or with other formats (HILIGHT is also available)

DIRECT RECORD ACCESS

FIELD NAME	EXAMPLES		
DIALOG Accession Number	TYPE 3603153/5	DISPLAY 3603153/TI,CS	PRINT[6] 3603153/3

†TAG may be used for tagged fields, e.g., TYPE S2/AN,SO/1-5 TAG.

SCISEARCH®

ONTAP® SCISEARCH® (294)

Information Retrieval Service

FILE DESCRIPTION

SCISEARCH® is an international, multidisciplinary index to the literature of science, technology, biomedicine, and related disciplines produced by the Institute for Scientific Information® (ISI®). SCISEARCH contains all of the records published in the *Science Citation Index*® (SCI ®), plus additional records from the *Current Contents*® series of publications.

SCISEARCH is distinguished by some important and unique characteristics. First, journals indexed are selected on the basis of several criteria, including citation analysis, resulting in coverage of the most significant publications in the scientific, technical, and biomedical literature. Second, in addition to the more conventional retrieval methods, SCISEARCH offers citation indexing, which permits searching by cited references. For records added since January 1991, author abstracts, author keywords, and KeyWords Plus™ may be searched.

File coverage is as follows: File 434 covers 1974 to the present; File 34 covers 1988 to the present. File 294 is available for **ON**line **T**raining **A**nd **P**ractice and contains SCISEARCH selected records from the beginning of 1991.

SUBJECT COVERAGE

SCISEARCH covers virtually every subject area within the broad fields of science, technology, and biomedicine, including but not limited to:

- Agriculture and Food Technology
- Astronomy and Astrophysics
- Behavioral Sciences
- Biochemistry
- Biology
- Biomedical Sciences
- Chemistry
- Computer Applications and Cybernetics
- Earth Sciences
- Electronics
- Engineering
- Environmental Science
- Genetics
- Instrumentation
- Materials Science
- Mathematics
- Medicine
- Meteorology
- Microbiology
- Nuclear Science
- Pharmacology
- Physics
- Psychiatry and Psychology
- Veterinary Medicine
- Zoology

SOURCES

SCISEARCH indexes all significant items (articles, review papers, meeting abstracts, letters, editorials, book reviews, correction notices, etc.) from approximately 4,500 major scientific and technical journals. Approximately 3,800 of these journals are further indexed by the references cited within each article, allowing for citation searching. The other additional 700 journals indexed have been drawn from ISI Current Contents® series of publications.

DIALOG FILE DATA

	File 434	File 34	File 294
Inclusive Dates:	1974 to the present	1988- to the present	Early 1991
Update Frequency:	Weekly	Weekly (Approximately 14,000 records per update)	Not applicable
File Size:	Over 10.8 million records as of July 1991	Over 2.4 million records of July 1991	30,000 records

ORIGIN

SCISEARCH is produced by the Institute for Scientific Information. Questions concerning file content should be directed to:

ISI Technical Help Desk
Institute for Scientific Information (ISI)
3501 Market Street
Philadelphia, PA 19104
USA
Telephone: 800/336-4474
215/386-0100, ext. 1591
Fax: 215/386-6362
DIALMAIL: 15010

ISI Technical Help Desk
Institute for Scientific Information
132 High Street
Uxbridge, Middlesex UB8 1DP
United Kingdom
Telephone: (+44) 895-70016
Fax: (+44) 895-56710
Telex: 9933693 UKISI

Full copies of publications indexed in the SCISEARCH databases may be ordered online via DIALORDER® using The Genuine Article Number (ORDER ISI). See the ISI Yellowsheet for details.

SCISEARCH®

SAMPLE RECORD

```
DIALOG Accession Number
                                                                              GA=
   10513897   Genuine Article#: EJ220   Number of References: 48 ----------- NR=
   Title: THE EFFECT OF DIETARY FIBER TYPE ON GLYCATED HEMOGLOBIN AND RENAL   /TI
      HYPERTROPHY IN THE ADULT DIABETIC RAT
AU=Author(s): GALLAHER DD; SCHAUBERT DR
CS=Corporate Source: UNIV MINNESOTA,DEPT FOOD SCI & NUTR,1334 ECKLESAVE/ST
      PAUL//MN/55108; N DAKOTA STATE UNIV,DEPT FOOD & NUTR/FARGO//ND/58105 -- PY=
JN=Journal: NUTRITION RESEARCH, 1990, V10, N11, P1311-1323
LA=Language: ENGLISH   Document Type: ARTICLE ------------------------------- DT=
GL=Geographic Location: USA
   Subfile: SciSearch; CC LIFE -- Current Contents, Life Sciences ----------- SF=
   Journal Subject Category: NUTRITION & DIETETICS ------------------------- SC=
   Abstract: The effect of various dietary fiber sources on glycated          /AB
      hemoglobin and renal hypertrophy, two long-term indicators of blood
      glucose control, was studied in diabetic rats.  Streptozotocin-treated
      rats were fed a fiber-free diet or diets containing 8% dietary fiber,
      using one of the following fiber sources:  cellulose, sugar beet fiber,
      beet fiber treated with calcium carbonate, oat bran, rye bran, barley
      bran flour, wheat bran, or guar gum.  After 28 days, only guar
      gum-feeding reduced the % glycated hemoglobin relative to the
      fiber-free control group.  Renal hypertrophy was seen in animals from
      all diabetic groups and was not diminished by any of the fiber
      sources.  In meal-fed animals there was no evidence of expansion of
      the intestinal contents volume by feeding of any fiber sources.  Guar
      gum substantially increased the viscosity of the intestinal contents in
      3 of 4 animals.  The results are consistent with the use of purified,
      highly viscous fiber sources for improving glycemic control in
      insulin-dependent diabetes.
   Descriptors--Author Keywords: DIABETES; DIETARY FIBER; GLYCATED            /DE
      HEMOGLOBIN; RENAL HYPERTROPHY; VISCOSITY
   Identifiers--KeyWords Plus: INSULIN-DEPENDENT DIABETICS; INTESTINAL        /ID
      GLUCOSE-ABSORPTION; NON-STARCH POLYSACCHARIDES; SUGAR-BEET PULP;
      GUAR-GUM; BLOOD-GLUCOSE; GLYCOSYLATED HEMOGLOBINS;
      CARBOHYDRATE-METABOLISM; LIQUID-CHROMATOGRAPHY; CONSTITUENT SUGARS
RF=Research Fronts: 88-0752 002   (DIETARY FIBER; HYDROGEN BREATH TEST; RATE  /RF
      OF STARCH DIGESTION INVITRO)
      88-0909 002   (DIETARY FIBER; GLYCEMIC RESPONSE; ANORECTAL STRICTURES
      COMPLICATING CROHNS-DISEASE)
   Cited References:
      ABRAHAM EC, 1983, V102, P187, J LAB CLIN MED
      ARO A, 1981, V21, P29, DIABETOLOGIA
      BLACKBURN NA, 1981, V46, P239, BRIT J NUTR
      BOOKCHIN RM, 1968, V32, P86, BIOCHEM BIOPH RES CO
      BOSELLO O, 1981, V3, P29, DIABETES CARE
      BUNN HF, 1981, V70, P325, AM J MED
      CANNON M, 1980, V27, P1397, LIFE SCI
      CANNON M, 1980, V27, P1397, LIFE SCI
      CHIU SS, 1985, V34, P481, METABOLISM
      CHRISTIANSEN JS, 1981, V20, P451, DIABETOLOGIA
      COHEN M, 1980, V1, P59, MED J AUSTRALIA
      DAUMERIE C, 1982, V8, P1, DIABETES METAB
      EDWARDS CA, 1987, V46, P72, AM J CLIN NUTR
      ELSENHANS B, 1980, V59, P373, CLIN SCI
      ENGLYST H, 1982, V107, P307, ANALYST
      ENGLYST HN, 1984, V109, P937, ANALYST
      GABBAY KH, 1977, V44, P859, J CLIN ENDOCR METAB
      GALLAHER D, 1986, V45, P596, FED PROC
      GARLICK RL, 1983, V71, P1062, J CLIN INVEST
      HAGANDER B, 1987, V716, P1, ACTA MED SCAND S ------------------------ CA=
      HAGANDER B, 1986, V3, P91, DIABETES RES CLIN EX --------------------- CY=
      HALL SEH, 1980, V3, P520, DIABETES CARE
CR=   HARMUTHHOENE AE, 1979, V23, P399, NUTR METAB
      HEATON KW, 1988, V47, P675, AM J CLIN NUTR
      ISAKSSON G, 1982, V82, P918, GASTROENTEROLOGY
      JARJIS HA, 1984, V51, P371, BRIT J NUTR
      JENKINS DJA, 1977, V86, P20, ANN INTERN MED
      JENKINS DJA, 1978, V1, P1392, BRIT MED J
      JOHANSEN K, 1981, V7, P87, DIABETES METAB
      JOHNSON IT, 1981, V22, P398, GUT
      KOENIG RJ, 1976, V25, P230, DIABETES
      MAHALKO JR, 1984, V39, P25, AM J CLIN NUTR
      MONNIER LH, 1981, V20, P12, DIABETOLOGIA
      PARSONS SR, 1984, V40, P66, AM J CLIN NUTR
      PHATAK L, 1988, V53, P830, J FOOD SCI
      RASCH R, 1979, V16, P125, DIABETOLOGIA
      SCHWARTZ SE, 1980, V79, P833, GASTROENTEROLOGY
      SCHWARTZ SE, 1980, V79, P833, GASTROENTEROLOGY
      SELS JP, 1987, V57, P177, BRIT J NUTR
      SEYERHANSEN K, 1976, V51, P551, CLIN SCI MOL MED
      SEYERHANSEN K, 1977, V13, P141, DIABETOLOGIA
      SHAH N, 1982, V112, P658, J NUTR
      SKYLER JS, 1987, V16, P713, PEDIATR ANN
      SMITH CJ, 1982, V61, P196, S AFR MED J
      TREDGER J, 1981, V7, P169, DIABETES METAB
      VAALER S, 1983, V31, EUR ASS STUDY DIAB I
      VORSTER HH, 1986, V69, P435, S AFR MED J
      WOOD PJ, 1978, V55, P1038, CEREAL CHEM
```

SCISEARCH®

SEARCH OPTIONS

BASIC INDEX

SEARCH SUFFIX+	DISPLAY CODE	FIELD NAME	INDEXING	SELECT EXAMPLES
/AB	AB	Abstract[1]	Word	S OAT(W)BRAN/AB
/DE	DE	Author Keywords[1]	Word & Phrase	S DIETARY(W)FIBER/DE S GLYCATED HEMOGLOBIN/DE
/ID	ID	KeyWords Plus[1]	Word & Phrase	S INSULIN(W)DEPENDENT(W)DIABETICS S SUGAR-BEET PULP/ID
/RF	RF	Research Fronts[2]	Word & Phrase	S DIETARY(W)FIBER(S)STARCH/RF S DIETARY FIBER/RF
/TI	TI	Title	Word	S RENAL(W)HYPERTROPHY/TI

+If no suffix is specified *all* Basic Index fields are searched.
1For records added to the database from January 1991.
2Research Front names are also rotated word by word. EXPANDing is recommended.

ADDITIONAL INDEXES

SEARCH PREFIX	DISPLAY CODE	FIELD NAME	INDEXING	SELECT EXAMPLES
—	AN	DIALOG Accession Number		
AU=	AU	Author Name	Phrase	S AU=GALLAHER DD
AV=	AV	Abstract Availability	Word & Phrase	S AV=ABSTRACT S AV=ABSTRACT AVAILABLE
CA=	CA	Cited Author or Cited Inventor[3,4]	Phrase	S CA=HAGANDER B S CA=KOBAYASHI T
CP=	CP	Cited Patent[3]	Phrase	S CP=JA 102915, 1978, KOBAYASHI T
CR=	CR	Cited Reference[3,5]	Phrase	S CR=HAGANDER B, 1987?
CS=	CS	Corporate Source	Word	S CS=(UNIV?(W)MINNESOTA(S)FOOD)
CW=	CW	Cited Work[3,4]	Phrase	S CW=ACTA MED SCAND S
CY=	CY	Cited Year[4]	Phrase	S CY=1987
DT=	DT	Document Type[6]	Phrase	S DT=MEETING ABSTRACT S DT=ARTICLE
—	FN	File Name		
GA=	GA	Genuine Article Number	Phrase	S GA=EJ220
GL=	GL	Country Name from Corporate Source	Phrase	S GL=USA S GL=FEDERAL REPUBLIC OF GERMANY
JN=	JN	Journal Name[7]	Phrase	S JN=NUTRITION RESEARCH
LA=	LA	Language[8]	Phrase	S LA=RUSSIAN
NR=	NR	Number of References	Numeric	S NR=2:50
PY=	PY	Publication Year	Phrase	S PY=1990:1991
RF=	RF	Research Front Code Number and Weight[9]	Phrase	S RF=88-0752 S RF=88-0752 002
SC=	SC	Journal Subject Category	Word & Phrase	S SC=(NUTRITION(S)DIETETICS) S SC=PHYSICS, ATOMIC?
SF=	SF	Subfile[10]	Word & Phrase	S SF=SOCSEARCH S SF=CC LIFE
SO=	SO	Source Information[11]	Word	S SO=(NUTRITION(W)RESEARCH)
UD=	—	Update[12]	Phrase	S UD=9101W1:9999
ZP=	ZP	Zip Code of Corporate Source	Phrase	S ZP=55108

3EXPANDing is recommended to verify forms of entry.
4Extracted from the Cited Reference field. Display includes entire Cited Reference.
5Refer to the Cited Author, Cited Work, and Cited Year fields for searching on individual parts of the Cited Reference field.
6To restrict results to journal articles, refer to the Limiting section.
7Also word searchable using SO=.
8To restrict results to the English language, refer to the Limiting section.
9Second example includes the "weight" (002) assigned to the RF Code (indicates number of citations in common with the original Research Front cluster).
10Refers to ISI publications or databases.
11Also searchable using JN=. Display includes: Journal Name, Publication Date, Volume, Issue, and Pagination.
12Not available in File 294.

SCISEARCH®

LIMITING

Sets and terms may be limited by Basic Index suffixes, i.e., /AB, /DE, /ID, /RF, /TI (e.g., S S5/TI,DE,ID), as well as by the following features:

SUFFIX	FIELD NAME	EXAMPLES
None	DIALOG Accession Number	S S3/10565265-99999999
None	Publication Year	S S2/1990:1991
/ART	Journal Article	S LOGISTIC(S)MAP?/ART
/NART	Non-Article	S S3/NART
/ENG	English Language	S STRANGE(W)ATTRACTOR?/ENG
/NONENG	Non-English Language	S S6/NONENG
/REV	Review or Bibliography[13]	S CHAOS/REV
/NREV	Not a Review or Bibliography	S S6/NREV
/CR	Contains Cited References	S (CHAOS OR NONLINEAR(W)DYNAMICS)/CR
/NOCR	No Cited References	S S5/NOCR

[13]Limits only to Reviews from 1989 forward.

SORTING

SORTABLE FIELDS	EXAMPLES
Online (SORT) and offline (PRINT[12]) AU, CS, JN, PY, RF, TI	SORT S13/ALL/CS/AU PRINT S5/5/ALL/JN/PY,D

MAPPING

MAP FIELDS	EXAMPLES
AU, GA, PY	MAP AU TEMP /CA=

OUTPUT OPTIONS†

USER-DEFINED FORMAT OPTIONS

User-defined formats may be specified using the display codes indicated in the Search Options tables, e.g., TYPE S3/AU,TI,SO,NR/1-5.

PREDEFINED FORMAT OPTIONS

NUMBER	RECORD CONTENT
Format 1	DIALOG Accession Number
Format 2	Full Record except Cited References and Abstract
Format 3	Bibliographic Citation
Format 4	Full Record[1] with Tagged Fields
Format 5, 9	Full Record[1]
Format 6	Title, Genuine Article Number, and Number of References
Format 7	Bibliographic Citation and Abstract
Format 8	Title, Genuine Article Number, Journal Subject Category, Number of References, and Research Fronts
Format 25	Full Record[1] plus Keywords and Abstract minus Cited References
Format K	KWIC (Key Word In Context) displays a window of text; may be used by itself or with other formats (does not display Author Keywords) (HILIGHT is also available)

DIRECT RECORD ACCESS

FIELD NAME	EXAMPLES		
DIALOG Accession Number	TYPE 00270906/2	DISPLAY 00270906/AU,CR	PRINT[12] 00270906/5

†TAG may be used for tagged fields, e.g., TYPE S2/AU,TI,SO/1-5 TAG.