Advances in Complex Systems, Vol. 10, No. 1 (2007) 93–115
© World Scientific Publishing Company

MAPPING TECHNOLOGICAL TRAJECTORIES AS PATENT CITATION NETWORKS: A STUDY ON THE HISTORY OF FUEL CELL RESEARCH

BART VERSPAGEN
Eindhoven Centre for Innovation Studies (Ecis),
Eindhoven University of Technology,
PO Box 513, 5600MB Eindhoven, The Netherlands
b.verspagen@tm.tue.nl

Received 1 July 2005
Revised 15 March 2006

Technological change is argued to be taking place along ordered and selective patterns, shaped jointly by technological and scientific principles, and economic and other societal factors. Historical, descriptive analysis is often used to analyze these "trajectories". Recently, quantitative methods have been proposed to map these trajectories. It is argued that such methods have, so far, not been able to illuminate the engineering side of technological trajectories. In order to fill this gap, a methodology proposed by Hummon and Doreian (1989) is used and extended to undertake a citation analysis of patents in the field of fuel cells.

Keywords: Patent citations; technological trajectories; history of fuel cell research.

1. Introduction

This paper is aimed at increasing the understanding of how technological knowledge develops and is applied in concrete commercial innovations. The emphasis is on studying the development of detailed engineering knowledge in a particular field, which is fuel cell research. Fuel cells are interesting from the point of view of the "hydrogen economy," a notion that has recently received a lot of attention (to the extent of perhaps even been hyped) and has raised great expectations of a "clean" energy system.

The paper starts from the notion that such technological developments are shaped by societal factors, among which economic factors play a large role. Economic goals and restrictions related to innovations shape the direction in which technology develops. This has been the main argument in the literature on technological paradigms and technological trajectories. This literature argues that technological development is above all "selective in technology space," i.e. of all the possible directions that a technology could have taken, only a small fraction tends to be realized.

Although this theory has been used frequently in historical research as well as, e.g. in debates around technology policy, there is surprisingly little systematic empirical material available that can be used to put the idea of a technological trajectory to the test. Most of the work in the area relies on specific data material that is often collected at relatively great cost, but cannot be applied beyond the specific case under consideration.

While nothing is wrong with such an approach, it would undoubtedly be an advantage if a general database could be applied in the analysis of trajectories in a range of technological fields. An obvious candidate for such a data source is the information available in patent documents. Patents are available for a long period (more than two centuries in the case of the US patent system), and contain detailed technical information.

This paper applies a methodology proposed in Ref. 1 and applies it to the fuel cell case. A similar analysis has been undertaken in Ref. 2 in the field of health care technologies. The present paper will attempt to draw conclusions on the specific results on fuel cells, as well as on the usefulness of the method for analyzing trajectories in general.

The rest of the paper is organized as follows. Section 2 summarizes (very briefly) the most important notions in the literature on technological trajectories. Emphasis is on the notion of a technological paradigm [3], and the role of trajectories in these. Although similar concepts have been proposed by others (e.g. natural trajectories [4] and technological guideposts [5]), these will not be discussed at length.

Section 3 introduces patent citation data for analyzing technological trajectories, and presents the methodology of citation networks, as proposed in Ref. 1, but with several extensions. This section also introduces and operationalizes the research questions, by focusing on the selective and persistent nature of technological trajectories. Section 4 presents a brief primer in the technological history of fuel cells.

Section 5 presents the core of the analysis. The database of fuel cell patents is described, and the citation network is analyzed. This section provides both a number of general network maps, which provide an impressionistic view of developments in the field, and a relatively detailed discussion of specific paths in the network that emerge as important ones. An attempt is made to interpret the network structure and the individual citation paths, both at the micro-level of chains of individual inventions, and at the level of organizations active in fuel cell research.

Section 6 will draw some main lessons, both for the fuel cell case, and with respect to the prospects of the methodology.

2. Technological Trajectories

The history books are full of radical technological breakthroughs that have changed society in fundamental ways. Examples of this are the steam engine, the automobile, and the computer. But what is often less visible in the popular accounts is how relatively “minor inventions” contribute to these radical innovations. These incremental innovations are often linked to the further refinement and development of basic breakthroughs that set the direction for development for a long time. Consider, for example, the steam engine, which is commonly associated with the mechanization of industry in the 19th century [6, 7]. The prototype steam engine usually referred to in the accounts of 19th century industrialization is the engine invented by James Watt in 1775, who at the time was an instrument maker at Glasgow University. What James Watt did was in fact to improve greatly upon a design of a steam engine that was made earlier by Newcomen (1712). The Newcomen engine was useful for pumping water out of mines, but it worked only with atmospheric pressure, and could not be employed as a source of power in other industries. James Watt added a so-called separate condenser to the design, thereby changing the operation drastically, and actually making use of the steam pressure rather than atmospheric pressure.

Watt's engine greatly improved the possibilities for application of steam as a universal power source in the emerging manufacturing industry (although during the period that is now known as the Industrial Revolution, water wheels remained the dominant source of power). But the standard set by Watt's engine was hardly the best attainable in steam power. Alessandro Nuvolari [7], following on work by Nick von Tunzelmann [6], has assembled evidence on this for the region of Cornwall during the early 19th century. He shows how in a period of approximately 30 years, the performance of steam engines introduced in this region increased two- to threefold, without the basic design of the type of engine undergoing a major change. In other words, there were relatively minor or incremental changes to the engines' crucial parts, such as boiler and cylinders, that were responsible for the rapid increase in performance, rather than a revolutionary change of the underlying technology.

The history of technological change is abound with such examples of initial radical breakthroughs followed by incremental improvements. These improvements take place during the process of diffusion of the innovation. In fact, the invention–innovation–diffusion distinction has sometimes been taken too literally as a sequential process. What the example of the steam engine shows is that diffusion of a major innovation is often associated with incremental innovations of the basic design, and these are often put on the market by firms that compete with (try to imitate) the original innovating firm. Thus, innovation, often incremental innovation, is an essential part of diffusion, rather than a predecessor of diffusion.

It is this combined process of radical breakthroughs, incremental innovations and diffusion that is described by the terms "technological paradigms" and "technological trajectories" [3]. By a technological paradigm, Dosi refers to a "model and pattern of solution of selected technological problems, based on selected principles from the natural science and on selected material technologies." The term is borrowed from Kuhn's philosophy of science, which posits that the normal development path of scientific knowledge is heavily selective in terms of a dominant framework jointly adhered to by the leading scientists in the field. From all the possible directions scientific, or in Dosi's notion, technological development may take, only a small portion gets realized, and we use the term paradigm to describe the first general selection made from all possible research directions.

We may think of a small number of basic innovations that set out a technological paradigm and dominate techno-economic developments for a long time. Along the paradigm, the basic design of the innovation is constantly altered by incremental innovations, but the basic directions in which technology develops has already been limited by the choice of paradigm. Still, there is some room for choices along the paradigm, and these choices are governed by the specific circumstances in which the technology develops. This development is termed a "technological trajectory" by Dosi. In the example of the steam engines in Cornwall, the engines were employed in copper and tin mines. This means that the coal needed to operate the engines needed to be brought into the mines (in fact, into the Cornish area over sea), and this made it relatively expensive to operate an engine. Hence, the engineers in the business of designing engines for Cornish miners set out to get as much power as they could per bushel of coal, and this goal dominated their designing efforts. The engines they developed became ever bigger (in terms of cylinder size) and more powerful. Under different circumstances, for example, steam engines operating in locomotives used for transport, designers (such as George and Robert Stephenson) had to work with a completely different aim, namely, to get as much power as possible while keeping the engine small so that it fitted on wheels and could be moved. One may imagine that again a completely different set of engineering aims applies to the case of a steam ship.

Thus, a basic innovation can be thought of as setting out developments in the techno-economic domain for a number of years to come, but the success of the paradigm, and hence of the basic innovation, depends crucially on how well incremental innovation is able to adapt the paradigm to local circumstances. For the latter question, both societal and engineering factors are important. Societal factors include the skills and capabilities of the workforce that has to work with new machinery, as well as even broader factors such as certain cultural aspects of the society in which the paradigm develops. Engineering parameters will obviously differ between technological fields, but in general one will find tradeoffs between performance in different dimensions.

This pattern of interaction between societal and technological factors along the development path (trajectory) of particular technologies is an issue that is still relatively mystified. Even the fundamental question of how to map a particular trajectory is one that has no obvious answer, and quantitative attempts to describe particular trajectories have been scarce. One notable exception to this has been the research tradition that started with the contribution of Ref. 8, whose author proposed to make a distinction between technological characteristics and service characteristics (consumer value). They then propose to use specific indicators for both and developing the relation between the two dimensions.

This approach was applied to a number of cases, among them helicopters, in Ref. 9. This contribution made it clear how the technological content of the trajectory becomes related to particular market niches, or a broader servicing of the market. The notion of variety (both on the technological and consumer side) became crucial in describing the trajectories under consideration. Reference 10 related the notion of variety in design space and function space to evolutionary dynamics, and includes notions such as branching and merging trajectories. References 11 and 12 took this one step further and applied the idea of NK-landscapes, which stems from theoretical evolutionary biology, to map evolutionary trajectories. In summary, what results from this line of research is an analytical toolbox that links the evolution of technology to the evolution of consumer demand, and provides useful insights into the joint causation between these two factors in shaping particular technological trajectories.

While this is a useful and fascinating undertaking, it is also fair to say that the approach lacks a perspective on what Ref. 3 has called the "inner dynamics" of the engineering dimension of technological paradigms, and the interaction with scientific developments. In fact, such a "backward extension" of the research on quantitatively mapping technological trajectories has largely been lacking in the literature. The insight on what occupies engineers who are actually involved in developing the (incremental) innovations along a trajectory has largely been limited to historical research with only limited quantitative underpinnings. Reference 13 is an example of the type of research that has been taking place in the historical tradition.

It is the aim of this paper to fill this niche of research that exists on mapping the engineering dimension of technological trajectories. Following Ref. 2, we propose to use patent citation networks to map technological trajectories. The particular case that will be studied is that of research on fuel cells, and US patents will be used to map the citation links between inventions in this field. The next section will lay out the basic methodology.

3. Using Citation Networks to Map Technological Trajectories

3.1. *Patent citations, trajectories and research questions*

The notion of a technological trajectory as outlined above points to technological innovations as sequential and interrelated events. One way that has been proposed in the literature to measure the interrelatedness between innovations is by means of patent citations.

The use of patent data as a technology indicator has a long tradition, but they are not, however, undisputed. Reference 14 provides a survey of the main advantages and disadvantages of using patent statistics. Patent statistics are an output indicator of innovation rather than an input indicator (such as R&D expenditures). Their main advantage is that patents are available for a rather long period, and provide detailed technological information. The main disadvantages are that simple patent counts do not take into account differences in the quality of innovations, that many patents do not lead to innovations (i.e. are not applied), and that the propensities to patent an innovation may differ between sectors and firms. In the present context of a detailed study of one technological field (fuel cells), these disadvantages do not seem to be very pressing. First of all, because we are looking at the "inner dynamics" of the trajectory, looking at commercial application is not the primary goal of the analysis. Rather, the interest lies in trying to map the technological interconnectedness between the patents. Second, the fuel cell developments indeed seem to be documented well in patents. The historical descriptions found in the literature (see, for example, Ref. 15) are indeed reflected in the patent database that we will present below.

Patent documents contain a detailed description of the patented innovation. In addition to this, the name and address of the innovator and the applicant are given. But most importantly for the present study, patent documents also contain references to previous patents, i.e. patent citations. These citations first of all serve a legal purpose: they indicate which parts of the described knowledge are claimed in the patent, and which parts have been claimed earlier in other patents.

Broadening this legal interpretation, it has been argued (see, for example, Refs. 16 and 17) that a reference to a previous patent indicates that the knowledge in the latter patent was in some way useful for developing the new knowledge described in the citing patent. This is exactly the type of interpretation that allows us to use patent citations as a tool for mapping technological trajectories in fuel cells. We will take individual patents as pieces of knowledge, or ideas, and the presence of a citation to patent X in patent Y as an indication that patent Y builds upon patent X. Obviously, a single patent may source knowledge from multiple previous patents. Also, citing patents may themselves become cited in the future, so that we will be able to map "chains" of ideas as they develop over time.

Thus, the set of patents and the citations between them naturally lend themselves to be viewed as a network of ideas and their relatedness. But a network of the size and density that we will consider for fuel cells (details will be given in the next section) is not very easily summarized beyond a rather general level that says little about the precise structure of the flows of ideas in it. However, the notion of a technological trajectory suggests that within this network, several main streams (or main paths) of knowledge exist that summarize the major developments in the field. It will be the general aim of our analysis to describe these main paths of knowledge flows in the fuel cells dataset that we employ.

More specifically, we will ask two interrelated questions about these main paths of knowledge. First, the notion of a trajectory suggests that there is a degree of selectivity about the main paths, in the sense that what emerges *ex post* as the main stream of ideas is focused in a rather limited neighborhood of technology space, and other neighborhoods, although they may have been searched to some extent, do not contribute to the main stream so much. We will first operationalize this issue by asking whether the main paths that we can identify in fuel cell research are convergent to a limited number of neighborhoods (this would be our expectation), or whether they wander in a non-converging way. We will also look at selectiveness of the technological trajectory by asking whether the main paths that we identify in the field of fuel cell research are selective with regard to the organizations that have added to the paths. More specifically, we will ask whether those patents that we will identify as belonging to the main paths in the development of fuel cell technology involve a limited selection of all organizations active in fuel cell research (this would be our expectation based on our interpretation of the idea of technological trajectories), or whether they are just a random sample of all organizations active in fuel cells (this would refute our expectations).

Second, the notion of a paradigm suggests that there is a high degree of cumulativeness about the development along the main paths. Each new innovation builds upon previous knowledge, and, in general, will extend existing ideas. Ongoing exploration of the technology space is guided by research results from the past, and, up to a certain degree, new patents are expected to extend existing paths. Hence, one would expect that a relatively high degree of persistence (over time) of the main paths is evident in the data.

On the other hand, due to the fundamental uncertainty that exists in technological search, and due to co-evolution with economic and other social factors, one may also expect that the main paths also occasionally change direction. Thus, although persistence is expected, we also expect to observe occasional splitting of the main paths, as well as convergence of separate paths (fusion). In particular, we may expect that the relative mix of persistence and exploration of new directions would change over the life time of a technology. In the beginning of the life cycle, not much knowledge exists, and hence we cannot expect much persistence. Instead, one would expect a relatively high degree of exploration and variety of main paths. Later on, one or several dominant main paths are expected to arise by the cumulative nature of knowledge building, and persistence is expected to become a dominant feature of the data, with only occasional switches of direction.

Below, we will propose a way of operationalizing the notions of persistence and exploration in our citations dataset. On the basis of that operationalization, we will ask the question whether we indeed observe a change over time in the relative mix of exploration, variety and persistence.

3.2. *Methodology*

The methodology used in this paper draws on Ref. 1. They analyze the network of citations between scientific publications on the discovery of DNA. Their aim is to construct a "main path" through this network that corresponds to the main flow of ideas in this field as represented in the formal publications. Similarly to Ref. 1, we are interested in discovering the "main flows of ideas" through a field of technological development (by means of a patent citation network), and confronting these flows with the notion of a technological trajectory.

This section will describe the methodology proposed by Ref. 1, as well as several modifications to it. The methodology rests on a number of basis concepts from network analysis, which will be explained first.

We represent a patent citation network as a collection of *vertices* and *edges*. The vertices (patents) represent pieces of knowledge that depend on each other. The edges are connections between them, in this case citations between two patents. In the particular case of citation networks, the edges are directed, i.e. they have an origin (the cited patent) and direction (the citing patent). This convention corresponds intuitively to the idea of a piece of knowledge flowing from the earlier patent to the later patent.

We represent the citation network by means of a matrix **C**, in which the element c_{ij} is equal to 1 if patent j cites patent i, and zero otherwise. Define the matrix **C*** as the symmetric matrix in which the elements are formed by taking the maximum value (in **C**) of below and above diagonal elements. We define a (weak) component in the network **C** as a subset of patents in which for every patent i and j, a path from i to j exists in the network represented by **C***. We use the concept of a (weak) component to represent a subset of the network that is somehow connected by a complex set of relations.

The citation networks that we consider are acyclic, i.e. if a path from i to j exists in **C**, no path exists from j to i. This follows logically from the nature of a citation: a patent can only be cited by patents that are published after itself, but this implies at the same time that the original patent cannot cite these later patents. This logic becomes flawed when citation to "forthcoming" patents (or patent applications) is allowed, but this does not happen in the database that will be used in this paper (it happens more frequently in scientific papers).

In the network matrix **C**, vertices may be distinguished into three categories: sources, sinks and intermediate points. Sources are vertices that make no citations, but are cited, i.e. a node i is a source if $\forall j : c_{ji} = 0$. Sinks are the opposite: they are not cited, but make citations, i.e. $\forall j : c_{ij} = 0$. Intermediate points both cite and are cited, i.e. $\exists j : c_{ij} \neq 0 \wedge \exists j : c_{ji} \neq 0$. A somewhat trivial case is the isolate: a patent that does not cite and is not cited, and hence, according to the above definitions, is both a source and a sink. Below, we will use the term startpoint to refer to a node that is a source but not an isolate, and the term endpoint to refer to a node that is a sink but not an isolate. We will use the (relative) number of isolates and the number of (weak) components of a citation network as (loose) measures for the connectedness of the network.

The most important notion in Ref. 1 for our purposes is that we can use the network structure to say something about the importance of the various individual edges (citations) in the network. Their analysis starts from the notion of a search path, which is any path from i to j in the network represented by **C**. The simplest indicator from Ref. 1 is the so-called *search path link count* (SPLC) measure. This simply enumerates all possible search paths in the network, and counts how often an edge lies on such a search path. Although Ref. 1 (seems to) suggest including even those search paths involving intermediate vertices as the origin or final destination, one may also consider the case of using only search paths from startpoints to endpoints.

A different measure is the search path node pair (SPNP) indicator. This "accounts for all connected node pairs along the search paths" (Ref. 1, p. 51), which is, as Ref. 18 observes, a somewhat unclear and imprecise statement. We follow Ref. 18 in using the following definition for the SPNP value of the edge c_{ij}. First, count all nodes in $\mathbf{C}$ for which a path to i exists, and include i itself in this count. We denote this count by n_i. Then, count the number of patents to which a path exists from j, and include j itself in this count. Call this number m_j. Now the SPNP value for c_{ij} is defined as $n_i \times m_j$. Thus, SPNP represents the number of pairs of patents that can be formed by taking one patent that lies "upstream" the edge c_{ij} and one patent that lies "downstream" this edge. As Ref. 1 observes, as a result of the multiplication, compared to SPLC, SPNP tends to weight patents on the middle of a path more heavily.

Once a measure of the importance of the edges is calculated (using either SPNP or SPLC), Ref. 1 proposes to define the "main path" through a network using the following heuristic algorithm.

(i) For each startpoint in the network, pick the (outward) edge that has maximum value, among all edges going outward of the startpoint, of SPx (where SPx is either SPLC or SPNP). If there is a tie in SPx values, take all edges that tie.
(ii) Select the startpoint(s) for which the value obtained in Step (i) is maximal. This is the startpoint(s) of the main path.
(iii) Take the target(s) (citing patent) of the edge(s) identified in the previous step.
(iv) From the target(s) identified in the previous step, pick (again) the outward edge that has a maximum value among all outward edges from this target. In case of a tie, pick all edges that tie. Add this edge(s) to the main path. If (all) these edge(s) point to an endpoint of the network, exit the algorithm, otherwise go back to Step (iii) and continue.

The intuition behind this main path is that it represents at each step (edge) the option that has attracted most weight in the SPx procedure, i.e. it represents the largest flow of ideas in the network. In the small network of Ref. 1 (40 nodes), the single main path that is identified in this way indeed represents a path that corresponds to the authors' expectations based on a (loose) historical analysis of the field. But as they observe, it is possible to construct a main path for each startpoint in the original network $\mathbf{C}$ by keeping, in Step (i) of the above algorithm, the edge(s) with maximum value obtained for each startpoint. This is the approach that we will take here.

In the case of Ref. 1, these main paths for all startpoints converge to the single main path obtained in the original version of the algorithm. Obviously, once a main path p converges to another main path q, the two will never part again. The convergence, which is by no means implied by the method, is taken in Ref. 1 as evidence of coherence in the network, and on the basis of this, the authors argue that the single main path they identify can indeed be used as a representation of the single main flow of ideas through the network (Ref. 1, p. 53ff). However, in a larger network than that to be studied below, such convergence cannot be taken for granted, and hence we will explicitly consider the main paths originating from each startpoint in the network **C**. Obviously, these main paths together can be considered as a network themselves, and we will use the symbol **C'** to denote the matrix corresponding to this "network of main paths."

Note that the procedure that we adopt here is heuristic, in the sense that it involves an arbitrary decision [at Step (ii)] to consider only a single branching point at each node of the network [or multiple branching points in the special case that outward SPx in Step (ii) ties]. It might well be the case that in this way, meaningful paths are deleted from the network. Other criteria could be used at Step (ii) (e.g. Ref. 18 proposes to use a cut-off value, resulting almost by definition in multiple out-branches at each node). It would be interesting to investigate the differences and similarities between such different methods, but this is out of scope here for space reasons.

We will use **C'** to investigate our research questions about selectivity of the main paths in the fuel cells citation network. This is done in two different ways, corresponding to the two research questions identified above. First, we will look at how many endpoints exist in the network represented by **C'**. If convergence to a limited number of neighborhoods in technology space is a real phenomenon, we would expect that in **C'**, the number of startpoints is relatively high as compared to the number of endpoints, or, in other words, that the main paths originating from multiple startpoints will converge to only a few endpoints.

Second, we will look at the presence of different organizations (patents applicants) in the patents in **C'**, as compared to the distributions of organizations in **C**. Here, we would expect selectivity to lead to the overrepresentation of a limited number of organizations in **C'** as compared to **C**.

In order to investigate our research question about persistence versus exploration, we will modify methodology of Ref. 1 on one other account. We will construct a network of main paths $\mathbf{C'}_t$ (as described above) for every sub network $\mathbf{C}_t$, where $\mathbf{C}_t$ is defined as the subset of **C** that includes only rows and columns corresponding to patents granted in a year smaller than or equal to t. Hence, $\mathbf{C'}_t$ corresponds to the network of main paths that reflects the flow of ideas up to and including the year t.

Next, identify in the network $\mathbf{C'}_t$ the single main path on which the sum of all edges is maximal, and call this path $\mathbf{p}_t$. Note that $\mathbf{p}_t$ may or may not correspond to the single main path in $\mathbf{C}_t$ of Ref. 1. The first edge on the single main path of Ref. 1 is found by maximizing the value of SPx among all startpoints, and this does not necessarily correspond to the main path found by our procedure of maximizing the sum of SPx over the whole path. In particular, we have observed cases where the two methods yield different results, and in general, the method of Ref. 1 seems to yield shorter paths.

The final step in our procedure consists of merging together the paths $\mathbf{p}_t$ for all values of $t = T_0 + a \ldots T_1$, where T_0 (T_1) is the earliest (latest) patent grant year in the dataset, and a is some non-negative number. We denote a matrix that corresponds to the resulting network $\mathbf{P}$, and take this as a representation of the temporal evolution of the main paths in the original network $\mathbf{C}$.

Our research question about the persistent nature of this network can be investigated by observing the nature of this network $\mathbf{P}$. In particular, we would expect to observe the initial exploration phase of the technology by way of a part of $\mathbf{P}$, corresponding to an early phase of the period over which we map the developments, that is characterized by a limited degree of overlaps of the paths $\mathbf{p}_t$ for different values of t. After this, we would expect the overlap of paths $\mathbf{p}_t$ to grow, with only occasional splits. Below, we will investigate this graphically by plotting the network corresponding to $\mathbf{P}$.

4. Fuel Cells: A Short History and Overview[a]

The modern fuel cell is a relatively simple piece of equipment that consists of a two electrodes "sandwiched" around a layer of material called the electrolyte. Hydrogen passes over one electrode, oxygen over the other. The hydrogen atom splits into a proton and electron. The proton passes to the other electrode through the electrolyte, while the electron becomes electric current. After being used, the current is returned to the oxygen side of the fuel cell, where it unites with the protons and the oxygen to produce water. A so-called "reformer" can be included in a fuel cell to produce hydrogen from other sources (typically hydrocarbons such as natural gas or methanol).

The history books on fuel cells usually start with the work by William Grove in Wales, 1839. Grove described an experiment in which he placed electrodes in aqueous sulfuric acid. He administered hydrogen gas to one of the electrodes and oxygen gas to the other, and observed a current flowing between the two electrodes. The "prototype for the practical fuel cell" [15] was built by Ludwig Mond and Carl Langer in 1889 in the UK. The line of research was carried on by Emil Baur in Switzerland (over the first half of the 20th century) and Francis T. Bacon in the UK (during the Second World War and onwards).

A second line of research was set in in pursuit of "electricity direct from coal." Here, solid coal was oxidized in a reaction with air. The work of William Jacques in 1896 in the US was instrumental in this line of research. Although the low efficiency of the "coal batteries" ultimately led to a dead end, the materials used in this line of research were applied in fuel cell research after the Second World War, especially so in high temperature fuel cells.

---

[a] This section is largely based on Refs. 15 and 19.

[b] The Direct Methanol FC (DMFC) is often considered as a subtype of the PEMFC. It also uses a polymer electrolyte, but it uses methanol as a fuel, by inclusion of a catalyst, which takes hydrogen directly from the methanol.

Table 1. Fuel cell types.

Fuel cell type	Technical characteristics	Operating characteristics
Polymer Exchange Membrane (PEMFC)[b]	A polymer electrolyte, clothed by a catalytic material (usually platinum), electric efficiency $\sim 50\%$	Low operating temperatures (80°C), operates only on pure hydrogen, 50–250 kW
Alkaline (AFC)	Alkali electrolyte, electric efficiency up to 70%	High operating temperature (200°C), operates only on pure hydrogen, 300 W–5 kW
Phosphoric Acid (PAFC)	Electrolyte made of concentrated phosphoric acid, electric efficiency $\sim 40\%$	High operating temperature (200°C), tolerates CO_2 in fuel
Solid Oxide (SOFC)	Electrolyte of ceramic material (e.g. zirconium dioxide), O^{2-} diffuses through the electrolyte, electric efficiency $\sim 65\%$	Operates on very high temperatures (1000°C), uses hydrogen or hydrocarbons as fuel (automatic reforming through high temperatures)
Molten Carbonate (MCFC)	Electrolytes of molten salts, still in development phase, automatic reforming	Very high operating temperature (650°C), uses many different fuels

After the Second World War, military uses were foreseen in the US, and this led to investment by the Department of Defense. A landmark in postwar research was the work of Bacon, which led to the successful demonstration of a 5 kW fuel cell in 1959. Space missions also became a successful market for fuel cells. The first Gemini mission used a so-called Polymer Exchange Membrane Fuel Cell[c] (PEMFC), but later on the Alkaline Fuel Cell developed (AFC) developed by Bacon (and patented by Pratt and Whitney) became the NASA standard. The Institute of Gas Technology (a research institute of the American gas companies) developed a research line in the 1960s on low temperature fuel cells used for generating electricity in houses that were connected to a gas network.

In the 1970s, efforts devoted to fuel cell research diminished. However, after the oil crises, interest was slowly regained, but now aimed at other applications, e.g. vehicles. During this period, government attention in other countries than the US was also on the rise.

Currently, five types of fuel cells are distinguished, mainly based on differences between the type of electrolyte used. These types are summarized in Table 1.

5. The Patent Citation Network in Fuel Cells

We use the US Patent Office database to map the citation network in fuel cells research. The field is defined in terms of technological classes, covering the (current) US patent classes 429/12–429/46 (inclusive). This is the heading in the classification

---

[c] Also: Proton Exchange Membrane.

that is called “fuel cells,” but it provides only a relatively narrow definition of fuel cell research. Developments outside this range of classes are important for fuel cell research, especially in other subclasses under the class 429. In order to keep the dataset manageable, we do not, however, include these patents.

Data are available to us for the period from the beginning of the US patent system up to and including 2002. 3,371 patents are found in the technology classes under consideration, covering the period 1860–2002. The sources of the patent citations is the NBER citations dataset [17], supplemented with data downloaded from the USPTO website. Unfortunately, the NBER dataset only contains patent citations in which the citing patent was published in 1975 or later. Hence, the only way in which patents published before 1975 can be included in the NBER dataset, is when they are cited by patents published in 1975 or later. Also, patent citations were not systematically added or documented in the early days of the US patent system: the first patent in our data set that systematically documents citations is from 1948.

The problem of having only partial citation data was solved by manually adding patent citations from the pre-1975 period to the dataset. This information was taken from the online database of the US Patent and Trademark Office (http://www.uspto.gov). In the final data set, citations appear for the period 1948–2002 (dated by cited patents). For the periods 1860–1948, citations are limited to cited patents from these periods, while the citing patents are outside these periods.[d] 15,506 citations are found between the 3,371 patents in the database (which corresponds to a 0.3% density).

Figure 1 provides some basic statistics on the citation network that is obtained. The line on top (labeled “patents”) shows the number of patents in the data set up to the year specified on the horizontal axis. Growth is slow in the beginning, with about 100 patents up to the point in the 1950s where growth takes off at a higher speed. The spell of extremely rapid growth holds on until the 1970s, when growth slows down to exponential speed (a linear trend on the logarithmic vertical axis). Because of the absence of any citations until 1948, the network consists of isolates only up to that year, and this is indicated by the overlap between the “patents” line and the “isolates” line up to that point. After this, the number of isolates fluctuates at about 200, despite the growing number of patents in the network. The number of startpoints and the number of endpoints keeps growing over the whole period since 1948, although at much slower rates over the period 1970–2002 than before. Towards the end of the period, intermediates are by far the most general type of nodes. The general level of connectedness of the network is fairly high, as the number of components (indicated in the line in the bottom of the graph) shows. The maximum number of

---

[d] This is obviously a limitation of the data of which we are unable to gauge the exact consequences in terms of outcomes of the analysis. One may argue, however, that this impact is not very severe, since there are only a limited number of patents (about 100, or 3% of the total sample) dating from the pre-1948 period; see Fig. 1).

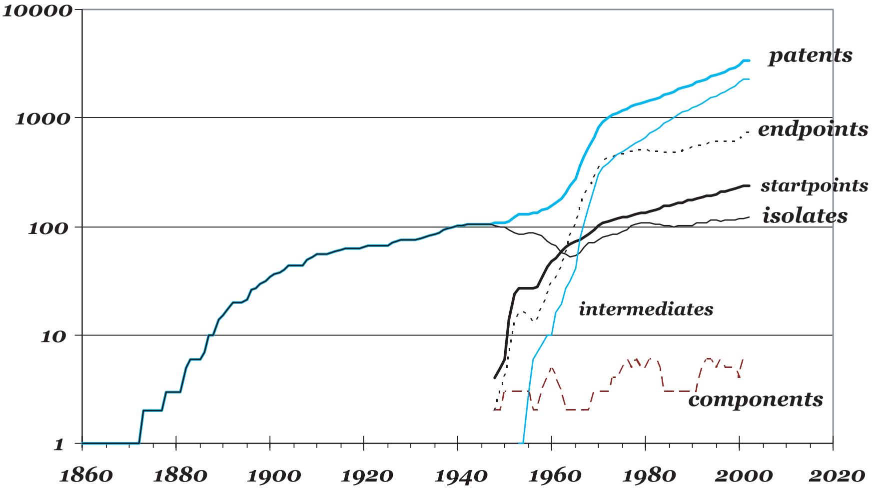

Fig. 1. Development of the fuel cell citation network over time.

components found in the network is six. There is always one large component accompanied by a number of smaller ones (these smaller components are typically smaller than five patents and never larger than ten).

In order to identify the main knowledge flows in the network as they appear towards the end of the period, we first look at the network of main paths as it emerges for the complete period 1860–2002. This is the network of the matrix **C'** as identified above. We document results based on SPNP (instead of SPLC) values throughout the whole analysis (SPLC values were also calculated, and the results are rather similar). The network of main paths has nine components of three or more patents, six of those have less than eight patents. Of the remaining three components, one has 86 patents, one 91 patents, and one has 337 patents. In the first instance, we focus attention on the largest component in the network of main paths (i.e. the one with 337 patents). This component is mapped separately in Fig. 2. The two other large components will briefly be discussed at the end of the section.

The environments of converging paths in Fig. 2 are identified by different shades. The dominating feature in the graph is a separation in two parts: a large set of patents on the right side (light shade), and a smaller number of patents on the left (dark shade). The right-hand side environment consists of paths that all converge to the single node in the center. Within this environment, there is a backbone that runs northeast to center, and to which a multitude of other (shorter) paths converges. The patent to which this environment converges, as well as the two that lie just before it, deal with air metal batteries. This is a hybrid technology between conventional batteries and fuel cells. An air metal battery functions as a conventional battery, but uses oxygen from an inflow of air as the cathode to react with water. Zinc-air batteries have been used, for example, in hearing aids, but recent applications are also aimed at automotive applications.

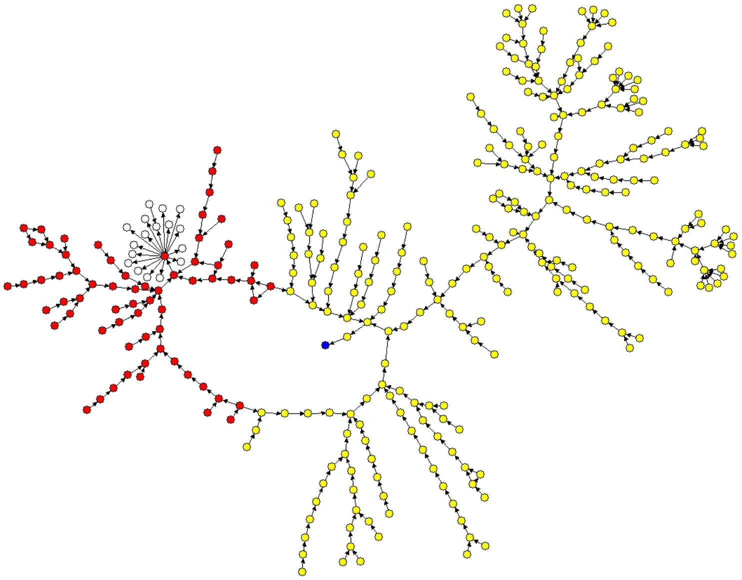

Fig. 2. The largest component in the network of main paths in fuel cell technology, 1860–2002.

The left-hand side environment converges to a set of 16 patents, which share one common "ancestor" (this is the star structure on the top-left). This specific structure is undoubtedly due to the truncation of the patent data towards the end of the period (the common ancestor receives one citation by each of the patents in the star, and then the data ends), and it would make more sense to look at the one common ancestor than at the 16 individual endpoints. This common ancestor also deals with air metal batteries.

Based on the picture of Fig. 2, we can indeed conclude that the network of main paths (**C'**) is characterized by selectivity in the sense that the main paths from the various startpoints converge to only a limited number of endpoints. In the complete network **C**, there were 721 endpoints (see also Fig. 1) for a total of 3,371 nodes (i.e. 21.4% of the nodes are endpoints). In **C'**, we have 17 endpoints for 337 nodes, i.e. 5.0%. In fact, while there are 19 startpoints in the right-hand side environment in Fig. 2, all of these converge to a single direction (the star structure). For the left-hand side environment, this is even more extreme: there are 80 startpoints that all converge to a single endpoint. Overall then, Fig. 2 has 99 startpoints that converge onto only two attracting neighborhoods. Selectivity in this sense is indeed a relevant feature of the fuel cell citations network.

Next, we turn to the issue of the organizations in the network **C'** in Fig. 2. In order to provide a benchmark for comparison, Table 2 documents all organizations in the database that hold more than 25 patents. The procedure to draw up this list is based only on the name of the patent assignee listed on the patent, and did not attempt to take into account ownership relations between organizations (e.g.

Table 2. Number of patents per company or organization in the fuel cell patents database (only patents that are cited or cite).

Rank	Organization name	Patents	Fraction of total
1	UNITED TECHNOLOGIES CORPORATION	159	0.05
2	UNITED STATES OF AMERICA	139	0.04
3	INTERNATIONAL FUEL CELLS	133	0.04
4	GENERAL ELECTRIC COMPANY	102	0.03
5	WESTINGHOUSE ELECTRIC CORP.	88	0.03
6	LEESONA CORPORATION	85	0.03
7	SIEMENS AKTIENGESELLSCHAFT	84	0.03
8	UNION CARBIDE CORPORATION	72	0.02
9	EXXON RESEARCH + ENGINEERING CO.	68	0.02
10	BALLARD POWER SYSTEMS INC.	58	0.02
11	FUJI ELECTRIC CO., LTD.	49	0.02
12	HITACHI, LTD	49	0.02
13	MITSUBISHI DENKI KABUSHIKI KAISHA	49	0.02
14	ENERGY RESEARCH CORPORATION	48	0.02
15	GENERAL MOTORS CORP.	47	0.01
16	PLUG POWER INC.	39	0.01
17	ENGELHARD CORPORATION	37	0.01
18	VARTA AG	37	0.01
19	INSTITUTE OF GAS TECHNOLOGY	35	0.01
20	NGK INSULATORS LTD.	34	0.01
21	AER ENERGY RESOURCES INC.	33	0.01
22	ALLIS-CHALMERS CORPORATION	30	0.01
23	MATSHUSHITA ELECTRIC INDUSTRIAL CO., LTD.	30	0.01
24	ASEA BROWN BOVERI LTD.	28	0.01
25	RAYOVAC CORPORATION	28	0.01
26	TOSHIBA CORPORATION	28	0.01
27	SANYO ELECTRIC CO. LTD.	26	0.01
28	TEXAS INSTRUMENTS, INCORPORATED	26	0.01
	SUM OF ABOVE	1,641	0.52
	OTHERS	1,507	0.48

mother- and daughter-firms), or mergers, acquisitions and split-ups.[e] As an exception to this rule, it was decided to create a category that unites many instances where a government organization, such as a ministry or the army, actually holds a patent (universities are not included here). This category is called "US government." It was decided to create this category to bring out the message that US government interest, although perhaps not fully coordinated between the different public agencies involved, has always been an important factor in fuel cell research.

There are 28 organizations with more than 25 patents, and together they account for about half of all patents in the database. The top five organizations in this table hold 20% of the patents. United Technologies Corp. is the largest firm, holding 159 patents. The broad US government category ranks second. Note that this category includes only those patents where such a US government organization acts as the patent holder, there are also cases where the patents document a role of the

---

[e] Small variations in names and spelling mistakes were corrected for, however.

Table 3. Number of patents in Fig. 2 per company or organization.

Rank	Company name	Patents	Fraction of main path	Patents fraction of total
1	UNITED TECHNOLOGIES CORPORATION	28	0.08	0.05
2	UNITED STATES OF AMERICA	20	0.06	0.04
3	REVEO, INC.	13	0.04	< 0.01
4	AER ENERGY RESOURCES, INC.	12	0.04	0.01
5	EXXON RESEARCH + ENGINEERING CO.	11	0.03	0.02
6	GENERAL ELECTRIC COMPANY	11	0.03	0.03
7	UNION CARBIDE AND CARBON CORPORATION	11	0.03	0.02
8	LEESONA CORPORATION	10	0.03	0.03
9	ALLIS-CHALMERS MANUFACTURING COMPANY	7	0.02	0.01
10	ELTECH SYSTEMS CORPORATION	5	0.01	< 0.01
11	INSTITUTE OF GAS TECHNOLOGY	5	0.01	0.01
12	THE STANDARD OIL COMPANY	5	0.01	< 0.01

US government in the research that led to the patent (e.g. funding), but where a non-government organization is the patent holder. Overall, US firms or organizations dominate the table, with only relatively few European, Canadian or Japanese organizations being present as main players.

Table 3 compares the dominance of the key players in fuel cells research from Table 2 between the set of main path patents in Fig. 2, and all patents in the database. The table documents which organizations are the main players in the network of main paths in Fig. 2. The top two in this table is the same as in Table 2, but the dominance of these players is even stronger than in Table 2 (United Technologies holds 8% of all patents in the network of main paths versus 5% of all patents; for the US government, this is 6% and 4%, respectively). In fact, for ten out of 12 organizations in Table 3, the fraction of patents they own in the network of main paths is larger than the fraction they own in the total network. The exceptions are General Electric Corporation and Leesona Corporation, which both show equal shares between Table 2 and 3. There are only three organizations that are present in Table 3, but not in Table 2, i.e. they are top players in the network of main path, but not so in the overall data set. These are the Standard Oil Company, Eltech Systems Corp. and Reveo Inc.

The comparison between Tables 2 and 3 also confirms that our second interpretation of selectivity (i.e. firm-level) is present in the data: the network of main paths involves indeed a selective set of patent holders. The implication is that many companies or organizations that are active in fuel cell research are not present in the network of main paths.

Finally, we turn to the temporal network of main paths **P**, which we construct for the period 1960–2002. This is displayed in Fig. 3. In this figure, the lightly

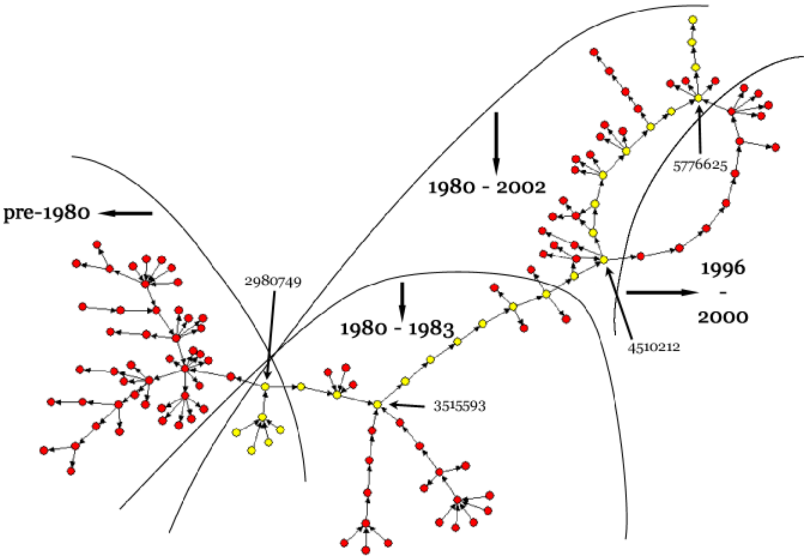

Fig. 3. The evolution of main paths over time (network **P**).

shaded dots represent patents that are on the maximal SPNP main path from $\mathbf{C'}_{2002}$, i.e. the backbone that we identified in Fig. 2. This path represents the most recent map of the technological trajectory in fuel cells research. The darkly shaded dots represent patents that were at one stage part of the main trajectory, but subsequently dropped out. Note that in Fig. 3, the years denote the t in $\mathbf{p}_t$, i.e. they indicate the periods for which the trajectories were calculated rather than the year in which the patents were filed (e.g. $\mathbf{p}_t$ for, say, 1980, may include a patent from, say, 1963, or 1923).

On the left of Fig. 3, we see a collection of patents that represent the initial phase of the development. This is the set of patents in the area marked “pre-1980.” Indeed, 1980 emerges from the analysis as the demarcation year between the phase of exploration and the building-up of the cumulative trajectory identified by the lightly shaded nodes. Before 1980, the main path changes direction quite often, resulting in the relatively complex (compared to the 1980–2002 part of the graph) structure of the network. In this period, there are three main branches (coming from above, below, and from the right) that converge on the path that stretches into the lower-left corner. But as it turns out, this is a dead end, and, instead of continuing this path, the 1980–2002 developments continue in a different direction, re-starting from one the three braches that emerged in the pre-1980 period.

The “dead end” path in the pre-1980 period is one that is focused on the development of the electrodes of fuel cells. It does not involve a single one of the organizations in Table 3. On the other hand, the (lightly shaded) path that does become the main focusing device after 1979, is one that is associated with larger systems. Initially (this is the part of the trajectory that lies before 1980), work on electrodes is also included (e.g. the work by Justi *et al.* on electrodes for the AFC that comes just before the patent labeled 2980749). The latter patent, invented by Broers in the Netherlands and filed in 1958, also includes an important part on electrodes. This is the patent at which the two trajectories, i.e. the one that runs into a dead end in 1979 and the one that continues into the period 1980–2002, split off.

The resulting main trajectory (lightly shaded) has three main sources, of which the one that originates in the 1970s is the main one. These three sources converge on to the main trajectory at the patent no. 3515593, which is a patent owned by the United Aircraft Company. This patent describes a lamination bond used in joining fuel cells in a stack in such a way that waste heat generated in the stack can be disposed of properly. The three streams that feed into the main trajectory at this point represent different types of fuel cells (in terms of Table 1). The part of the trajectory that is lightly shaded and is in the pre-1980 part of the graph brings together knowledge on the MCFC, both in the patent by Broers that was already mentioned (2980749), and three of the four patents at the basis of this part of the trajectory, which are invented by Gorin at the Pittsburgh Consolidation Coal Company in the early 1950s [20]. The second path that feeds into the main trajectory at patent no. 3515593, is based in the "direct coal" tradition of Jacques and later on Baur [15, 20]. At the basis of the third path that feeds into the main trajectory is the work by William Grubb at General Electric Corporation on the PEMFC [20].

From patent 3515593 onwards, the trajectory follows a completely linear path until 1996, at which point it splits up. Thus, the 1980–2002 period indeed shows a high degree of persistence as the main trend of fuel cell research. The trajectory that emerges in the years 1980–1995 is always along the main part of the lightly shaded path in Fig. 3. Between patents 3515593 and 4510212, this trajectory addresses the issue of temperature control (cooling) of fuel cells and stacks of fuel cells. The only split-off that occurs, besides the small diversions from the main track that are the result of truncation of the dataset at various points, is after the patent no. 4510212. This patent is owned by the U.S. Department of Energy, and results from research undertaken at the Argonne National Laboratory at the University of Chicago. This is one of the basic patents for the SOFC. Both paths that diverge from this patent are based initially in the development of the SOFC, but later on switch to the PEMFC. Interestingly, towards the end of the period, the two paths converge again into the stream on air metal batteries that was already observed in Fig. 2.

We may summarize the picture in Fig. 3 as one in which after initial exploration of various paths, the main trajectory in fuel cell research emerged at the early 1980s. This main trajectory was initially centered around several aspects of the basic design of a fuel cell, moving on in the direction of building larger systems of fuel cells, with issues of cooling becoming dominant, but turning back to the level of designing a specific type of fuel cell in the mid-1980s (SOFC), and later on to the PEMFC and air metal batteries. Thus, we indeed observe the life cycle effect discussed above, and eventually observe a high degree of persistence in the trajectory. Interestingly, the main trajectory that we find encompasses work in all main types of fuel cells that were introduced in Table 1 above.

We now briefly discuss the other two large components in the network of main paths. The second component (91 nodes) is focused exclusively on the development of the solid oxide fuel cell. It contains a backbone of development that starts early 20th century, and is joined later by six smaller branches. Although, obviously, the earlier part of this backbone does not focus specifically on solid oxide fuel cells (this is a relatively recent invention), the latter part does exclusively so. We may thus consider this second largest component as a map of a specific and isolated trajectory in the development of the solid oxide fuel cell. This is quite different from the largest component analyzed above, in the sense that this second component is much more focused. The third largest component in the network of main paths (86 nodes) has a less clear focus. It contains patents that deal with “smaller” issues that are important, but not core to the design of a (system of) fuel cell(s). Examples of this include controlling the level of CO in the so-called reformer part of a fuel cell, or specific issues in stacking and cooling. We may see this as a map of some of the more detailed issues in fuel cell technology that is not necessarily connected in a strong way to the main path in the field.

6. Conclusions

What can we learn from this analysis of citation networks in fuel cell technology? At the level of fuel cell research, the main lessons are as follows. First, the citation paths that we have analyzed in detail suggest that technological trajectories in fuel cell research are indeed selective and cumulative. Selectivity occurs both in terms of a small number of technological neighborhoods to which the main flows of fuel cell research converge, and in terms of small number of the organizations that are found to play an important role in the main paths found in the analysis. Cumulativeness is observed by the persistent nature of the main trajectory that comes out of the analysis. Persistence is a feature of the later stage of fuel cell research (1980 and beyond), while exploration of a larger number of paths is the dominant feature in the period before 1980.

We also find that the main trajectory in fuel cell research is not compartmentalized between the five types of fuel cells that are often distinguished (PEMFC, AFC, PAFC, SOFC, MCFC). Instead, the trajectory is fed by developments in different types of fuel cells, and seems to transform itself rather gradually from one type of fuel cells into another type (e.g. from AFC/MCFC to SOFC to PEMFC). In accordance with the general historical overview provided by Refs. 15 and 20, the research Alkaline FC (AFC) and Molten Carbonate FC (MCFC) are particularly important for setting out some basic principles in the late 1950s and early 1960s (e.g. the work by Justi and Winsel, Broers, and Gorin). Development of the SOFC and PEMFC dominates much of the subsequent paths in the late 1980s and 1990s.

Second, the development in part of the citation network seems to move quite naturally between different levels of aggregation. This starts with work at the level of components of fuel cells (such as electrodes or the electrolyte), then moves to single cells as the unit of research, and then moves to systems (stacks) of fuel cells. Associated with this trend is a focus on different problem areas (e.g. from micro processes at the surface of electrodes to macro problems such as cooling a stack of cells). This is, however, not a linear trend towards ever-higher levels of analysis, because the trend may also reverse.

More generally, with regard to the usefulness of the citation network method for mapping technological trajectories, the results seem somehow promising, but also point to potential problems. The main conclusions with regard to this part of the analysis are as follows. First, there is a whole spectrum with regard to the level of detail that one may look at in the citation networks. At one extreme is the single main path proposed in Ref. 1, which reduces the 3,192 patents in our data set to a path of 26. Looking at this path is insightful, but the 26 patents (obviously) ignore important trends in other parts of the network. At the other extreme is the complete set of patents (337) that is in the large component of the network of main paths for the period 1860–2002. This is too large a set of patents to easily summarize. Looking at particular environments of this set of citation paths reveals, again, interesting developments, but these are not obviously summarized into a single characteristic of the technological trajectory in a field.

Of particular interest seems to be the application of the overview of the historical development of the main paths as proposed in our network **P**. This is a methodological novelty as compared to Ref. 1 that proved particularly insightful in the case of fuel cell research. It brings out clearly the life cycle of fuel cell research as starting with broad exploration of a number of different directions, and the subsequent lock-in to a persistent path along which technological evolution is clearly interpretable. The application of this way of representing the temporal evolution of main paths in a citation network may well be fruitfully applied to other fields as well.

Second and lastly, it seems particularly promising to combine the quantitative analysis of citation paths with a detailed narrative history of the field. Citation analysis can provide additional insights in terms of outlining the important main lines of development and the interaction between different inventions. A narrative history can provide the necessary background to interpret the citation analysis results. Thus, it is obvious that the network analysis heuristics that were applied in this paper are in no way a substitute for detailed study of the engineering trends in the field. Without additional (qualitative) insights into the engineering history of a field, it is hard to make any sense of the paths in the citation network. Putting the result of the current citation analysis to further historical scrutiny must remain a priority for further research.

Acknowledgments

The author benefited from comments by and discussions with Gerald Silverberg, Önder Nomaler, Alessandro Nuvolari, Geert Verbong, Koen Frenken, two anonymous referees, participants at a workshop at the Vienna Exystence Thematic Institute for Complexity and Innovation, 2004, and participants at seminars at INRA, Grenoble, CESPRI, Bocconi University and GREDEG, University of Nice/Sophia Antipolis.

References

[1] Hummon, N. P. and Doreain, P., Connectivity in a citation network: The development of DNA theory, *Soc. Networks* **11** (1989) 39–63.

[2] Mina, A., Tampubolon, G., Ramlogan, R., Venetucci, L., McMeekin, A. and Metcalfe, J.S., Problem sequences and innovation systems: The emergence of a medical industrial complex in 'percutaneous transluminal coronary angioplasty (PCTA)', paper presented at the workshop on 'Healthy Innovation', ESRC Centre for Research on Innovation and Competition, University of Manchester, 8–10 July, 2004.

[3] Dosi, G., Technological paradigms and technological trajectories, *Res. Policy* **11** (1982) 147–162.

[4] Nelson, R. R. and Winter, S. G., In search of a useful theory of innovation, *Res. Policy* **6** (1977) 36–76.

[5] Sahal, D., *Patterns of Technological Innovation* (Addison-Wesley, 1981).

[6] Von Tunzelmann, G. N., *Steam Power and British Industrialization to 1860* (Clarendon Press, 1978).

[7] Nuvolari, A., The making of steam power technology. A study of technical change during the British Industrial Revolution, Ph.D. thesis, Eindhoven University of Technology, 2004.

[8] Saviotti, P. P. and Metcalfe, J. S., A theoretical approach to the construction of technological output indicators, *Res. Policy* **13** (1984) 141–151.

[9] Saviotti, P. P., *Technological Evolution, Variety and the Economy* (Edward Elgar Publishing, 1996)

[10] Frenken, K., Saviotti, P. P. and Trommetter, M., Variety and niche creation in aircarft, helicopters, motorcycles and microcomputers, *Res. Policy* **28** (1999) 469–488.

[11] Frenken, K., Understanding product innovation using complex systems theory, Ph.D. thesis, University of Amsterdam and University of Grenoble, 2001.

[12] Frenken, K. and Nuvolari, A., The early development of the steam engine: An evolutionary interpretation using complexity theory, *Industrial and Corporate Change* **13** (2004) 419–450.

[13] Bradshaw, G., The airplane and the logic of invention, in *Cognitive Models of Science* ed. Giere, R. (University of Minnesota Press, 1992).

[14] Griliches, Z., Patent statistics as economic indicators: A survey, *J. Econ. Lit.* **28** (1990) 1661–1707.

[15] Chen, E. History, in *Fuel Cell Technology Handbook*, ed. Hoogers, G. (CRC Press, 2003), Chap. 2.

[16] Jaffe, A. B., Trajtenberg M. and Henderson, R., Geographic localization of knowledge Spillovers as evidenced by patent citations, *Q. J. Econ.* **108** (1993) 577–598.

[17] Jaffe, A. B. and Trajtenberg, M., *Patents, Citations & Innovations. A Window on the Knowledge Economy* (MIT Press, 2002).

[18] Batagelj, V., Efficient Algorithms for Citation Network Analysis, University of Ljubljana, Department of Mathematics, cs.DL/0309023 (2003).

[19] Ulutas, Ö, Diffusion of stationary fuel cells, MSc thesis, Eindhoven University of Technology, *De Diffusie van Stationaire Brandstofcellen in de VS*, Afstudeerscriptie TeMa, Technische Universiteit Eindhoven, 2003.

[20] Perry, M. L. and Fuller, T. F., A historical perspective of fuel cell technology in the 20th century, *J. Electrochem. Soc.* **149** (2002) S59–S67.