

# TOWARDS TRANSITION THEORY



J. de Haan

TOWARDS TRANSITION THEORY

J. de Haan



ISBN: 978-90-8570-572-7

# TOWARDS TRANSITION THEORY

The research resulting in this thesis was funded by the Dutch Knowledge network on System Innovations (KSI).

ISBN 978-90-8570-572-7

All design, layout, and graphics by the author.

Printed by Wöhrmann Print Services [ [www.wps.nl](http://www.wps.nl) ].

Copyright © 2010, J. de Haan

# TOWARDS TRANSITION THEORY

## TRANSITIETHEORIE TEGEMOET

### **Proefschrift**

ter verkrijging van de graad van doctor aan de  
Erasmus Universiteit Rotterdam  
op gezag van de  
rector magnificus

Prof.dr. H.G. Schmidt

en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

Vrijdag 10 september 2010 om 9<sup>30</sup> uur

door  
Johannes de Haan

geboren te Rotterdam



**Promotiecommissie**

**Promotor:** Prof.dr.ir. J. Rotmans

**Overige leden:**

Prof.dr.ir. W.A.H. Thissen

Prof.dr. W.J.M. Kickert

Prof.dr. J.P.M. Groenewegen

# Preface

This thesis is a treatise on a theory for societal transitions: pillar theory. Societal transitions are complex processes taking place in complex systems, large-scale, long-term processes in which societal systems radically change the way they are composed and function. Since we all are part of societal systems, it speaks for itself that we ought to want to understand transitions.

Nevertheless, although several aspects of transitions have been studied from various perspectives in various disciplines, the study of societal transitions *as such* is a relatively recent development. Consequently, the knowledge on transitions is scattered over disciplines and rather fragmented.

Understanding requires theory, for even articulating what one doesn't know about a certain subject inevitably requires phrasing it as a question using concepts. Theory is an intellectual tool. It also works the other way around. Theory is also the result of understanding and it is this Janus-faced property on which this thesis is based.

The theory presented in this thesis brings together much of the current state-of-the-art on transition theorising as well as expanding it to be a complete and integrative theoretical framework for future research on societal transitions. It tries to consolidate the current knowledge on transitions and frame it in a more solid and consistent manner, flexible enough to learn from new knowledge and rigorous enough to use it to obtain that knowledge. This suggests that this theory is constructed from, rather than emerged as the gradual result of, past research. It is a piece of technology made for a purpose — as opposed to an organism, evolved for that purpose — and made to be adaptable.

In any case, this invited an unorthodox structure for the thesis. That is, this thesis is actually a tryptich; it consists of three books that, although they build forth from each other, stand on their own. The theory itself is presented in book II. The grounding of the theory can be found in book I and book III presents some applications and elaborations.



# Towards Transition Theory

## **Book I** — FOUNDATIONS

Grounding of the theory. Essayistic investigations into the philosophy of science concerning the nature of explanation and theory. Furthermore an inquiry on the consequences of complexity in the context of making theory about phenomena in complex societal systems.

## **Book II** — PILLARS

Presentation of the theory itself. Introduced without assuming prior transition knowledge, then built up by giving an exposé of the language it uses, the flesh of the theory and a formalisation in a pseudo-mathematical symbolic language.

## **Book III** — BUILDINGS

Applications of the Pillar theory. All chapters are independent elaborations, building on Book II of the thesis.



# Contents

|     |  |          |
|-----|--|----------|
| I   | FOUNDATIONS  | <b>1</b> |
| 1   | Explanation and theory . . . . .                               | 5        |
| 1.1 | Understanding explanation . . . . .                            | 6        |
| 1.2 | Explaining theory . . . . .                                    | 8        |
| 1.3 | Constructing theory . . . . .                                  | 11       |
| 2   | Consequences of complexity . . . . .                           | 13       |
| 2.1 | What consequences? . . . . .                                   | 14       |
| 2.2 | On emergence . . . . .   | 14       |
| 2.3 | On self-organisation . . . . .                                 | 22       |
| 2.4 | On causality . . . . .   | 25       |
| 3   | Towards transition theory . . . . .                            | 27       |
| 3.1 | Towards transition theory . . . . .                            | 28       |
| 3.2 | Dealing with the consequences and making good theory . . . . . | 28       |
| 3.3 | Apologetics . . . . .  | 31       |
|     | Literature . . . . .   | 33       |
| II  | PILLARS  | <b>1</b> |
| 1   | Introduction . . . . .   | 5        |
| 1.1 | About this book . . . . .                                      | 6        |
| 1.2 | About societal systems and transitions therein . . . . .       | 7        |
| 1.3 | About pillar theory . . . . .                                  | 15       |
| 2   | Language . . . . .   | 25       |
| 2.1 | Beforeword . . . . .   | 26       |
| 2.2 | The words . . . . .  | 26       |
| 2.3 | Etymology . . . . .  | 30       |
| 3   | Pillars . . . . .  | 37       |
| 3.1 | Preliminaries . . . . .  | 38       |
| 3.2 | Pillar I: Conditions . . . . .                                 | 46       |
| 3.3 | Pillar II: Patterns . . . . .                                  | 55       |
| 3.4 | Pillar III: Paths . . . . .                                    | 67       |
| 4   | Formalisation . . . . .  | 77       |
| 4.1 | Beforealisation . . . . .                                      | 78       |
| 4.2 | Functioning and constellations . . . . .                       | 79       |
| 4.3 | Conditions . . . . .   | 83       |
| 4.4 | Patterns . . . . .   | 87       |
| 4.5 | Paths . . . . .  | 90       |

|  |          |
|--|----------|
| <i>CONTENTS</i>  | vii      |
| 4.6 Modelling . . . . .  | 91       |
| Literature . . . . .   | 93       |
| Glossary . . . . .   | 101      |
| <b>III BUILDINGS</b>   | <b>1</b> |
| 1 “Transitions in Dutch Health Care” . . . . .                 | 7        |
| BUILDING NARRATIVES — A HISTORICAL CASE STUDY . . . . .        | 7        |
| 1.1 Introduction . . . . .                                     | 8        |
| 1.2 Methodology: applied pillar theory . . . . .               | 9        |
| 1.3 Dutch health care as a societal system . . . . .           | 11       |
| 1.4 The landscape . . . . .                                    | 14       |
| 1.5 Transition overview and questions raised . . . . .         | 15       |
| 1.6 Patterns in health care . . . . .                          | 17       |
| 1.7 A closer look: addressing the questions . . . . .          | 38       |
| 1.8 A closer look on the big picture . . . . .                 | 43       |
| 1.9 Conclusion and discussion: ironies and paradoxes . . . . . | 44       |
| 1.10 Appendix: the pattern matrix . . . . .                    | 47       |
| References . . . . .   | 54       |
| 2 “Building a Model on Pillars” . . . . .                      | 57       |
| BUILDING MODELS — RIGOROUS THOUGHT EXPERIMENTS . . . . .       | 57       |
| 2.1 Introduction . . . . .                                     | 58       |
| 2.2 Outline of the model . . . . .                             | 58       |
| 2.3 Model behaviour . . . . .                                  | 62       |
| 2.4 Conclusions and discussion . . . . .                       | 65       |
| 2.5 Plots and graphs . . . . .                                 | 67       |
| References . . . . .   | 72       |
| 3 “The Dynamics of Functioning” . . . . .                      | 75       |
| 3.1 Introduction . . . . .                                     | 76       |
| 3.2 Formalism . . . . .  | 77       |
| 3.3 Elaborations . . . . .                                     | 83       |
| 3.4 Conclusions and discussion . . . . .                       | 88       |
| References . . . . .   | 91       |
| <b>CODA</b>  | <b>1</b> |
| Biographical sketch . . . . .                                  | 3        |
| Samenvatting . . . . .   | 5        |



Book I  
Foundations



# Table of Contents

|     |  |           |
|-----|--|-----------|
| 1   | Explanation and theory . . . . .                               | <b>5</b>  |
| 1.1 | Understanding explanation . . . . .                            | 6         |
| 1.2 | Explaining theory . . . . .                                    | 8         |
| 1.3 | Constructing theory . . . . .                                  | 11        |
| 2   | Consequences of complexity . . . . .                           | <b>13</b> |
| 2.1 | What consequences? . . . . .                                   | 14        |
| 2.2 | On emergence . . . . .   | 14        |
| 2.3 | On self-organisation . . . . .                                 | 22        |
| 2.4 | On causality . . . . .   | 25        |
| 3   | Towards transition theory . . . . .                            | <b>27</b> |
| 3.1 | Towards transition theory . . . . .                            | 28        |
| 3.2 | Dealing with the consequences and making good theory . . . . . | 28        |
| 3.3 | Apologetics . . . . .  | 31        |
|     | Literature . . . . .   | <b>33</b> |



# Chapter 1

## Explanation and theory

What exactly does explanation entail? What is a good theory?  
What does it all *mean*?



## 1.1 Understanding explanation

Science strives for understanding — or at least it ought to. Doing science is questioning in the broad sense of the word: asking questions and doubting answers. Scientific understanding is for those who are not satisfied with just answers to questions, this is for the religious, but need also to know *why* it is an answer to the question. In other words, in scientific understanding it is of central importance that one can understand *how* an answer answers a question. Answering questions in such a way is explanation and consequently explanation is a way to understanding.

Explanation is by no means the only way to understanding. One might even argue that it is not the preferred way. Intuition, for example is in many cases the better way of understanding something. One could live a very unhappy life indeed following a rationally correct reasoning that nevertheless is incompatible with some deep conviction or feeling about something, things that cannot perhaps be made clear. Belief, is another way of understanding, of knowing rather. Who knows in this sense does not need this belief to be justified or confirmed true<sup>1</sup>. A believer needs no, nor even wants, explanation of what he knows. This does not hold for religious belief only, but also for the truths imprinted in our early childhoods, our cultural norms and values, prejudices and superstitions. Yet other ways of knowing and understanding are insight and the romantic notion of *Verstehen*. Both of which can be regarded the opposite of explanation in that the way one arrives at understanding is not, or cannot be, made explicit. These ways of knowing rely on the individual perception of the world and its cultural embedding therein to give meaning to phenomena.

If so many forms of knowing, such various ways of understanding, exist, why should explanation be the preferred one for science? One could argue that it is the conviction of the author, because he believes this to be so. Maybe he feels it is right or perhaps he is being true to what he was thought as a child. This would probably not be convincing. The point is that some questions ask for explanation. Scientific questions are such.

Some more explanation on this latter observation seems appropriate. Science is a collective endeavour, a building of mankind. Although it is true that in some cases great minds have individually contributed a lot to this building, it is the fact that it is accessible to each and everyone to build on and learn from that is its pivotal quality. If one is to build from existing scientific knowledge it needs to be trustworthy, and trust in knowledge can be gained by explaining how it became known. Therefore, if everyone is in principle to be able to use and contribute to scientific knowledge, it is a simple necessity that it is clear how this knowing came about. This is why scientific understanding necessarily comes in the form of explanation.

But what then is an explanation? Mostly, explaining why something is the case takes the form of stating what caused it to become the case. Scientific explanation would then amount to referring to some law of nature or accepted

---

<sup>1</sup>Slightly bypassing here the age-old foundational question of epistemology: What actually *is* knowledge? One of the oldest definitions, attributed to Plato, holds that for someone to know something, it must be true, this someone needs to believe it to be so and this belief needs to be justified. But how *can* one know, if not only *believe*, that the justification for one's belief that something is true is itself justified? This implies either that one can *know* only those things that are objectively proven true — and these facts need not one's belief — or knowing does not depend on the truth of the matter, rendering it equivalent to faith.

### 1.1. UNDERSTANDING EXPLANATION

scientific theory and showing that the circumstances were so that under that law it was imperative that something is the case. This view — although widely criticised, old-fashioned and covered under a pile of newer ideas on what explanation is or ought to be — is so appealing that even expounding another view would take it as a point of reference. This ‘covering-law’ model of explanation is advocated by many but popularly attributed to Hempel (see e.g. Hempel & Oppenheim, 1948). Much critique focuses on the problems concerning the nature of the law and consequently the law-like statement is by-passed in many attempts to overcome this<sup>2</sup>.

These attempts for instance replace the law-like statement by a statistical pattern and an explanation takes the form of “if the circumstances are like this then in most cases this is the case. The circumstances are like this therefore this is the case.” Obviously statistical explanations can be presented much more convincingly, but they remain essentially platitudes of the above type. What is clear is that these are not explanations at all, since one does not understand something knowing that it happens nine out of ten times. What about quantum mechanics then? Quantum mechanics explains, and explains well, in statistical terms doesn’t it? This is precisely correct, in that the answers that quantum mechanics give are statistical but the explanation itself is one of the clearest examples of a covering-law type explanation. There is a law (Schrödinger’s equation) and circumstances (a potential) and with these the probability distribution of a particle is explained.

Other attempts explain by accounting the chain of events that led to that which is the case in terms of a causal chain in the most mechanical way. These explanations take the form of “my cat broke the vase. While cleaning the mess I cut myself picking up the pieces. I had no band aids and called my brother to bring me one. He had none either and dropped by the drugstore to buy some. His future girl friend worked there. That is why my brother fell in love with her.” This, although ‘causally correct’ is at best a description of *how* my brother fell in love. Without reference to some law-like statement giving a causal chain is just a description systematically narrating what preceded what.

The fundamental problem of the covering-law model of explanation is not that the explanations make reference to laws, but that they are just that. Laws are not to be questioned and this was exactly what was desired in scientific explanation. At least one can investigate how the law came about and two main routes can be identified: 1. The law is deduced. From other laws for instance, with proven explanatory power. Or from something that has more general validity. Something of which the phenomenon to be explained is a special case. 2. The law is a generalisation or abstraction of empirical observations. Arriving at laws via either route or combinations of them is a different matter however. The problem of covering-law explanations is that in the end they too answer questions *why* something is the case in terms of the causal chain that led to *how* something came to be the case. This is the point where one could become cynical or nihilistic, since either explanation does not really exist and true understanding can not be attained, or understanding has little to do with explanation and understanding is an illusion that can be produced by explanations.

---

<sup>2</sup>See for instance Woodward’s (Spring 2009) overview of some the various models of explanation.

## CHAPTER 1. EXPLANATION AND THEORY

To not throw out the law-like baby with the philosophically problematic bathwater, Bunge (2003) suggests to make the black boxes that laws are translucent. He suggests that explanations be mechanistic and defines

“to *explain* X is to propose the mechanism(s) that give(s) rise to (or else maintains or destroys) X.”

Bunge (2003, p. 23, original emphasis).

To do this one has to analyse the system in which the explanandum resides, to find out what makes it ‘tick’ and then propose the mechanism that, according to such an analysis, produces what is the case. This gives rise to explanations of the form: “The cuckoo cuckoos because it is a part of a clock, a time-telling system comprised of cogwheels, levers and dials arranged in such a way — *insert description of ‘way’ here* — that a mechanical cuckoo is shown accompanied by the appropriate sound effects, every hour on the hour.” This is an example of a mechanistic explanation of a mechanical phenomenon, but, obviously, non-mechanical phenomena could be explained mechanistically as well.

The covering-law model of explanation and the mechanistic sophistication of it, might feed the misconception that an explanation is always to a certain extent a causal account. In the sense that the explanandum must follow from the explanans, certainly. In the sense that the explanation must be in terms of the explanandum being a situation or event preceded by a situation or event that was its cause, certainly not. This temporal interpretation of explanation and causation, though appealing, is not strictly necessary. In other words, to have explained something is not the same as to be able to have foretold it from what preceded it. For example one can explain the state of a certain physical system by observing that, given the circumstances, the composition and dimensions of the system, number of particles, etc., it has assumed the state of lowest energy or highest entropy. Indeed in thermodynamics the parameter time is peculiarly absent. Something similar holds for explanations using geometrical and symmetry arguments.

However, if something that is the case does not necessarily follow — in the temporal sense — the circumstances that cause it, then separating causes and consequences can become precarious. Especially in social scientific explanations, it sometimes appears possible to take an explanation, reverse the argument, and explain the cause with the consequence. Or does this suggest that both are epiphenomena of some ‘deeper’ cause?

### 1.2 Explaining theory

A scientific question seldom comes alone. Often various questions arise about various phenomena that are in some sense connected, for instance, because they share the same subject matter or because they treat a similar behaviour in different subjects. In any case, scientific questions often invite to be addressed in coherence. To understand some subject, some class of phenomena and the various questions it could raise scientifically, requires that the explanations too are coherent. An additional, even deeper, reason for this is the Duhem-Quine thesis (Duhem, 1906; Quine, 1951), which asserts that a hypothesis cannot be tested in isolation. Therefore a hypothesis alone is scientifically meaningless,

## 1.2. EXPLAINING THEORY

or at least senseless, and a body of hypotheses is required. This is where explanation moves towards becoming theory. But theory still is more than just a coherent set of hypotheses on a certain subject.

That a theory is, too, a coherent set of hypotheses puts some demands on what it needs to encompass as well. To address various questions in a coherent manner requires that a theory comes equipped with tools to frame the phenomena to be explained on equal terms. Or *in* equal terms in fact, since these tools are concepts to describe the phenomena or the conditions under which they occur. Simply put: a theory needs to be capable of *description*.

If this was all there was to theory then normal language was a theory and storytelling explanation. For most scientists this would sound all too post-modern and to most post-modernists this would probably be science, though they would not call it theory. Which is exactly the point: a theory should have deductive capacities; it needs to be possible to *infer* something from it. A theory allows the construction of a statement about its subject using its concepts and hypotheses, a statement that can, or that could in principle, be tested. In other words: a theory needs to be capable of *prediction*.

Note that prediction is by no means the same as forecasting. Obviously the words are synonymous in everyday speech, but in the context of theory it is useful to make a distinction. A prediction is a statement explaining the (possibly, though not necessarily, future) state of a system with the relevant circumstances and mechanisms. A forecast is a statement about the future state of a system, or the evolution towards it, based on information on some earlier state of it. A forecast always involves a temporal element, a prediction not necessarily<sup>3</sup>. In the natural, and broader, the mathematical, sciences prediction and forecast do often overlap. Moreover, in the mathematical sciences an explanation often *is* a forecast of the current state from a preceding state.

In biology, for example, one can predict that evolution will lead to species adapting to their environments and thus explain the long necks of giraffes, but one couldn't have foretold them. In social sciences this difference is even more apparent. To explain a revolution is in a certain sense possible. A prediction could take the form of a statement like: given the circumstances, these mechanisms worked that brought about a revolution, or these circumstances must have been present, given that a revolution took place. If you could foretell a *coup d'état*, you should have been arrested.

A body of concepts and hypotheses that can describe and be a basis for explanation could be called a theoretical framework. A theory then, is a coherent system of concepts and hypotheses capable of *description*, *explanation* and *prediction*.

The demands on a good theory in the previous are for the most in line with the falsificationist position like propagated by Popper (1959, 1963), if 'good' is read as 'rational' or 'scientific'.

Does not a good theory need to be true? Does it not need to correspond to facts? Strange as it may appear, this is not necessary. Not even Popper thinks so, as it appears he is only interested in the falsifiability of theory, and improving theory is making it more falsifiable<sup>4</sup> (e.g. Popper, 1959, p. 54). Obviously a theory that is blatantly untrue does not provide understanding and therefore

<sup>3</sup>For this Mayr (1982) introduces the distinction between logical prediction, the broader conception intended here, and temporal prediction which is a subclass thereof.

<sup>4</sup>Some exaggeration here, obviously, Popper does have ideas about theory and truth.

## CHAPTER 1. EXPLANATION AND THEORY

cannot be a good theory. But a theory that does not contradict the broader body of knowledge in which it is embedded, that consistently describes and explains phenomena and allows predictions can very well be a good theory even if it is based on untrue assumptions.

For example, the epicycle theory of planetary motion is based on the false assumptions that the planets move in circles called epicycles that in turn move on circles called deferents centred around a fixed point called the equant. That these assumptions were false was not known at the time the theory was designed (3<sup>th</sup> century BC) nor was it testable if these assumptions were right. It appears that this all did not matter, and rightfully so since this theory is clearly elegant, elegantly clear, and fruitful in the sense that it helped understanding the heavens and easily accommodated new knowledge, by adding epicycles. Even more important, the theory worked and allowed predictions of planetary motions, moreover it still works. However, adding epicycles obviously reduces elegance, in this sense it is interesting to note that Copernicus apparently needed far less epicycles by adopting the heliocentric hypothesis<sup>5</sup>. More to the point is the fact that Copernicus himself advertised his theory as ‘harmonious’, from which he suggested that truth was a consequence (Martens, 2009).

Another example of a good theory that is untrue is the miasma theory of diseases. This theory stated that diseases, especially contagious ones, were caused by miasma, the foul reeking vapor that rises from decomposing material. This theory was replaced by the germ theory of disease, which might be composed of statements that are themselves more true. Considering the phenomenon to be explained (cholera and malaria epidemics), the empirical data available (if the water stinks, people near get sick) and the internal consistency of the reasoning, the miasma theory was at the time a good theory. Even more so, it was elegantly simple and clear. Fruitful it was too, in that it provoked research on the relation between hygiene and disease. Moreover, it was a good theory in the societal-normative sense because the decisions that were made on its basis (hygienic measures) were necessary and proved effective.

This raises the question of what the characteristics of a *better* theory would be. In the above examples, several aesthetic criteria were used to describe the merits of a theory: harmony, simplicity, elegance. . . Appeals to beauty are quite common in assessing theory, as McAllister (1996) shows, and there is indeed a great deal to be said that aesthetic arguments are superior. Well then, a theory is better if it is more:

### **Elegant**

By which is meant exactly that. If two theories explain something as well, the more beautiful is the better one. Elegance could be found in the parsimony of a theory, though this is not strictly necessary; conceptual economisation should not come at the expense of explanatory power. Elegance can be found in the structure of the theory, in its use of symmetries or systematic composition. There is a practical side to beauty too, in that an elegant theory is more comprehensible, and if a theory is more understandable then likely its subject can be easier understood with it.

### **Clear**

Theories are to provide understanding. For this, it does not help that

---

<sup>5</sup>Although this is disputed.

### 1.3. CONSTRUCTING THEORY

the theory itself is hard to understand. This is not to say that a theory ought to be simple, or easy to apply. This is nonsense. A theory should be straightforward. The subject matter dictates whether a theory needs serious mathematical skills, is highly abstract or makes use of a large conceptual vocabulary. But it should be straightforward why this is necessary, that without the long list of concepts a certain subtle distinction can not be appreciated and the theory would not provide any insight, or that without differential calculus Newton's laws do not explain planetary motion.

#### **Fruitful**

A good theory does more than make formulation and addressing of questions possible. In working with a good theory, answering a question raises new ones, not because one is left more puzzled but rather since the understanding gained leads further inquiry towards more understanding. It is therefore also a quality of a good theory that it is compatible with other theories, so that questions at the boundaries of the subject matter can be addressed and the subject can also be better understood in its relation to other subjects. Similarly, a good theory needs to be able to adapt, grow and incorporate new insights. This ability can be the result, for instance, of a systematic, modular composition. This latter point could be interpreted as a matter of elegance as well.

## 1.3 Constructing theory

The first two demands on better theories, apart from being appeals to beauty, can be seen as applying Occam's razor in theorising, and all of them are recognised by Kuhn (1962) as reasons why new paradigms are being adopted. It appears that the main reasons for adopting or rejecting theories have little to do with truth. However, the reason that science is practised *does*, as a component of the knowledge it would like to produce. This chapter began by characterising science as asking questions and doubting answers. Obviously, if one doubts answers, one doubts their truth, and science should at least strive to arrive at true insight, with truth interpreted in the practical sense of correspondence to facts.

In this way one arrives at Feyerabend (1975), holding that there is nothing against unfalsifiable hypotheses or untruths in scientific theories, claiming even that this sometimes is essential in the progress of science. Here appears the paradox that theory is at once the result of, and the instrument to, acquire scientific knowledge. Theory is an instrument of science, a tool for scientific inquiry. Theory is the precipitation of knowledge, a representation of the scientific understanding. In the dynamic sense, it therefore is not fatal that theory, or part thereof, is untrue if it is a good instrument to arrive at knowledge that is more true. As Kuhn (1962) describes, scientific research is done *within* the constraints of paradigms, of which theories are certainly part, if not synonyms sometimes. This implies that the way empirical reality is observed and scientifically interrogated is via theory.

This is somewhat cross with the idea of science, or at least with naive interpretation of the scientific method, as induction of hypotheses and concepts

## CHAPTER 1. EXPLANATION AND THEORY

from empirical reality to subsequently verify, falsify or otherwise confirm or reject them. Rather, it invites the idea of theory as human construct used in the practice of science, like a scalpel, sharpened instead of blunted through use, and discarded if a better tool can be conceived. Theory is technology.

If this is the case, then in principle there is no objection to an *a-priori* synthesis of an entire theory of a class of phenomena. Indeed, such theories have been, and are attempted, with varying success: Freud's psychoanalysis, Parson's sociology, string theory in physics, Einstein's general relativity, as well as the quantum theories by Schrödinger and Heisenberg and their integration by Dirac, Democritus' atomic theory, the four humours physiological theory, etc. Whether these were or are all *good* theories is a different matter. In fact, systems theory as a whole can be viewed as an attempt to a meta-theory facilitating this kind of theory construction to achieve unity of science.

## Chapter 2

# Consequences of complexity

What comes with complexity? What are typical traits of complexity that one needs to understand to understand complex societal systems?



## 2.1 What consequences?

This book is part of a bundle on transition theory. Transitions occur in societal systems which are undoubtedly complex and, moreover, transitions themselves are complex phenomena. Therefore it is more than appropriate to explore the consequences of complexity.

Complex system is somewhat of a tautology since both words refer to something that consists of parts making up some whole. But in the sciences complex systems are generally considered to be a special class of systems. Sometimes the complexity lies in the build-up of the system, that is the system is composed of many parts or different kinds of parts or the parts are interrelated in a complex way. Sometimes complexity refers to the behaviour or emergent properties of the system. Sometimes, and certainly with societal systems and transitions, it refers to both. It is not the aim of this chapter to provide an overview of the state-of-the-art on complexity, since the field is broad and the literature too vast for such an endeavour. Some of the relevant classics will of course *en passant* be referred to.

The principle aim of this chapter is to provide insight into what the consequences of complexity are for theory on complex societal phenomena. This will be done in a conceptual journey. The central station of this journey is the concept of *emergence*. When a whole has a quality that its parts are lacking, one speaks of emergence. Much fog surrounds this subject, for apparently the ways systems acquire such emergent properties are mysterious.

One special such property, which is as well a special such way of acquiring it, is *self-organisation*. This next station is relevant since it is recognised that many societal systems are in some sense self-organising. One concept leads to another and when exploring self-organisation one arrives naturally at related topics like *adaptation* and *coevolution*.

Consequences presuppose causes, and causality is another seemingly mysterious force in complex systems. Causation is often an important component in explanations. Though explanations, and thus theory, need not be built on causal accounts, one still needs to understand what the peculiarities of causality in complex systems are to better understand the dynamics in them. Causes and consequences are sometimes hard to distinguish in complexity phenomena and this again has to do with emergence. Thus causality will be the last stop.

## 2.2 On emergence<sup>1</sup>

One of the most striking properties of complex systems, and complex societal systems specifically, is that they can be studied at various *levels*. Different properties or phenomena reside on different levels that have a certain autonomy with respect to each other but are also in interaction.

If one has a strong inclination to reductionism one would maintain that this is all a matter of how close one ‘zooms in’ on the system, but this would not do justice to the fact that some properties cannot be well understood by merely looking at aggregated properties of the finest grains. In a societal context for instance it is not possible to understand a healthcare system well by just studying

---

<sup>1</sup>Large parts of this section are taken from the article published as de Haan (2006), “How Emergence Arises” in *Ecological Complexity*, volume 3, pages 293-301.

the interactions of individuals. The levels in between, insurance companies, the medical professions, the medical infrastructure all have certain properties that are necessary to explain the dynamics of the healthcare system. These levels are undoubtedly *produced* by the interactions of the individuals on the ‘underlying’ level, but nevertheless need to be treated in their own right.

Conversely, taking a holist stance and only studying the properties of the system as a whole will not provide understanding either. For example, it is hardly elucidating to study an economic system in terms of macro-economical figures only, without reference to how they came about in terms of the interactions of financial institutions, labour, companies and industries.

The connection, so to speak, between levels, between systemic properties and the interactions that produce them is what the concept emergence refers to. Indeed, sometimes complex systems are simply described as systems exhibiting emergence and the levels spoken of in the above as levels of emergence. This concept therefore deserves a thorough treatment.

Not only do the levels in complex societal systems invite, or even necessitate, the use of the concept of emergence. Conversely, in speaking of emergence it is useful to employ the notion of levels. Levels allow for a conceptual separation of the emergent behaviour and the interactions that produce it. The objects and interactions causing the behaviour are considered to be residing on an underlying level, whereas the emergent behaviour itself is considered to reside at a higher level. Depending on the specific case of emergence the criteria for distinguishing the levels can be aggregation, scale, time, organisation, function or other notions.

Often the observer is only aware of one of the levels. As an illustration one can think of a fly with a degree in hydrodynamics that is caught in a tornado. Although it is familiar with the Navier-Stokes equations and thus in principle knows everything there is to know about fluid flows, it would be quite surprised to learn that the turmoil around it is in fact a highly structured flow (Goldenfeld & Kadanoff, 1999). The opposite situation is perhaps even more common, many times a structure or pattern is observed of which the underlying mechanisms are unclear. Think for instance of social structure that is observed all around while its underlying mechanisms remain elusive. Another example would be the levels of organisation in ecosystems.

Whatever the nature of the higher level may be in comparison to the underlying level, in all cases the emergent phenomenon takes place on wholes that are comprised of parts that reside on the underlying level. Emergence as qualitative novelty of a composite whole is extensively elaborated by Bunge (2003). The new properties provide the higher level with a certain autonomy. Although the higher level is composite it has a unity that is apparently bound by the novel property. This is illustrated by the fact that emergent phenomena are named and referred to as objects themselves, such as: city, forest or mind. What underlies this apparent autonomy?

Two explanations can account for the autonomy of levels. The first is a statistical argument, or rather an argument using ensembles, whereas the other one draws upon nonlinearity. The statistical reasoning follows an argument by Bedau (1997, 2002) that when several micro states correspond to one recognisable macro state the phenomenon on the macro level is autonomous. Changing from one such micro state to another then, does not lead to a perceivable difference on the higher level, although those micro states can differ greatly when observed

## CHAPTER 2. CONSEQUENCES OF COMPLEXITY

at the underlying level. Large changes on the underlying level need not lead to large changes at the higher level, in fact they need not lead to change there at all. Therein lies the autonomy of the phenomena on the higher level, and thus of the levels themselves.

The nonlinearity argument is straightforward. Since nonlinearity implies that cause and effect are not coupled with a simple proportionality relation, scales can become separated. Small changes on the underlying level can lead to severe changes on the higher level and a large perturbation on the underlying level could dampen out to be unperceivable on the higher level. This apparent causal decoupling is then interpreted as autonomy. The nonlinearity argument and the statistical argument are in fact equivalent, the first emphasising a dynamical approach and the latter a static analysis.

The autonomy of emergent phenomena makes anticipating emergence a difficult task. Unless the mechanism of the system on the underlying level is known in sufficient detail it remains a perilous undertaking to predict even the scales at which a phenomenon is to be observed. Moreover, the processing of the known system information can very well be beyond the scope of analytical or computational capabilities. Fortunately, the same autonomy allows us to study and describe an emergent phenomenon, once discovered, on its own terms. That is, the properties or phenomena that emerge do not need reference to the underlying interactions for a proper description.

### 2.2.1 The conjugate

One could ask the question whether the emergent phenomenon has meaning to the system exhibiting it. In other words, is there a difference between a phenomenon simply observed by an external observer and a phenomenon that is somehow used by the system. Intuitively it is easy to make such a distinction, for it is clear that the emergence of a layer of fertile soil is purposeful for the ecosystem producing it and that the patterns produced by bark beetles are not.

An objective criterion for emergence to be noticed by the system itself can be given in terms of its complexity<sup>2</sup>. Complexity has been recognised by several authors as a way to make emergence detectable in an objective manner, see for instance (Bonabeau & Dessalles, 1997). Crutchfield (1994), however, goes one step further and distinguishes what he calls intrinsic emergence from discovery. Intrinsic emergence is characterised by closure of newness, the effect that the computational effort necessary for the self-description of a system by the system drops at the instant emergence arises. The computational effort can be regarded as a neutral observer; it is neither part of the system nor disturbs it. Since this is the case we can state that the system recognises its emergent phenomenon.

In systems exhibiting emergence where the property or phenomenon is relevant for the system two properties emerge simultaneously. At the higher level a global property or phenomenon emerges while simultaneously the objects on the underlying level obtain an emergent property. This can be illustrated with the example of the emergence of value. Consider for example a sac of cotton as Marx (1867) does in his labour theory of value. The sac of cotton is placed in a relational context by virtue of its intrinsic properties such as its potential to

---

<sup>2</sup>Complexity here loosely defined as the amount of information necessary for the self-description of a system. Various definitions of complexity exist see for instance (Bar-Yam, 1997).

weave it into cloth. In this context it becomes possible to exchange or trade the cotton, whereby it obtains an emergent property, namely its exchange value. This exchange value emerges simultaneously with the higher-level phenomenon of cotton exchange or trade. The higher-level phenomenon, the exchange or trade, together with the emergent property on the underlying level, the exchange value, are denoted as the emergent conjugate.

Many cases of emergence in physics have the characteristics sketched above. When in such a system emergence arises, it is possible to introduce a so-called order parameter (Landau *et al.*, 1980). An order parameter is a parameter on the underlying level used to describe the higher-level phenomenon. The order parameter is, however not an intrinsic, but an emergent property of the underlying level objects. As illustrated with the example of the sac of cotton this applies to much more forms of emergence than those in physics alone. In general, when emergence consists of more than just the discovery of a phenomenon, that is when there is closure of newness, it is possible to find an emergent property on the underlying level. Since the couple of the higher-level phenomenon and its underlying level emergent property emerge simultaneously they are referred to as the emergent conjugate. In short, the concept of emergent conjugate addresses the observer from the perspective of the system itself. The system observes its emergence when the emergence comes in conjugate form.

### 2.2.2 The observer

If the observer is not considered to be the system itself, but rather as an observer in the ordinary sense, other perspectives on the emergent phenomena are possible. A perspective is dependent on the place of the observer relative to what it observes. So far the systems exhibiting emergence have been treated as if they were observed by an external observer, having no influence on what it observes.

It can be the case, however, that the observer is part of the system. If the objects on the underlying level are somehow equipped with a reflective capacity, the newness closes in a different way. Emergence in this form necessarily comes in conjugate form since the perception of the emergence is itself the emergent property on the underlying level. If the objects with reflective capacities on the underlying level can alter their interactions in response to what they perceive<sup>3</sup>, which can be readily assumed, this has complicated dynamical consequences. Moreover it has consequences for the chains of causality, which now become closed through the perceptions of the reflective objects.

Examples are readily found. Every social system is a system comprised of reflective objects and the circular causality that results from that has long been a subject of study and topic of debate. Virtual systems provide examples too. A typical and conceptually clean example would be an agent based model where the agents are adaptive with respect to their system perspective in the sense that Gell-Mann (1994) and Holland (1995) describe. Ecosystems are examples in between the former two. Species in an ecosystem perceive the behaviour of the ecosystem, or facets thereof, and adapt their behaviour. The circular causality, again, determines the dynamical behaviour of the system and therefore its evolution.

---

<sup>3</sup>If they couldn't it wouldn't be possible to know if they had reflective capacities.

### 2.2.3 Typology

In the preceding, the concept of the conjugate and the place and role of the observer with respect to the system exhibiting emergence have already been used to distinguish between qualitatively different types of emergence. If the conjugate and the observer are taken as criteria of distinction, there appear to be three distinct types. The first type is a property or phenomenon that is discovered by an external observer that has no causal effect on the system that produces the emergence. The second type is that in which the produced emergence is used by the system and this results in a reduction of complexity. This closure of newness is manifest in that the emergence comes in conjugate form, with an emergent parameter on the underlying level that is coupled to the pattern, property or phenomenon on the higher level. The third type is a property or phenomenon produced by the interactions of objects that perceive the emergence they cause. Their ability to alter their interactions accordingly introduces a circular causality to this type of emergence.

The logic of emergence coming with a conjugate or not and the observer as acting part of the system or not leads one to think of a typology with four, rather than, types of emergence. However, the case of a system with reflective objects exhibiting emergence without a conjugate can be excluded on general grounds. For on the one hand, if the objects perceive the emergence a conjugate would have emerged with this perception and on the other hand, if the objects do not perceive the emergence it effectively collapses to the first type. Hence three types of emergence.

#### I — DISCOVERY

The property or phenomenon is observed only because there is an external observer that does so. This is emergence in the eye of the beholder, hence discovery. The objects forming the system cannot be attributed an emergent property as a result of the emergent phenomenon. Only the whole acquires an emergent phenomenon or property. No conjugate emerges. Although this type of emergence allows for a more compact description of the system or its properties, which is a defining characteristic for any form of order, the observer needs reference to the higher level only.

This is conceptually the broadest class of emergence, containing not only the dynamical cases but also the more general properties of wholes that cannot be simply be viewed as aggregated properties of their parts. This type is very dependent on the observer; moreover, scientific micro reduction can kill it. Take for instance the property of temperature. As a thermodynamic quantity it is an emergent property of a whole. With the scientific micro reduction of temperature as a statistical entity defined upon the ensemble of the constituting parts, however, it becomes the simple aggregated property of the sum of its parts.

*Examples:* Self-similar patterns in a strange attractor, fractal patterns in coastal lines or river basins, rainbows, the property of colour itself.

#### II — MECHANISTIC EMERGENCE

For the recognition of the emergent phenomenon as such, an external observer is still necessary. But in contrast to type-I emergence reference to

the higher level alone does not suffice to describe the emergent behaviour and the system is not oblivious to it. The system is closed with respect to its qualitative novelty in the sense that the dynamics that produces the emergence actually depends on the emergent pattern, property or phenomenon. This has been addressed qualitatively, with Marx's (1867) example of commodities obtaining exchange value as the system obtains the emergent phenomenon of a market, and quantitatively by referring to Crutchfield's (1994) noticing of the drop in computational effort in virtual systems the instant emergence occurs. The conceptual synthesis is found in the emergent conjugate which couples the higher-level emergent phenomenon with an emergent property on the underlying level.

Since the conjugate allows the observer a more compact description of the emergent phenomenon with the use of its accompanying emergent property on the underlying level, it is reminiscent of the use of order parameters in theoretical physics. It must be noted, however that an order parameter can be defined by the observer, whereas the emergent parameter on the underlying level is a proper emergent property of the system's objects.

*Examples:* Convection rolls in hydrodynamic systems, spontaneous magnetisation, the emergence of exchange value in markets.

### III — REFLECTIVE EMERGENCE

In systems exhibiting this type of emergence the observers are among the objects of the system so no external observer is required. The objects have some reflective capacity, which enables them to observe the emergence they produce. Emergence of this type necessarily comes in conjugate form, since the perception of the emergence of the objects is itself an emergent property of the objects.

Causality in these systems is circular, since in addition to the type-I or -II emergence with upward causal direction, the system facilitates a form of downward causation. The objects perceive the emergent behaviour and are able to alter their interactions accordingly, which is in effect equivalent to downward causation. The dynamics that results from this will be denoted as recursive adaptation

*Examples:* Social systems like corporations or political parties, most any social system that can be distinguished as one by people that are part of it, ecosystems, virtual systems containing agents adapting to global systemic properties.

#### 2.2.4 Elaboration

The three types of emergence found by employing the concepts of conjugate and observer, also attach a different status to the newness that emerged. By considering the consequences of the place and role of the observer with respect to the system exhibiting emergence the boundary fades between the ontology and epistemology. This is because the question of what emerges is answered by studying to whom it emerges. The two extremes are, type-I where a system exhibits novelty that is noticed because it is noticed by an external observer, and type-III where the system is in a sense composed of observers.

## CHAPTER 2. CONSEQUENCES OF COMPLEXITY

In the case of type-I emergence the observer notices a system that has properties or exhibits behaviour that cannot be reduced to that of its constituent objects. Although the observer might find an explanatory mechanism in terms of the properties and relations of objects on the underlying level and thus find the emergent properties or behaviour to be supervenient (Kim, 1984) upon them, the description of the behaviour or the denoting of the emergent phenomenon still needs the extension of the set of properties or behaviours that can be observed<sup>4</sup>. This is what corresponds to nominal emergence as introduced by Bedau (1997), a term well chosen since it is about naming and thus acknowledging novelty. Bedau contrasts his notion of nominal emergence with weak and strong emergence.

In terms of the framework of this discussion weak emergence would be the appearance of a pattern, property or phenomenon on the higher level that cannot be explained in a trivial way by the objects and their interactions on the lower level. Not trivial is then later specified as only explainable by simulation, by letting the objects interact and observing the emergence as the system evolves. Several examples of weak emergence would still fall under type-I emergence such as the appearance of fractal patterns from simple interactions and all patterns emerging from the simple rules of cellular automata (see e.g. Wolfram, 2002). So although all nominal emergence is type-I emergence, the latter encompasses a broader set of phenomena. Type-I emergence is perhaps more akin to what Crutchfield (1994) calls discovery, hence the term discovery's attachment to type-I in the typology. Although Crutchfield's notion of emergence refers predominantly to pattern formation in dynamical systems, it appears that the scope of the term discovery can be easily extended to fit all type-I emergence by identifying pattern with property and phenomenon.

The other extreme with respect to the role and place of the observer is type-III emergence. In systems exhibiting type-III emergence the objects have a capacity to observe and their interactions can be altered according to their observations. The alteration of the interactions can in turn lead to alteration of the produced emergent behaviour. This is reminiscent of the downward causation in Bedau's strong emergence. Strong emergence however requires the emergent phenomenon or property to have irreducible causal powers over the underlying level and with type-III emergence the causality is still always bottom up. Downward causation and thus strong emergence are philosophically troubling, since they demand simultaneously that the emergence is caused by interactions on the underlying level and that the emergent phenomenon has causal influence back on the underlying level.

The statement that all type-III emergence still only features upward causality and thus would classify as weak emergence appears trivial. To illustrate that it is far from that, consider the case when an object has no choice other than to alter its interactions in a prescribed way in reaction to the perceived emergent phenomenon. Still, when one keeps in mind actors or animals this appears not controversial. If one, however, goes to the conceptual extreme of considering a mathematical object reacting to a global system property as Bar-Yam (2004) does, one touches the boundary of what observing means. Bar-Yam presents this as an argument for strong emergence in the Bedau sense of the word. However, one can also consider this as touching upon the essence of what an observer

---

<sup>4</sup>See (O'Connor, 1994) for more on supervenience and emergence.

is, as an object able to alter its interactions in response to the properties or behaviour of a higher level. The word higher here provides the distinction between a normal and a reflective object.

Even with more conventional reflective objects in mind the dynamics of type-III emergence is worth some more thought. From a more general complexity perspective it appears valid to state that systems capable of exhibiting type-III emergence are examples of what is referred to as complex adaptive systems<sup>5</sup> (CAS). Although many definitions of CAS can be given (e.g. Gell-Mann, 1994; Holland, 1995), the common denominator is a system that contains agents, that is objects, adapting their behaviour to their environment. The emergence of structures in such systems and the interactions between layers of emergent phenomena can be viewed as the result of recursive adaptation in a system exhibiting type-III emergence. Adaptive agents in CAS are attributed a systems perspective, sometimes called schema (Gell-Mann, 1994) or internal model (Holland, 1995). This systems perspective can be regarded as the emergent property, that is, the part of the emergent conjugate of the emergence in a CAS on the underlying level. This reminds in turn of what Castelfranchi in (Gilbert & Conte, 1995) calls representational emergence in his heuristic typology of forms of emergence: the emergence of a representation of an emergent phenomenon in an agent.

If such a schema allows the prediction of the future state of the system or parts thereof the agents can become anticipatory agents and the system an anticipatory system in the Rosen sense<sup>6</sup>. The paradox that a future event appears to cause an effect in the present is now simply the temporal case of the apparent downward causation in type-III emergence. It is still only apparent since the causation depends on a prediction of a future event by an agent based upon its present perceptions and not on the future event itself.

Between the two extremes regarding the position of the observer relative to the system is type-II emergence. In a sense, this is type that is conceptually the hardest to grasp. The external observer observes a system exhibiting emergent behaviour observing its emergent behaviour. The system itself becomes in a way an observer when the newness is closed in the form of the conjugate, that is, in addition to the observed phenomenon the objects on the underlying level acquire an emergent property. In the closing of this newness the complexity, the computational effort for self-description of the system, is reduced, as elaborated earlier. It is the coupling of the higher and the underlying level by the emergent conjugate that allows the system to utilise the emergence, as for instance the convection rolls in fluid dynamical systems allow the system to efficiently transport heat and the emergence of the use value allowed the cotton to function in a market context. In both cases the emergent conjugate is recognised as qualitatively fulfilling the same role in a quite different mechanism.

The difference between type-II and type-III emergence can be illustrated by considering flocks. Flocking can be regarded as an emergent phenomenon, arising from local interactions of birds. If an observed flock is the result of purely local interactions like the simulated virtual ones in (Reynolds, 1987), the emergence is type-II. If however the birds perceive the flock as such and use this perception to maintain the flock form, the emergence is type-III.

<sup>5</sup>Save for the systems of the kind studied by Bar-Yam (2004) perhaps, if they were to be considered pathological examples of type-III emergence.

<sup>6</sup>See e.g. (Dubois, 1998) on several types of anticipatory systems.



## 2.3 On self-organisation

Societal systems are self-organising systems in various senses. It is therefore relevant to discuss this subject in some depth. Apparently the term self-organisation goes back to Kant (1790) and his *Kritik der Urteilkraft* in the context of living beings. The term then becomes adopted by cybernetics (see e.g. Ashby, 1962) and later on by complex systems science (Nicolis & Prigogine, 1989; Kaufmann, 1993, for example), where it by then has obtained the connotations for which it is discussed here.

Thinking of self-organisation one intuitively thinks of a system creating or maintaining a certain order without there being some single entity imposing this order. One would almost say that the order is emergent. This is indeed an appropriate view on self-organisation. Self-organisation presupposes emergence, is a manifestation of it. The converse does not hold obviously, not all emergence is self-organisation, as the example of colour shows. However several examples of emergence in the previous section are examples of self-organisation: the Bénard convection rolls, the flocking of birds, etc.

On a more abstract note the distinction can be made clearer. Emergence is about a system having a quality that its constituent parts do not, whereas self-organisation is about the processes by which systems acquire and maintain a certain emergent property: organisation.

If, from local interactions, an organisation emerges then this emergence necessarily is type-II. The local interactions occur by virtue of the *intrinsic* properties of the entities on the lower level. The organisation, the higher-level emergent phenomenon, produced in this fashion however also makes the lower-level entities acquire an emergent property, which is how they relate to the higher-level organisation. In physical self-organising systems this is reflected in the fact that an order-parameter can be defined, coupling the microscopic variables to the macroscopic phenomenon. In a broader sense it means that self-organisation, as an emergent phenomenon, comes in conjugate form as discussed in the previous section, and therefore is of type-II.

If the emergent conjugate in type-II emergence is understood as closure of newness in the Crutchfield (1994) sense, this closure is obviously organisational closure (not thermodynamical closure, more on that shortly). This organisational closure also defines a boundary between the system and its environment.

There appears to not be a specific mechanism underlying all self-organisation, at least there is no consensus about it. Several authors highlight different mechanisms and it might be illustrative to give a tentative overview here.

### CLOSED CAUSAL CHAINS

This appears in various forms. Heylighen (2003) speaks of a sequence of events, each one causing the following to happen, until one eventually causes an event that already occurred earlier, closing a causal cycle. This is reminiscent of a point attractor in an iterative map (see e.g. Strogatz, 1994). The autocatalytic reactions that Kaufmann (1993) studies can also be viewed as closed causal chains. In such reactions the product of one chemical reaction serves as the catalyst for another, which if the chain of reactions is closed produces a self sustaining process, if an inflow of reactants and energy is available. Generally speaking all sorts of feedback mechanism would fall under this moniker.

### 2.3. ON SELF-ORGANISATION

#### VARIATION AND SELECTION

That evolutionary dynamics can lead to organisation is no surprise. The idea is that there is an underlying level with entities that are produced with variation, some of which having better chances of survival than others. Selection consists then of having the better chance. The criterion against which is selected, referred to as fitness, can be imposed by the direct environment, the neighbouring entities so to speak, or can be exogenously defined.

#### LOCAL ALIGNMENT

This refers to the lower level entities adjusting their behaviour or value according to that of their neighbours. This can be imitation when it concerns consumers or the minimisation of some potential function in the case of crystallisation. It could be the alignment of spins in a ferromagnet or more abstract alignments.

This overview is probably incomplete, nor is it to say that all these mechanisms inevitably produce self-organisation. For example several mechanisms have been shown to be able to produce self-organised criticality, a dynamic driving the system into a critical state (see e.g. Bak, 1996), but *sufficient* criteria for it have not been identified.

Thus the question remains why a system would self-organise in the first place. An organised state is not the simplest state, not the most symmetric, not the state one would expect. It is not the most symmetric in the sense that the constraints that the organisation imposes on the lower-level entities and their interactions can be seen as a preferred direction, often literally such as with the flocking birds. Indeed, the process of self-organisation has been compared to or explained as symmetry breaking (e.g. Heylighen, 2003). It is not the simplest state, that is, there are far more ways in which a system can be *disorganised*. This simple statistical observation is the basis of the second law of thermodynamics: for a given system with known energy, volume, and number of particles, there are so incredibly more possible states of that collection of particles possible corresponding to a disordered system, that the chance that one observes it in an organised state is effectively zero. The law does not *forbid* organisation, it merely states its extreme unlikelihood. Thus the second law is popularly rendered as systems tend to maximise disorder, which is measured with their entropy.

Apparently a self-organising system defies this law, but it would be better to say that it operates out of its jurisdiction. Thermodynamics applies for systems that are in quasi-equilibrium as any introductory text on the subject can testify (Callen, 1985, for example) and some, not all, self-organising systems simply are not even near equilibrium. Although they are organisationally closed they are thermodynamically open (Heylighen, 2003) and through the boundary a flow of matter, energy or information drives the system from thermodynamical equilibrium towards an organised state (Nicolis & Prigogine, 1989), which in another sense optimally facilitates this flow. The Bénard cells are of this type.

Some self-organising systems, however, do operate near equilibrium. Still they form and maintain a boundary with respect to their environment. But once organised, interaction with the environment stops until the environment changes. Examples are flocks of birds and the formation of crystals. It is interesting to notice the difference in ‘drive’ between these apparent classes of

## CHAPTER 2. CONSEQUENCES OF COMPLEXITY

self-organisation. The latter is completely internally driven, the local interactions produce a global organisation which separates it from its environment after which it freezes. The former, the Bénard cells class, is externally driven and the local interactions produce a global organisation that facilitates its own persistence in this external drive. Thus two ways to self-organisation:

### INTERNALLY DRIVEN

Local interactions favour an organised state which leads to an equilibrium.

*Examples* Flocking birds, schools of fish, magnetisation and crystallisation.

### EXTERNALLY DRIVEN

Local interactions organise the system through and with respect to a flow across the boundary of the system, leading to a maintained out-of-equilibrium state.

*Examples* The Bénard convection rolls, biological cells.

What about systems that cannot readily be attributed ‘mechanisms’ such as social and societal systems? Luhmann (1984) theorises that social systems self-organise as systems of communications. The flow driving the organisation here would be the stream of information that permeate the boundaries of social systems. Communications are produced by the system that allow it to filter the information that has meaning for it. Such a system thereby creates and maintains a boundary with its environment according to the kind of information that can be processed as meaningful. Luhmann applied this reasoning to social systems as diverse as law and love.

### 2.3.1 Adaptation and coevolution

Thus far self-organisation has been discussed as if it were a sort of process, that once having occurred remained unchanged. There is however a taste of robustness to self-organisation, that is it appears that these systems are able to remain organised even in changing circumstances. Perhaps robustness is not the appropriate word and resilience or even adaptivity are the better words.

That self-organisation is robust is not really surprising, since it is the consequence of myriads of local interactions. Any perturbation in the environment of the system, or even within the system would induce those local interactions to restore or maintain the organisation. The case is different if the changes in the environment of the system demand it to change, rather than restore, its organisation. In other words when it needs to adapt. That the more complex self-organising systems do so is well known and we, as living beings, in several ways serve as examples.

As with self-organisation itself it is hard to make general statements about the mechanisms that underly adaptation. Organisation scientists like McKelvey (1997), following Kaufmann’s (1993) interpretation, see a large role for coevolution. In fact this means evoking variation and selection again as a mechanism, but now for self-reorganisation. This interpretation of coevolution deviates somewhat from how Ehrlich & Raven (1964) introduced the concept as reciprocal selective responses. The idea here being more that self-organising systems, especially the ones studied by McKelvey (organisations, firms) and Kaufmann (living beings in general), usually find themselves enveloped in an environment that itself consists of self-organising, evolving systems.

## 2.4 On causality

Causation is an important matter since it is so central in many an explanation. But causality seems to be obscure in complex systems. Why is this?

In treating emergence the matter of causality already came up and indeed it is emergence that makes causality a complex matter. If every phenomenon could be explained in terms of the lowest level elements of a system then — at least in principle — causation would not be problematic. The interactions between those elements would produce dynamics that were to be interpreted or their interrelations would yield some property to be studied. It is clear that this is only the case in some of the more ideal typical systems studied in the natural sciences and it is also clear that in almost any societal system this explanatory mode does not apply.

To sketch the problem somewhat more clearly, consider a societal system, say healthcare. Typically one chooses some level of analysis, like ‘the patients’ or a higher level of aggregation like ‘insurance and financing’. When the dynamics on one of these levels, the latter for instance, is to be explained, one would find that it is partially caused or influenced by the former. For example, the choices of insurance company made by the patients shape to some extent the landscape of healthcare insurance and the financing of healthcare is influenced by the state of health of those patients and their individual lifestyles. Conversely, the individual patients and their interactions could be explained as constrained or even caused by the dynamics on the higher level. This ranges from how an individual arranges his personal financial affairs to course-of-life influencing decisions based upon being able to afford some treatment or not.

This is all very much reminiscent of Giddens’s (1984) way of explaining social structure as discussed earlier in this chapter, but the matter is aggravated here in the sense that one does not want to explain how social structure comes about, but rather that here one tries to explain societal phenomena that occur in systems that are patchworks of social structures that are in constant and continuous interaction. In the case of societal transitions one even has to deal with the situation that the structure of such a patchwork is changing.

Reiterating the problem in the emergence terminology introduced earlier: How does causality work in systems that are composed of several, possibly overlapping, systems exhibiting type-III emergence in coevolution?

There is no single answer to this question. For each phenomenon in a complex societal system, in general several causal chains could be identified ‘explaining’ how it came about. This causal blur, as it might be called, is the source of many never-ending debates in social science. And not only does this make explanation to a certain extent a matter of choice and taste, it makes prediction virtually impossible and constructing trustworthy models difficult. Predicting how a societal system evolves involves making a dynamical model of it that captures enough of the causal influences to mimic its behaviour within acceptable margins of error. Not only is this difficult since knowledge and knowability of complex societal systems is limited, but the often non-linear behaviour of these models can amplify small perturbations or dampen out influences that in reality would have large consequences.

Another consequence of the causal blur is that causes and consequences can become interchangeable. More precisely put, the explanans in one explanation could very well be the explanandum in another and vice versa. For instance, is

## *CHAPTER 2. CONSEQUENCES OF COMPLEXITY*

the troublesome introduction of sustainable or renewable energy sources to be explained in terms of the powerful present infrastructure and vested interests in fossil fuels, blocking upcoming alternatives or, on the same level of emergence, is one to explain the current dominance of fossil fuels by the lack of viable alternatives? Or, going to a different level of emergence, are both to be considered epiphenomena, mere by-products of consumer behaviour, enabled and constrained by the market dynamics? Each explanation is in principle just as valid, provided that it is valid. One could call this causal nihilism. From a theoretical perspective however this is far from satisfactory since one would like some uniformity, some system in the explanations.

One might be inclined to think that the problem with causation here has to do with the problematic status of downward causation in philosophy of science literature on emergence. This is however not the case. In the section on emergence the problems of downward causation were already effectively circumvented by considering type-III emergence. The problem has to do with how causation relates to explanation. If one insists that an explanation of some phenomenon is a causal account of how it came about then emergence poses a problem, since influences cross levels and this obscures explanation.

## Chapter 3

# Towards transition theory

Knowing about explanation and theory, aware of the consequences of complexity, how to move on towards transition theory?

### 3.1 Towards transition theory

The title of this chapter, section and of the thesis in its entirety “Towards Transition Theory”, invites three questions: First, what is necessary for a good transition theory?, second, what *kind* of theory is needed?, and third, why is it necessary to have a transition theory?

The first question can be approached from two paths. One path has been cleared for in the previous chapters, that is: what is demanded of a good theory. The other is clearly what the demands are that the subject matter — societal transitions — impose on this theory. The latter demands will be explored somewhat further in the next section.

Addressing the last two questions amounts to an apologetics for the rest of the thesis, in which such a theory is actually presented and elaborated. Addressing these questions is a reproach to the attitude that says: Why not wait for theory to come naturally? Why not wait for more data and in the meanwhile use the theories that are around already?

First things first.

### 3.2 Dealing with the consequences and making good theory

What are the challenges that societal transitions pose to a theory for them? What transitions are will be more properly elaborated in the next book. Here it will suffice that they are processes of fundamental change that occur in complex societal systems.

Complexity poses a twofold challenge for a theory in that complex dynamics is difficult to describe in a sensible way and in that *explaining* phenomena in the context of complexity is difficult. Some of the consequences of complexity that appear to be problems, however, are also keys to dealing with them.

A different challenge for theorising on transitions is that of empirical grounding. Part of this is because transition studies, as such, is still a nascent field of research. The consequences of this are that there is a limited amount of studies of societal transitions and, moreover, that the demarcation of what is to be taken into account in such a study is still far from clear. As transitions are processes of fundamental change in societal systems, virtually everything can bear some relevance in a transition process. This is of course related to the problem that the description of complex dynamics poses, since any solution or approach to it will involve dealing with the scale, level and properties of the concepts involved.

In summary the challenges going towards transition theory are: 1. The description of complex dynamics. 2. Explanation in the context of complexity. 3. Empirical grounding.

The following sections will elaborate how these challenges can be met and what it would mean for a good theory for societal transitions.

#### 3.2.1 Describing complex dynamics

The change that happens to a societal system is said to be fundamental, which implies that it occurs on the scale of the system itself. Thus it appears logical to

### 3.2. DEALING WITH THE CONSEQUENCES AND MAKING GOOD THEORY

find descriptive forms of that change on a scale just under that of the system<sup>1</sup>, which means descriptions on the level of societal subsystems. Then a choice can be made: will the change itself be treated as a *continuous* description in terms of variables, or qualities of these subsystems or rather as a *discrete* concatenation of typical processes of change? The former can be temporally as fine-grained as one wants, for the latter it again makes sense to choose a scale just under that of the phenomenon of interest. This would amount to a sub-transition. As it is the case that much of the current theory on transitions is in narrative, descriptive form, recounting typical transition processes, it appears sensible, or at least practical, to choose for the latter here. There are other reasons as well, though.

Many phenomena in complex systems, emergent phenomena, are examples of pattern formation. Indeed it is the very occurrence of regularity or pattern in societal dynamics that often makes it an object of inquiry. As was argued earlier, emergent phenomena have a certain autonomy with respect to the underlying dynamics that produce them. In other words, to a certain extent they can be described and considered in their own right, without reference to their direct causes. And precisely this provides an opportunity for systematic theorising. If a certain class of societal phenomena can be ‘reduced’ to the occurrence of certain patterns, then at least the *dynamics* of these phenomena can be understood in terms of concatenations of such patterns<sup>2</sup>. Concatenating patterns also introduces a sense of modularity into a theoretical framework, more on that later.

#### 3.2.2 Levels, complex causation and explanation

If one considers the patterns that were referred to earlier, which can be empirically found or theoretically proposed, as the mechanisms for explanations in the Bunge (2003) sense, then one is already quite close to theory indeed. What is now a necessity is to provide a consistent conceptual framework that can adequately describe the conditions under which these patterns form. To avoid the problems produced by the causal blur, the problem of influences crossing levels of emergence so to speak, it is important that these conditions are explicitly not properties or phenomena on the levels that underly the levels on which the patterns are found. Slightly differently put, the conditions are to be only the circumstances under which the patterns form, not the causes. This also circumvents the use of a contested concept like downward causation in one’s explanatory framework.

Such conditions can then be considered the circumstances in the covering-law type of explanation that make up the theory to be. This covering law then might not give a causal account of the phenomenon it is to explain, but as the same Bunge (1979) states, this is the case for many scientific explanations. In fact, as he points out, scientific explanations that employ differential equations, considered by some as the purest form of scientific law, merely recount how a phenomenon evolves over time, not what causes it to evolve in that way.

---

<sup>1</sup>Also the logic behind Coleman’s (1990) explanatory frame for social dynamics.

<sup>2</sup>Such patterns themselves can consequently *not* be explained anymore in this framework, any explanation of why and how these patterns emerge would entail an account of their underlying mechanisms, which is explicitly abstained from here. Within this framework these patterns can only be defined or described.



## CHAPTER 3. TOWARDS TRANSITION THEORY

The conditions are to be used together with the patterns to explain societal dynamics; the conditions are *not* to be used to explain the patterns. Again, conditions thus need to be properties or descriptions thereof on the same level of emergence as the patterns. That this, paradoxically, is a possibility is precisely because the causal blur provides the theoretician with a large degree of freedom in his choice of explanans and explanandum.

### 3.2.3 Falsifiability and empirical grounding

Lack of empirical basis as alluded to in the previous section makes an empiricist approach to theorising somewhat precarious. That is to say, the naive idea of the ‘scientific method’ seems hardly applicable. There is little data to hypothesise from and few experiments to test them with.

However, there are two important reasons why this should not be a reason to refrain from building ambitiously towards a transition theory.

The first is that the above naive ‘scientific method’ isn’t the way scientific knowledge is produced anyway, as Kuhn (1962) pointed out. This holds all the more for the emergence of theory. Feyerabend (1975) makes clear that amongst the various reasons for the adoption or rejection of a hypothesis or theory, verification or direct correspondence to data is only one. In fact, he provocatively states that any method is probably more hindrance in the progress of science than help. This epistemological anarchism, as it is called, is summarised in his phrase: anything goes. Mintzberg (2006) describes how he constructs theory in a way that can definitely not be described as following the scientific method. Anything might go, but scientists are apparently inclined to assess theories on *aesthetic* grounds, as McAllister (1996) discusses<sup>3</sup>.

The first chapter already ascertained that there is in principle, as far as scientificness is concerned, no objection to constructing a comprehensive theory. This is not to be confused with producing a grand theory like the type Merton (1968) objected against when he advocated social theories of the middle range. Merton suggested that a theory should be of the middle range, in that it didn’t get lost in abstraction nor in meaningless data-details. A modern stance that is quite similar is taken by Hedström (2005). He suggests that social theory should be build up using social mechanisms, pieces of theory, providing explanations and subject to empirical validation.

This invites a modular approach to theory construction. The main reasons for a modular approach are: It makes it possible to incorporate existing theory, just as it makes for flexibility with respect to future developments and it makes it more scientific in Popper’s (1959) sense, the parts that need to be falsifiable can be separated so that refutation of one part does not affect the rest of the theory. Constructing theory in a modular way also has spin-off advantages. Namely that parts of the theory, for instance just the descriptive language, can be used in separation. It also brings it one step closer to formalisation and modelling approaches.

Another reason for a modular approach is that a theory also provides a perspective on the phenomenon one is interested in. Theory is, as referred to

---

<sup>3</sup>Using beauty as a criterion for good theory is not as unscientific as it might appear, it is illuminating to read the accounts of Einstein and Dirac on why beauty is a superior criterion to experimental success. The author of this thesis, by the way, is of the opinion that beauty is the only criterion of any real value in any judgement.

before, also the instrument with which science is practiced. It is the lens through which one observes, and constructing a comprehensive theory in a modular way amounts to allowing oneself a broad vision. Making it good is polishing the lens.

### 3.3 Apologetics

The theory presented in the remainder of this thesis obviously follows the line suggested in the previous pages. The explanatory core of the theory is given by a set of patterns, empirically suggested but theoretically proposed, that can be concatenated to form paths describing transitions. These transition patterns form under certain conditions that are provided as well. The theory becomes scientific, in the sense of testable, in that the hypotheses about which patterns form under what conditions, are separated from the rest of the theoretical framework. In other words, the conceptual language is presented separately from the explanatory mechanisms and causal hypotheses are integrated in such a way that if they are found to be invalid they can be discarded without ‘damaging’ the larger theory. The theory was constructed using as much as possible of the already available theoretical body of knowledge, and adding to this conceptual lexicon to provide an as comprehensive as possible basis to work from.

But why not wait for theory to come naturally? Why not wait for more data and in the meanwhile use the theories that are around already? Why all *this*?

Well,

- Current theory is not sufficient

The state-of-the-art in transition theorising still is rooted very much in the technological and socio-technical transition research. A rather one-sided perspective towards how novel societal structure affects societal systems still prevails. To prevent the risk of theoretical lock-in, the theory needs to be re-framed for a broader perception of societal systems. Although the current concepts and terminology can be, and are, taken over, the broader frame invited additional concepts and mechanisms. That this was useful is shown in its application in a narrative case study (chapter 1, book III).

- There is need for a certain type of theory

Progress in the natural sciences can be partly attributed to the successful interplay of empirical research, computational and mathematical approaches and theorising. The theory presented in this thesis is explicitly constructed to be a basis for such interplay. It is first presented in a narrative, though fairly rigorous, form that is suitable for the kind of case study that is common in current transitions research. To be a stepping stone for modelling approaches it was subsequently recast in a formal form using a pseudo-mathematical symbolic language. The computational and mathematical possibilities are further explored in chapters 2 and 3 of book III.

- There might be more down the road

For as far as the theory is *not* scientific it still claims to contribute to the research field as what Balashov & Rosenberg (2002) referred to as a *metaphysical research programme*, a framework for possible theories. This goes one step further on the meta-level than the earlier remark that

### *CHAPTER 3. TOWARDS TRANSITION THEORY*

theory provides the perspective on one's subject matter. The theory, as a metaphysical research programme, provides a way of thinking about the phenomena<sup>4</sup>. In the case of the theory in this thesis this could refer for instance to viewing transitions as the interplay of a limited number of patterns and that transition paths can be viewed as concatenations of them. Or, to the idea that conditions for transitional change can be distilled from the composition of the societal system.

And now, towards transition theory. . .

---

<sup>4</sup>Very similar to Kuhn's (1962) idea of a paradigm, obviously, but the author however thinks that if finds oneself in a paradigm, passively, whereas a metaphysical research programme at least suggests the researchers active participation in it and influence on it.

# Literature

- ASHBY, W. 1962. *Principles of the self-organizing system*, chapter 1, pages 255–278. Pergamon Press: London.
- BAK, P. 1996. *How Nature Works: The Science of Self-Organized Criticality*. Copernicus Books.
- BALASHOV, Y. & A. ROSENBERG, editors 2002. *Karl Popper: “Darwinism as a Metaphysical Research Programme”*, chapter 18, pages 302–320. Routledge.
- BAR-YAM, Y. 1997. *Dynamics of Complex Systems*. Studies in Nonlinearity. Westview Press, Boulder, Colorado.
- 2004. A mathematical theory of strong emergence using multiscale variety. *Complexity*, 9(6): 15–24.
- BEDAU, M. 2002. Downward causation and the autonomy of weak emergence. *Principia*, 6: 5–50.
- BEDAU, M. A. 1997. Weak emergence. *Nous*, 31(Supplement: Philosophical Perspectives, 11, Mind, Causation, and World): 375–399.
- BONABEAU, E. & J. DESSALLES 1997. Detection and emergence. *Intellectica*, 2(25): 85–94.
- BUNGE, M. 1979, 3rd edition. *Causality and Modern Science*. Dover Publications.
- 2003. *Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge*. Toronto studies in philosophy. University of Toronto Press Incorporated, Toronto.
- CALLEN, H. 1985. *Thermodynamics and an Introduction to Thermostatistics*. John Wiley & Sons, New York, Chichester, Brisbane, Toronto, Singapore, 2nd edition.
- COLEMAN, J. 1990. *Foundations of Social Theory*. The Belknap Press of Harvard University Press, Cambridge, Massachusetts and London, England.
- CRUTCHFIELD, J. 1994. Is anything ever new? In *Integrative Themes*, edited by G. Cowan, D. Pines, & D. Melzner, volume XIX. Addison-Wesley, Reading, MA, Santa Fe.
- DUBOIS, D. M. 1998. Introduction to computing anticipatory systems. *International Journal of Computing Anticipatory Systems*, 2: 3–14.

## LITERATURE

- DUHEM, P. 1906. *La théorie physique, son objet et sa structure*. Chevalier et Rivière, Paris (Translated as: *The Aim and Structure of Physical Theory*, by Philip P. Wiener, Princeton, NJ: Princeton University Press, 1954).
- EHRlich, P. R. & P. H. RAVEN 1964. Butterflies and plants: A study in coevolution. *Evolution*, 18: 586–608.
- FEYERABEND, P. 1975, revised edition 1988. *Against Method*. Verso, New York and London.
- GELL-MANN, M. 1994. Complex adaptive systems. In *Systems Thinking*, edited by G. Midgley, volume IV. Sage publications.
- GIDDENS, A. 1984. *The Constitution of Society*. University of California Press, Berkeley and Los Angeles.
- GILBERT, N. & R. CONTE 1995. *Artificial Societies: The computer simulation of social life*. UCL Press Limited, London.
- GOLDENFELD, N. & L. P. KADANOFF 1999. Simple lessons from complexity. *Science*, 284(5411): 87–89.
- DE HAAN, J. 2006. How emergence arises. *Ecological Complexity*, 3(4): 293–301.
- HEDSTRÖM, P. 2005. *Dissecting the Social*. Cambridge University Press, Cambridge, New York, Melbourne.
- HEMPEL, C. G. & P. OPPENHEIM 1948. Studies in the logic of explanation. *Philosophy of Science*, 15(2): 135–175.
- HEYLIGHEN, F. 2003. The science of self-organization and adaptivity. In *in: Knowledge Management, Organizational Intelligence and Learning, and Complexity, in: The Encyclopedia of Life Support Systems, EOLSS*, pages 253–280. Publishers Co. Ltd.
- HOLLAND, J. 1995. *Hidden Order: How Adaptation Builds Complexity*. Ulam Lectures Series. Helix books / Perseus books, Cambridge, Massachusetts.
- KANT, I. 1790. *Kritik der Urteilskraft*.
- KAUFMANN, S. A. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, Oxford.
- KIM, J. 1984. Concepts of supervenience. *Philosophy and Phenomenological Research*, 45(2): 153–176. FLA00318205 State University of New York at Buffalo Copyright 1984 International Phenomenological Society.
- KUHN, T. S. 1962, 3rd edition. *The structure of scientific revolutions*. University of Chicago Press, London, Chicago, second edition edition.
- LANDAU, L., E. LIFSHITZ, & L. PITAEVSKII 1980. *Course of Theoretical Physics, Volume 1-10*.
- LUHMANN, N. 1984. *Social Systems*. Stanford University Press, Stanford, California.

- MARTENS, R. 2009. Harmony and simplicity: aesthetic virtues and the rise of testability. *Studies In History and Philosophy of Science Part A*, 40(3): 258 – 266.
- MARX, K. 1867. *Das Kapital*. Verlag von Otto Meissner, Hamburg.
- MAYR, E. 1982. *The growth of biological thought: Diversity, evolution, and inheritance*. Belknap Pr.
- MCALLISTER, J. W. 1996. *Beauty and Revolution in Science*. Cornell University Press.
- McKELVEY, B. 1997. Quasi-natural organization science. *Organization Science*, 8(4): 351–381.
- MERTON, R. K. 1968. *Social Theory and Social Structure*. The Free Press, New York.
- MINTZBERG 2006. Developing theory about the development of theory. In *Oxford Handbook of Management Theory*, edited by M. Hitt & K. Smith.
- NICOLIS, G. & I. PRIGOGINE 1989. *Exploring Complexity*. W.H. Freeman and Company, New York.
- O’CONNOR, M. 1994. Emergent properties. *American Philosophical Quarterly*, 31: 91–104.
- POPPER, K. R. 1959. *The Logic of Scientific Discovery*. Hutchinson & Co Ltd. — 1963. *Conjectures and Refutations*. Routledge, London and New York.
- QUINE, W. V. O. 1951. Two dogmas of empiricism. *The Philosophical Review* (Reprinted in “*From a Logical Point of View*”, Harvard University Press, 1953; 2nd, revised, edition 1961), 60: 20–43.
- REYNOLDS, C. W. 1987. Flocks, herds and schools: A distributed behavioural model. In *SIGGRAPH ‘87*, edited by M. C. Stone, volume 21 of *Computer Graphics*.
- STROGATZ, S. H. 1994. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Studies in Nonlinearity. Perseus, Cambridge, MA.
- WOLFRAM, S. 2002. *A New Kind of Science*. Wolfram Media, Inc.
- WOODWARD, J. Spring 2009. Scientific explanation. In *The Stanford Encyclopedia of Philosophy*, edited by E. N. Zalta.



Book II

Pillars





# Table of Contents

|     |  |            |
|-----|--|------------|
| 1   | Introduction . . . . .                                   | <b>5</b>   |
| 1.1 | About this book . . . . .                                | 6          |
| 1.2 | About societal systems and transitions therein . . . . . | 7          |
| 1.3 | About pillar theory . . . . .                            | 15         |
| 2   | Language . . . . .                                       | <b>25</b>  |
| 2.1 | Beforeword . . . . .                                     | 26         |
| 2.2 | The words . . . . .                                      | 26         |
| 2.3 | Etymology . . . . .                                      | 30         |
| 3   | Pillars . . . . .  | <b>37</b>  |
| 3.1 | Preliminaries . . . . .                                  | 38         |
| 3.2 | Pillar I: Conditions . . . . .                           | 46         |
| 3.3 | Pillar II: Patterns . . . . .                            | 55         |
| 3.4 | Pillar III: Paths . . . . .                              | 67         |
| 4   | Formalisation . . . . .                                  | <b>77</b>  |
| 4.1 | Beforealisation . . . . .                                | 78         |
| 4.2 | Functioning and constellations . . . . .                 | 79         |
| 4.3 | Conditions . . . . .                                     | 83         |
| 4.4 | Patterns . . . . .                                       | 87         |
| 4.5 | Paths . . . . .  | 90         |
| 4.6 | Modelling . . . . .                                      | 91         |
|     | Literature . . . . .                                     | <b>93</b>  |
|     | Glossary . . . . .                                       | <b>101</b> |



# Chapter 1

## Introduction

## CHAPTER 1. INTRODUCTION

### 1.1 About this book

This book is an introduction to pillar theory, a treatise of a theory of societal transitions. It intends to provide understanding of the profound shifts parts of society sometimes go through. It intends to provide the conceptual tools to further deepen this understanding. Because one wants to know. Because, not only has society in the past altered the way it met its needs, it is doing so right now in various societal contexts. Because one wants to also know ways to change ways in various societal contexts. Perforce. In this sense these pillars are a classic scientific endeavour since they seek understanding out of sheer curiosity, while aware and having at heart that their knowledge be put to use.

The way this theory came into being however is far from a classic scientific endeavour. Since the study of transitions in their own right, that is, not from a disciplinary or methodological perspective, is a relatively recent development, theorising on transitions proper is still in its infancy. Pillar theory is an attempt to bring together in a coherent whole the present body of knowledge on transitions, and to facilitate further theory development. It is therefore a *constructed* theory. It is not the result of several iterations of the inductive cycle but rather the deductive result of a theoretical framework inductively erected from current knowledge. Since a theory devised in this fashion cannot hope to be the final word on transitions it is constructed to be flexible enough to accommodate new insights.

Scientific theories generally consist of a coherent collection of concepts and hypotheses. Pillar theory is no different. With the former it provides a language to describe transitions and the latter are the explanatory core of the theory aiming to provide a better understanding of transitions. Reading this book should enable the reader to describe and explain societal transitions and their various aspects with pillar theory's conceptual tools. Explanation of transitions is a threefold matter here. First, the conditions are considered that make a societal system prone to transitional change. Second, the patterns of transitional change that can form under these conditions are given. And third, the transition paths that result are constructed. Conditions, patterns and paths are the pillars the theory is built on, hence the pillar theory of transitional change<sup>1</sup>.

Separating in this way transitional change from what drives it and what it leads to is reminiscent of the way natural sciences treat dynamics. This is not coincidentally so. Apart from deepening the understanding of transitions and providing means to adequately describe them in a narrative form, this theory is meant as a basis for computational and mathematical approaches and consequently employs such a structure. The author is well aware that this latter sentence is enough to scare off one half of his potential audience and provoke the postmodern loathing of the other. This book, however, is intended for anyone with a general and genuine interest in societal transitions. All is built up from the ground up and no prior knowledge is necessary nor assumed, just some patience and intelligence. The structure of the book is straightforward and as follows:

The chapter following this one, chapter 2 is an exposé of the language. It acquaints the reader with the core concepts of pillar theory. Many of these

---

<sup>1</sup>The theory was originally dubbed *pillars of change*, which made place for *pillar theory* or simply *pillars*. Pillars of change is somewhat misleading also since it suggests that the theory is about change in general rather than transitional change in specific.

## 1.2. ABOUT SOCIETAL SYSTEMS AND TRANSITIONS THEREIN

concepts are widely used in transition studies or other fields. Consequently they might have several interpretations and to avoid unclarity there will be one section giving the definitions that will be used in this book and a section on their origins, their etymology. The pillar-definitions are also collected in a glossary, supplemented at the end of this book.

Chapter 3 treats the pillars themselves. The first section deals with the conditions for transitional change. How transitional change is different from normal societal change will be discussed and how one can diagnose from a description of a societal system if it is prone to it, is treated here. From the systems composition to its condition, so to speak. The second section deals with the transitional change itself. It proposes that three patterns of transitional change can be distinguished — no more and no less — driven by the aforementioned conditions. These patterns contain the processes and mechanisms that produce transitional change and can be used as building blocks for descriptions and models of transitions. The last section uses the conditions and patterns to describe the unfolding of transitions over time, the paths. From the interplay between conditions and patterns a typology of transition paths is deduced, although in principle arbitrarily complex paths can be described, which is elaborated as well.

Chapter 4 is a formalisation of the theory. The rather rigid structure of the theory is exploited by casting it in a quasi-mathematical formalism with notation to go with it. The aim of this is to bring everything down to its bare essence gaining clarity and overview. The whole theory can be concisely presented and its structure becomes self-evident. Removing or bringing to light the ambiguity of normal language in this way allows for greater analytical rigour and serves computational and mathematical modelling approaches based on the theory. Another advantage is that the notation allows transition narratives to be coded, as it were, into strings that still contain the relevant information to reproduce the narrative. This makes for greater comparability of cases and of cases with models.

The rest of this chapter continues with a general introduction to the subject matter the theory deals with. In other words: what are societal transitions actually and what happens to a societal system when they happen. Furthermore some examples of societal systems and transitions will be given, transitions that have taken place, could or should take place or that are happening at present. This will all be done in normal language since the language of the theory is not introduced at this point yet. This general introduction then continues to sketch the theory and its range of application, this means describing what aspects of societal transitions the theory applies to and in what form it does that as well as its limitations. Since the theory is new and, although building forth on an existing body of knowledge on transitions, for a large part speculative, this is also the place where the theory presents its ambitions and their converse, its restrictions.

## 1.2 About societal systems and transitions therein

In the centre of a large city a man walks in the middle of the main street; he smiles at the woman passing him by with large and elegant strides and politely greets her. He is going on seventy years now and his pace is not as fast as it once was. But, he ponders, when it was, when I was the age my grandson is

## CHAPTER 1. INTRODUCTION

now, it was sheer impossible to be walking in the middle of this street in the centre of this city, or any city for that matter. It is a sunny spring afternoon and the city smells of flowers and people smelling of perfumes smelling of flowers. A lot has changed he thinks. Although he was always aware that society was in a state of constant flux, the profoundness of the changes the cities have undergone only now strikes him. The realisation rises in his mind that his grandson does not even have a recollection of cities as collections of cluttered-up and dangerous traffic jams, insalubriously reeking of the gases of the internal combustion engines. He himself was of the first generation that grew up with mobile telephones omnipresent, and remembered not believing then that his father had lived in times when these simply were not around. The idea of cities dominated by automobility must be similarly unthinkable for his grandson.

How did it all happen, he wonders, hailing a riksja to still make it in time for his coffee appointment a quarter of an hour from now. He remembers that petrol prices were on the rise for years already, until it became such a scarce good that it could only be used profitably in high end specialised petrochemical industry; simply burning it quickly became ridiculous. In those times we still thought we'd be using bio fuels to replace their fossil predecessors, he thinks. But the political instability created by the coupling of food prices and fuel prices and the wars that it provoked urged us to seek other solutions as well. The riksja driver asks him for his destination and he replies by giving the name of the *chique* coffee bar in the second centre of the large city on the other, southern, bank of the river. Within the cities things were already changing, out of necessity but also out of want. Dense traffic in the city centres rendered them awful places to live in and unhealthy also, ironically the transport sector itself suffered greatly from congestion problems as it was virtually impossible to get a product from here to there. Personal mobility changed rapidly he recalled. All of a sudden it was *en vogue* to be transported instead of to drive around in your own car. Riksja's, bicycle-taxis, tuk-tuk, one couldn't keep up with the fashionable new modes of urban transport, even though they were old modes upon closer inspection. Especially the combination of clubs and restaurants picking up and dropping off their clientèle gave it a boost. The health problems related with fine dust gave a push from the other side, government decreed that city centres needed to gradually become 'carscarse' zones.

What also made a lot possible, he now thinks, were that things were changing in other parts of society as well. How energy was produced and stored changed greatly in those years too. Descending the bridge, enjoying the reflection of the sun in the river, he informs the riksja driver that his destination is just two blocks away. Some innovations in that area almost directly changed the appearance of the city. Electricity already made almost all other energy carriers obsolete once nuclear fusion suddenly broke through, combined with solar electricity for domestic use. But now that storage of electricity was miniaturised and electric engines more powerful than ever, the internal combustion engine was completely eradicated as a factor in traffic. But by then cars were already rarely seen in the cities, and the car ban stayed in effect although the health argument was no longer valid. Apparently the societal need for automobility just disappeared. Who could have imagined that ever happening?

Fantasy, certainly, but also certainly a transition during which mobility changed profoundly. Well then, imagine that agriculture is practised by farmers that produce for their direct environment, having a limited amount of livestock

## 1.2. ABOUT SOCIETAL SYSTEMS AND TRANSITIONS THEREIN

and growing various crops. With the troubles that come with such a limited system, which swings between overproduction and shortages. Then imagine that such an agricultural system in a couple of decades goes through a transition by intensifying, by employing technology and sophisticated distribution systems, aided by a government that created institutions like agricultural universities, gives subsidies and alters legislation. Imagine that it is hard to imagine how different agriculture was some decades ago.

If one can imagine this then it is likely that several different transitions now come to mind in various systems. Probably the development into the nowadays specialised healthcare systems has witnessed transitions and is it not a transition if a city once thriving by its harbour and its industry has become a centre for the financial sector and a major touristic destination? Surely revolutions and *coups d'état* are transitions and would it not be a transition if we would be able to have a sustainable energy system? This small brainstorm raises the rather fundamental question: what kinds of systems are these that one can imagine going through transitions? Before attempting to address the matter of what transitions are — which is what the rest of this book is about anyway — it appears then only natural to ponder somewhat on the systems they affect. The term societal system already came about earlier in this book as the generic name for the relevant systems. So what are these societal systems?

### 1.2.1 Societal systems

The transitions just passing by are apparently supposed to occur in systems that serve some purpose. Moreover they appear to be systems that are actively maintained to do so. Be it how people enable themselves to move about from one place to another, how they arrange for illnesses to be cured or public matters to be addressed, all these systems function to meet some societal need. The societal need defines the societal system. No clear system boundary can really be drawn between societal systems as societal functions often overlap and, moreover, depend on the societal need one is interested in. If one speaks of a region as a societal system then mobility aspects can be a part of it, whereas one could also be interested in the mobility system in that region as such. In thinking along these lines several societal systems readily come to mind: agricultural, education, healthcare and mobility systems; cities, harbour areas and nations; spatial planning, policy and legislative systems and probably various others.

Any such system, as said, is defined by the societal need it meets. More precisely put it is defined by the *way* it does so, the way it functions, so to speak. The functioning of a societal system with respect to the societal need it is supposed to meet enables some refinement of the concept. Although a social security system and an energy supply system both serve a societal need — *which*, is suggested by their respective names — they appear to be of a rather different kind. Similarly, the harbour area of Rotterdam and the policy system around the water management of the Rhine delta appear very different, though they have this river as a common denominator.

Thinking of the various forms of systems that would qualify as societal systems, it is not at all hard to produce a tentative classification, or hierarchy even, based on the societal need met and the way it does that. For instance, there are systems that fulfil a societal function while remaining ‘close’ to the natural resources they use. A region would be of that kind, or a mining area. One



## CHAPTER 1. INTRODUCTION

level of abstraction higher appear systems that are more or less defined through the infrastructures or the ways the resources are exploited, like energy supply systems or railways. Then systems appear around products, their production, distribution and consumption. In other words markets and industries viewed as societal systems. From these systems of production and consumption it is only a small step to systems around services like healthcare or education. Then the meta level comes in sight with systems concerned with the functioning of other systems, that is, systems of regulation like policy and legislative systems. A classification along this line of thought is summarised in table 1.1.

---

### Types of societal systems

---

#### Natural systems

Functioning defined by the natural resources involved, be they resources proper (ore, wood) or the geographical location (a river delta, a city). The geographical location therefore is often characteristic for these systems as the need they meet depends on it but also for logistic reasons (centrally located, for example).

*Examples:* Regions, fishing grounds, urbanised areas, rivers and deltas, forestland, countries/nations, harbours.

#### Infrastructure systems

Functioning defined by employed technology and infrastructure. Still very close to the natural resources, however more concerned with the way these are exploited and put to use than their sheer presence or availability.

*Examples:* Energy supply, communication, transport, railways, car-based mobility (roads, gas stations).

#### Sectoral systems

Or *industries*. Functioning defined by products, their production, distribution and consumption. Typically systems where societal need is situated in a market context.

*Examples:* Agriculture, music industry, fashion, consumer electronics, automobile industry.

Continued on next page . . .

## 1.2. ABOUT SOCIETAL SYSTEMS AND TRANSITIONS THEREIN

---

### Types of societal systems — *continued*

---

#### Service systems

Functioning defined by the service involved; Societal needs are met by delivering services, rather than products. Arguably the archetypical societal system in the sense that these also include the systems providing public services. Not publicly or state organised *per se*, organisation via markets or other solutions possible as well.

*Examples:* Healthcare, social security, education, housing, mobility.

#### Regulatory systems

Functioning defined by kind of regulation involved; systems concerned with the functioning of other societal systems. These systems are about decision making, planning, representation and regulation, in a sense societal meta-systems.

*Examples:* Policy, politics, spatial planning, law, economy.

---

Table 1.1: Types of societal systems

In reality however, no societal system will neatly fit one of the classes. This is often the case with classifications, but even more so when it comes to societal systems. The functioning of a societal system is produced on various levels of aggregation and can be considered on various levels of abstraction. This is not to be confused with the problems concerning the demarcation of societal systems and overlap between them, rather it has to do with the complex build-up of societal system: elements and agents can be heterogeneous and interact on and across different scale levels. Societal systems are complex, adaptive systems and their functioning is the emergent phenomenon they produces.

If table 1.1 is interpreted as ascending a hierarchy, raising to a higher level of abstraction with each step then it is very much a hierarchy of the ‘structural’ side of societal systems. The classification is in terms of structures and institutions, in general in terms of *how* the system is organised. Another angle towards societal systems and their functioning is also very well possible, that of the meaning that is attributed to them. And also *by* them, since societal systems are social systems and as such human systems. If this attributed meaning is interpreted as ‘sense-making’ then this might be referred to as the ‘cultural’ side of functioning (using Geertz’s (1973) use of the term culture). This division into a structural and cultural side to functioning will be elaborated later on when description of societal systems in terms of the triplet structures, cultures and practices is discussed. The dualism is the same as that of the interactive and reflective part of functioning introduced in (de Haan, 2007). In turn, the cultural side of functioning could be used to classify societal systems as well, using the way systemic properties are culturally framed for instance, how they are valued or their legitimacy.

With these two sides to fulfilling societal needs, a ‘plane’ of possible classes of societal systems exists. Consequently, at least the same amount of classes

## CHAPTER 1. INTRODUCTION

of societal transitions exists as well. One for each kind of societal system and then some. Such a classification of transitions will not be attempted here. In the rest of this book transitions, and the processes that are part of them, will be considered that are possible in *any* societal system.

### 1.2.2 Transitions

Now that an image has been sketched of what societal systems are, the matter of transitions in them can be discussed more thoroughly. It is appropriate to start with a tentative working definition like:

A transition is profound change in various or all aspects of a societal system's functioning.

This working definition is along the lines of the broad ones that can be found in for instance (Rotmans *et al.*, 2001; Rotmans, 2005). A more strict definition will be presented in the next chapter on language.

In defining transitions three matters apparently inevitably remain ambiguous and they all have to do with scale. The first is that of the demarcation of the societal system itself. The second has to do with the 'profoundness' of the change. The third has to do with the timescales at which the transition is considered.

#### Demarcation

How one draws the boundary of a societal system has consequences for what one is going to consider to be a transition in it. A transition in a certain societal system can, equally justifiable, be said to be a relatively minor shift in a larger societal system of which the former is part of. For example restructuring the financing of public healthcare would be a transition in healthcare financing, whereas the healthcare sector as a whole would not be functioning all that differently. This is also apparent when a socio-technical transition, say a switch to bio fuels for automobiles, is viewed in the context of a larger societal system, the mobility system which would still revolve around car-based, personal mobility.

This ambiguity is on the one hand fundamental but on the other hand resolved if a demarcation is chosen. This is, however, not always a straightforward matter. The two apparently most different approaches to coming to a definition of what actually *is* the system are the external analyst conceptually modelling it, and a group of people from within the societal system doing so in a participatory process. In both cases the main problem remains choosing what is in and what is out. Therefore it seems that the *basis* for these choices, what the system is, should *not* be in terms of the collection of elements that constitute it, but rather in terms of the functioning it produces. The leading question in demarcating a societal system in this fashion then ought to be: how does this system fulfil its societal need? How does this healthcare system care for the health of people? How does the mobility system in this region fulfil the needs for transportation? In what terms — actor networks, infrastructures, social institutions, legislation — these questions are answered is then, in a sense, of secondary importance.

## 1.2. ABOUT SOCIETAL SYSTEMS AND TRANSITIONS THEREIN

### Profoundness

The second matter of scale has to do with the ‘profoundness’ of the change. When is it a transition anyway? How much has to change and in what way to be able to speak of *transitional* change. For this it is convenient to exploit the dualism of structures and cultures in another sense.

Since both dimensions of functioning, the structural and cultural side of meeting a societal need, can be thought of as a spectrum, one can think of the part of this spectrum that is changed in a transition as a measure of the ‘profoundness’ of the change. Let this be called the *depth* of a transition. For instance, a transition where all fossil fuel would be replaced by bio fuels would be a major change in the energy or mobility system indeed, however, considering the entire spectrum it would only involve changing things concerning some technological facets of functioning, even leaving most of the infrastructure intact. On the other end one can think of a transition to sustainable housing, which might involve changing facets of functioning varying from the spatial planning policy to the road plans, not to mention building or rebuilding houses themselves.

Although a measure of the ‘profoundness’ of a transition this might be, another related scale aspect has not been considered yet, something which could be referred to perhaps as a transition’s *magnitude*. Exploiting the above fuel transition example once more, a small magnitude transition to bio fuel would, for instance, be limited to a certain city, or to public transportation only. In other words the magnitude of a transition has to do with the scale of the societal system itself, or the proportion of it that is affected.

The *impact* a transition has, or would have, needs to be assessed both in terms of its depth and its magnitude. The impact of a low depth, high magnitude transition, say achieving world wide carbon emission reduction, might very well be larger than that of a deep one, with low magnitude like a village becoming completely climate neutral<sup>2</sup>.

This ambiguity cannot be resolved and the question of what is profound enough a change to be called a transition is essentially a political one. What for one person is a mere optimisation of the status quo is a radical change for another. Although, apparently unresolvable, the ambiguity in any case concerns dichotomies like these. Some adjectives then. Transitional change is radical for instance, fundamental, profound, deep, structural, transformative, qualitative<sup>3</sup> and not, for example, merely incremental, quantitative, evolutionary or gradual.

### Time

The third matter of scale that introduces ambiguity in the concept of transition is that of time. In twofold manner in fact. Analogously to the ambiguity introduced by the system’s demarcation there is ambiguity that comes with the temporal demarcation. By choosing a beginning and end point for a certain transition, which is at times unavoidable, for instance in most historical case studies, one forces upon the analysis not only a time scale but also scales of depth and magnitude. Simply because the ‘amount’ of change will almost inevitably be

---

<sup>2</sup>Nothing said here about the possible effects such a transition would have by setting an example or becoming an icon of change, of course.

<sup>3</sup>In the sense that the nature of something changes and not merely becomes a bit more or less

## CHAPTER 1. INTRODUCTION

cast in terms of the change that actually took place in the timespan considered<sup>4</sup>.

The other, perhaps more fundamental way, time comes into play is when it comes to the pace of change. The concept of transition carries a connotation of abruptness, and one assumes that transitional change disrupts the normal cause of events. A graphical-intuitive representation of this would be a graph of an indicator of system state versus time going from convex to concave, an s-shape. Indeed, transition-like phenomena such as the demographic transition (Davis, 1945) or innovation diffusion processes are often represented by s-shaped curves or logistic curves. In transition studies this is called the multi-phase picture (Rotmans *et al.*, 2001) and it is customary to distinguish at least three phases, in succession: pre-development, acceleration and stabilisation. Between pre-development and acceleration a point of take-off is identified and similarly a point of touch-down could be recognised between acceleration and stabilisation (Timmermans *et al.*, 2008). In any case these points that separate the phases correspond to the nodes of the third derivative of the curve that plots the system state over time, for these are the points at which the pace of change changes maximally.

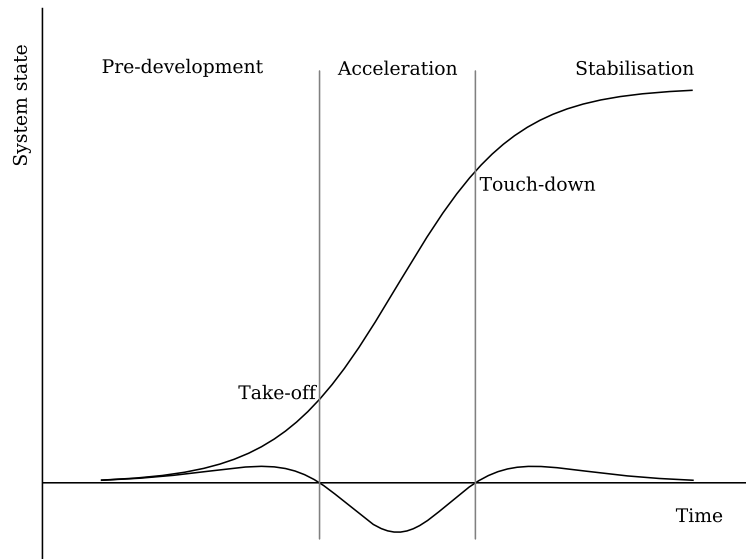


Figure 1.1: An s-shaped curve with 3<sup>rd</sup> derivative and phase demarcations.

But some issues remain. Although the convex-to-concave criterion *in principle*<sup>5</sup> provides a way to recognise transitional change, one can still imagine s-curves for small parts of the state of the societal system or for the entire societal system. Transitional change on, say, the level of one household is not

<sup>4</sup>Although this is not strictly necessary and especially comparing similar cases would probably mitigate this somewhat.

<sup>5</sup>If some suitable indicator can be identified, obviously.

transitional change on the level of the built environment system of a certain city. Moreover there is the appreciation of the change involved, convex-to-concave change affecting the entire societal system still needs not always to be considered transitional.

## 1.3 About pillar theory

Transitions and societal systems are what pillar theory is about, so it appears time to do some more thinking about pillar theory before actually unfolding it. How does this theory treat transitions? How does it conceptualise societal systems? In what *way* does it provide understanding? How does it address the peculiarities of its subject matter that were discussed in the previous section? And how does it relate to certain paradigms, theoretical schools of thought and similar theories?

As alluded to earlier, pillar theory conceptualises a transition as a profound shift in the way a societal system meets a societal need, its functioning. Since there are several, often qualitatively distinctive, ways in which a societal need is met, a societal system is a composite, a patchwork of societal *subsystems* meeting needs in various ways. These subsystems are referred to here as constellations and much in the theory revolves around them. Each constellation contributes, as it were, to the functioning of the entire system. Think for instance of a mobility system with a constellation around public transport and a constellation around car-based personal mobility. Constellations can be small, representing novel or deviant societal practice, or they can, comparable in scale to the entire system, represent the societal status quo. Or anything in between, continuously scalable. A constellation itself, being a societal *subsystem*, can in turn be thought of as composed of other constellations as well.

Constellations are complex, adaptive systems and their dynamics and interactions produce the dynamics of the societal system. This view on societal systems resolves at least some of the issues raised in the previous section regarding scale. Moreover, it eliminates the need for a rigid level division such as employed in various emanations of the multi-level perspective (e.g. Rip & Kemp, 1998; Geels, 2005*b*), while simultaneously allowing the conceptual language of niches and such still to be used.

Another consequence of this view is that the societal system is considered to be an open system, enveloped in an environment which is referred to as the landscape. In such a complex, adaptive system change can be the result of internal dynamics and of interactions with the environment. The theory argues, however, that transitional change only occurs under certain *conditions*. These conditions express that the meeting of the societal need is somehow being compromised, or that the functioning is somehow at cross with something. These conditions are symptoms of persistent problems (Rotmans, 2005), symptoms of unsustainability (de Haan, 2008*b*).

In the theory the conditions for transitional change are tension between the societal system and its environment, stress internal to the societal system or its constellations, and pressure due to upcoming alternatives. Though these conditions are considered to be driving transitional change in societal systems, the processes and mechanisms producing that change are treated separately. This is further elaborated by assuming that either a dominant constellation,

## CHAPTER 1. INTRODUCTION

commonly referred to as the regime, radically<sup>6</sup> changes, or that another constellation takes over dominance. In either case leading to the societal system as a whole assuming a qualitatively different functioning.

The functional reorganisation of a societal system is obviously a complex matter, with many processes and mechanisms involved. To avoid the inevitable getting lost in this, pillar theory approaches transitional change from a different angle. It proposes that transitional change is the emergent result of various processes and mechanisms occurring in a limited number of identifiable *patterns*. These patterns describe how societal systems, or rather their constellations, are affected by transitional change, they describe the mechanics of transitions. Each pattern is, as it were, a part of a transition story. By concatenating patterns it is possible to reconstruct the complex dynamics of a transition from relatively simple building blocks. Each pattern works under the influence of one or more of the aforementioned conditions. How the conditions drive patterns of transitional change is cast in the form of a set of to-be-tested coupling hypotheses, apart from the core of the theoretical frame. In this manner the theory itself can adapt to new insights and data.

Such chains of patterns produce transition *paths*. Depending on how the theory is implemented and to what end, transition paths can be interpreted in various ways. They can be the outcomes of a computer model or a path in the phase space of a set of differential equations representing societal dynamics. They can be narratives of a historical case study or other transition story lines. They can be used to investigate the conditions that made a transition take the course it took, or assess what policy measures seem appropriate given the patterns currently in motion under the present conditions.

Viewing a transition path as a chain of patterns eliminates, at least partly, the problems that arise through temporal demarcation. Choosing a timespan in a transition study will at times still be inevitable, but rather than investigating and defending a certain beginning and end point one can now focus on the patterns of transitional change that make up this transition, regardless of at what point, or if at all, it ended. This is of course also relevant when studying *ongoing* transitions.

Pillars is a constructed theory. The subject matter demanded that existing theoretical, empirical and intuitive knowledge about societal transitions was cast in an integrative theoretical framework. What results is inevitably an eclectic theory, drawing from various disciplines and traditions and employing a variety of concepts. Thus one could say that it is a systems theory because it treats societal systems; that it is functionalist since it talks about functioning; one could consider it to be an exponent of the punctuated equilibrium paradigm, as a revolutionary change theory; one could call it a complexity theory or a quasi-natural scientific one. And, all with some justification. Rather than trying to pigeonhole pillars in one-of-the-above it is more sensible to discuss some of them in some depth, to explore what is the link with this theory and to address some of their criticisms in relation to pillars.

---

<sup>6</sup>What exactly constitutes radicality, and thus sets apart transitional change from other forms of societal change is a recurring and complicated matter, which will be explored in some depth in section 3.1.3.

### 1.3.1 Systems theory

The term societal system already suggests that this is a systems theory. There are quite practical reasons that make this appropriate. As said, pillars is a constructed theory with elements from various theories and disciplines that needed to be brought together into a consistent whole. Synthesis and integration have always been keywords for systems theory. Even as early as in von Bertalanffy's (1956) work it is about bringing together different areas of science, different forms of knowledge in a 'system'. Necessarily then, a systems theory in a pure form is conceptually empty, like set theory in mathematics or logic. This is especially apparent in general systems' definitions like in (Hall & Fagen, 1956). Systems theory itself therefore is arguably not a theory proper, rather it is meta-theory, a way to structure theories. Systems theory is an attempt to a standardised ontology<sup>7</sup>. In this sense there is nothing intrinsically good about theories that explicitly or implicitly are systems theories, nor are they better than theories that are not, it only says something about their *physique*.

The use of systems theory in pillars also served to distinguish clearly between what undergoes transitions, the processes and mechanisms that produce them and the conditions under which they do so. It was a way to separate the theoretical framework, — that is, the concepts and their relations — from the hypotheses, the assumptions on what leads to what, that make it into a theory. It makes the theory intrinsically *modular*, which is a necessity because: first, if a hypothesis is proved wrong it ought to be possible to discard it without damaging the rest of the theory. Second, because it is not always desirable to have to use *all* of a theory in a case study, for a model, an analysis or whatever one wants to use it for. Third, modularity allows it to be elaborated, built from and on, allows it to be added to or become part of something. This is the kind of modularity which has recently been advocated as 'social mechanisms' (Hedström, 2005) in the world of social simulation, system thinkers by nature.

### 1.3.2 Functionalism

Speaking in terms of societal needs and functioning renders pillars an overtly functionalist theory. That the theory has systems as its elementary objects can be seen as an indication of its functionalism too. From Parsons's grand theoretical endeavours<sup>8</sup> onwards a 'system' has been introduced to attribute the functioning to. Not all theories that speak of systems are systems theories, however; discussing music is not the same as being a musician. Though it can be defended that most functionalist theories are systems theories in both senses.

In any case, a common trait in functionalist theories is precisely their view of society or parts thereof as systems that function to meet certain societal needs. Functionalism has been and still is one of the most successful and well known sociological paradigms. It has also, especially from the nineteen sixties on, been widely criticised. The two main points of critique that also affect pillars are the focus on explaining the stability of social structure — and with that its apparent inability to deal with societal change — and the limited roles for agency, actors,

---

<sup>7</sup>This attitude to systems theory is for instance advocated by Bunge e.g. in his 2003 work on emergence, among others.

<sup>8</sup>See for instance Turner (2003) on Parsons' work. To paraphrase Parsons: Who now reads Parsons?



## CHAPTER 1. INTRODUCTION

and individuals.

To start with the latter one could certainly maintain that this holds especially in the case of pillar theory since it is completely devoid of actors. It is a theory of societal systems proper and actors — in the sense of individuals — are only considered, and even then only implicitly so, as parts of constellations. The straightforward defence of this would be that this is largely a matter of scale and the scale of the individuals is not that at which transitions are studied here. The matter is, however, somewhat more subtle.

On the one hand the protagonists in the theory, constellations, are acting entities. This does not mean that these societal subsystems are personified. They have agency and can very well be considered actors, although they are not individuals, like for instance was done for the systemic entities regimes and niches in the simulation model by (Schilperoord *et al.*, 2008). In pillar theory proper, the agency of constellations is captured in the patterns. The patterns describe how the system changes, which is what the constellations do or what happens to them, under conditions that in turn depend on the relations between the constellations and the environment.

On the other hand pillar theory proposes that constellations do not simply have functioning as a property but rather produce it. This can be understood using Giddens's (1984) duality of structure and recognising that underneath such a social structure are the actors that through their actions — enabled and restrained by the structure — reproduce it. Another equivalent but more native way is to use a complexity argument and to understand functioning as an emergent property with the interactions of actors on the underlying level producing the structures and cultures on the higher level, coupled and looped back via the practices. This is also where the concept of emergence enters — more on that when the complexity aspects of the theory are discussed.

What about functionalism and its focus on societal stability rather than change? If pillar theory has some functionalist traits, how does it deal with this potential inability. If the whole theory is built up from constellations with a certain societal functioning then how is change, let alone a transition to take place?

This in fact is a pivotal point in the way transitional change is understood in this theory. Non-transitional societal change is considered as evolutionary, gradual and incremental and this fits well within the functionalist constraints. When conditions for transitional change are present however the system changes abruptly by the workings of a transition pattern. Since it explicitly addresses these punctuated equilibria the theory could also be regarded as a revolutionary change theory like Gersick (1991) would put it. In other words, in the absence of conditions for transitional change the theory works smoothly within the functionalist paradigm and if transitional change happens to it, the equilibrium is punctuated and the system jumps to a different state to resume its gradualism. Which appeals to the intuitive ideas of what a transition is.

The relation with functionalist theorising also places the theory in a certain tradition. It is far beyond the scope of this introduction to give even a brief overview of functionalism. The work of two theoreticians that are especially relevant for pillars will be briefly discussed: Merton and Luhmann, both of whom are students of that other eminent functionalist, Parsons.

Merton's work (Merton, 1968) is important to pillars since it acknowledges that in a societal system not all parts necessarily function to meet the same need,

### 1.3. ABOUT PILLAR THEORY

and that different parts can do so differently. There is room for dysfunction and deviancy, conceptually leading a way to the idea that a societal system is a composite of several constellations. Apart from this he was a strong advocate for theories of the middle range. These are essentially theories that hold the middle ground between the grand theories of Parsons or Giddens that attempt to encompass all things social and the other extreme where one stays at the level of statistical interpretation of raw data for instance. In both cases explanation, which is the main goal of theory, is lost, either in abstraction or in lack thereof.

Luhmann is one of those grand theorists, especially in his 1984 work on social systems. Its extreme level of abstraction, combined with a particularly nebulous writing style — in which he took great pride — renders this work quite unreadable, though the basic ideas are fairly simple. It is exemplary functionalist in that the individual is completely abstracted away. Central are social systems that reproduce themselves as systems of communication. These systems define themselves by maintaining a boundary with their environment, within which they attempt to reduce the complexity they perceive in that environment. This way of dealing with systems in interaction with an environment and their desire to reduce complexity, was an inspiration for the idea of conditions for transitional change.

#### 1.3.3 Natural science

Obviously this is precisely what pillar theory is not. It is not natural, arguably not science, let alone natural science. However, it does as a theory bear resemblance to natural science. This has to do with the way it deals with causes and effects. The distinction mentioned above, between the mechanisms behind transitional change, the conditions under which they work and what they work upon, is typical for natural science theories, more specifically dynamical systems approaches. Classical mechanics, for instance, describes how point particles move about under influence of forces, and how the particles react to the forces is governed by the dynamical equations of the system, typically some emanation of Newton's second law. That pillar theory mimics the structure of a natural science theory does not imply that it is a necessarily a quantitative theory. Nothing needs to be expressed in numbers to work with this theory. It is just that the theory is formulated in terms of concepts with certain properties, upon which certain mechanisms work under certain circumstances.

Especially in organisational studies, when it comes to describing change, often a different approach is used in which change is described as a sequence of events rather than in terms of changing attributes of an entity. This approach is dubbed process theory whereas the other one, that of the changing attributes or dependent variables, is referred to as a variance theory, see for instance (van de Ven & Poole, 2005; Pettigrew, 1997; van de Ven, 1992) for accounts of these approaches. Geels & Schot (2007) argue that narrative explanation, is essentially always process theory since it explains in "terms of *patterns* that result from interactions" (original emphasis). This is not necessarily true though, since it is easy enough to produce a — admittedly rather dull — narrative relating of the whereabouts of a point particle under influence of some force. Such narratives can — and are, even — be produced automatically from positively variance theory based agent-based models. The point here is that the distinction of theoretical approaches into process or variance ones does not apply well here

## CHAPTER 1. INTRODUCTION

and that pillar theory, although structurally akin to natural science theories and as such apparently a variance one, can be used for narrative explanation, thus taking the guise of a process theory.

This is possible by means of the way that the mechanisms are presented. Mechanisms in what might be referred to here as a ‘pure’ variance theory only ‘narrate’ when they are at work on the objects they are defined to work upon. In other words, Newton’s second law only comes to life when the particles and forces are specified and the equation is solved producing its trajectory<sup>9</sup>. In pillar theory, however, the mechanisms —they are even called *patterns*— are formulated as abstract pieces of transitions, chunks of a generic transition story line. Such a piece of transition story narrates about a constellation, which is a part of a societal system, being affected under influence of certain conditions. In this way the patterns are simultaneously the ‘equations’ governing the dependent variables describing the societal system, as well as the patterns in the process theory sense, that is parts of the sequence of events.

### 1.3.4 Complexity theory

The term ‘complex’ has dropped by a number of times now and the question arises whether pillar theory is a complexity theory. This has to be admitted, though with some hesitation. Complexity, especially in organisation sciences, has time and again been a label for theories promising explanation where none was to be given and has been a promise that was impossible to keep. Complex systems’ science<sup>10</sup> was a new interdisciplinary field of research where parallels and methods from mathematics, physics, chemistry, biology, and ecology were exploited. It dealt, and still deals, with non-linear dynamics, pattern formation (Cross & Hohenberg, 1993; Bar-Yam, 1997) and chaotic systems (Strogatz, 1994). It grew rapidly with the advent of increasing computing power and new insights in information and computer science (Shalizi, 2001), which in turn inspired some interesting philosophical debates on concepts like emergence and self-organisation (Bunge, 2003; Crutchfield, 1994; de Haan, 2006; De Wolf & Holvoet, 2005).

With the initial successes of the Santa Fe Institute and the school of thought that was emerging with people like Holland (1995), Kaufmann (1993) and Gell-Mann (1994) who were already suggesting that the ideas of complexity might be applicable to human and social systems, as did Nicolis & Prigogine (1989) who added the — for this theory interesting — observation that complex (adaptive) systems have an endogenous ability to go through transitions. Complexity theory became somewhat of a fad in organisation studies, with some researchers sceptical but hopeful (McKelvey, 1999) and other embracing it fully (Stacey, 1996). For organisation science, and perhaps for much of social science, using complexity theory can be problematic for several reasons. To begin with, there is no one complexity theory, like there is a relativity theory for instance. Complex systems’ science consists of a variety of methods applied on a variety of subjects, maybe with nonlinearity as the only real common denominator. A

---

<sup>9</sup>And even *this* is not entirely true since the equations themselves speak books, if the stretching of this metaphor be allowed here.

<sup>10</sup>Often complex systems’ science is viewed as distinct from complexity theory. The latter is regarded as also including the more ‘soft’ approaches, whereas the former is considered the ‘harder’ science.

### 1.3. ABOUT PILLAR THEORY

large problem also lies in the very phrase “using complexity theory”. What does *using* mean here exactly? For many scholars, for instance, it meant to say that the system they studied was a complex one and hence the processes they saw were processes of emergence and self-organisation. This is sometimes referred to as the metaphorical use of complexity. It is similar to saying that one’s research subject is relative so one can use relativity theory. Nevertheless, this metaphorical use inspired hordes of scholars (Byrne, 1998, for instance) and management guru’s (Senge, 1990, for instance).

The rise of agent-based modelling as a computational approach to social science and the associated emerging field of social simulation sparked a new, more seriously scientific complexity boom, however (Gilbert & Doran, 1994; Gilbert, 2008). In this field the problems of metaphor are of no significance since the complex systems that are studied are rebuilt as a computer model, which is often called an experiment *in silico*. The computer model is as a lab to these researchers, where they have control over the behaviour of the agents that make up their artificial societies (Gilbert & Conte, 1995). Typical for agent-based approaches is that the agents are fairly simple, as are the rules they follow, their decision rules, and that from this many of the complex patterns in real society can be reproduced. This way of pattern formation is what simulators refer to as a process of emergence.

This is precisely how complexity is important to pillar theory and why it can be regarded an example of a complexity theory. As with pillars being a systems theory, this has more to do with the form of the theory, the structure, than with the concepts and terminology used. Pillar theory is meant to reproduce, via modelling, narratives, or other ways, certain complex patterns, namely societal transitions. Here too the building blocks are conceptually fairly simple, the agents correspond to constellations and their behaviour is governed by ‘decision rules’ in the form of patterns affecting them under certain conditions. Some ambiguity appears to have slipped into the word ‘pattern’, since here it is meant in the sense of societal mechanisms or decision rules, whereas the emergent patterns referred to earlier correspond to the ‘paths’ of pillar theory.

Here one sees that complexity and the concept of emergence are central to the theory, without it being necessary to put it as a label on every second process encountered. One speaks of emergence when systems have properties its constituent parts do not have as such<sup>11</sup>. In other words: The whole has obtained a quality that the parts do not have, though it is in a certain sense produced by them. A very simple example is the property of colour. Atoms in a crystal do not have the property colour though it is these very atoms, due to their being part of the system of the crystal-lattice that produce the quality colour of the crystal as a whole. In much the same way one can speak of emergent phenomena, dynamic properties of systems, which are then often referred to as patterns.

An emergent phenomenon is in principle detected by an observer, and like in the case of colour, the system need not notice its emergent property. Sometimes an emergent phenomenon does have meaning to the system that exhibits it. In this case the emergence is called type-II, whereas in the former it would be type-

---

<sup>11</sup>The literature on emergence is vast and ever growing and a plethora of definitions is out there. For the view of the author see (de Haan, 2006). For extensive literature review see (De Wolf & Holvoet, 2005). For philosophical depth and discussion the reader is referred to (Bunge, 2003; Bedau, 1997, 2002; Emmeche *et al.*, 1997).

## CHAPTER 1. INTRODUCTION

I. An example of a type-II emergent phenomenon would be convection patterns in heated water. The pattern is an externally observable new property of the fluid system as well as serving to transport heat in it. Another example would be the emergent property of exchange value of a commodity (de Haan, 2006). In the context of trade and exchange a commodity obtains a value, which is simultaneously important in the very dynamics that determine it. A type-III emergence can also be distinguished, this is when the objects or agents that produce the phenomenon through their interactions observe it themselves and can act according to it. Social systems are of this type.

Emergent phenomena are *robust* with respect to what produce them. Though caused by objects and interactions on some underlying level, the emergent phenomenon has a certain autonomy with respect to it (de Haan, 2006). If societal systems and constellations are considered complex systems with their functioning as an emergent phenomenon, then this is exactly why it is allowed to abstract from the level of the individual, from the constituent parts of societal systems. Although the term emergence is not used by Giddens, structuration theory (Giddens, 1984) is in fact a description of a mechanism of producing the emergent phenomenon of social structure.

Levels and emergence go hand in hand (Emmeche *et al.*, 1997) and levels of all sorts are common in the social sciences. In transition studies the conceptual toolbox almost always contains the multi-level perspective (Rip & Kemp, 1998; Geels, 2005*b*), although pillar theory moves away from precisely this use of levels. Not all these levels are levels of emergence though. For instance, the micro-macro levels of Coleman (1990), which are often employed in the social simulation community, are not. The macro level there is the aggregate of the micro level which means that its properties can be straightforwardly reduced to sums or composites of micro level properties, for which no-one needs a concept of emergence. The multi-level perspective of the socio-technical transition literature does not contain levels of emergence for other reasons. There the micro, meso and macro level simply contain different entities, differing in role, function or scale — whatever is relevant. The properties on the higher levels are neither caused nor produced by those on the lower levels, just as they are not their simple aggregates.

### 1.3.5 Audience, scope, ambitions and limitations

A great deal more could be told about what kind of theory pillars is and its positioning with respect to other theories, paradigms and such. How it relates to evolutionary theories of social and economical change for instance. Or to innovation studies, theories of public administration or more broadly, policy analysis and organisation studies. About it being a revolutionary change theory and its relation to other work under the punctuated equilibrium paradigm. This all will not be done here.

At this point it is more interesting to know for whom this theory is meant. As already alluded to in the opening section the intended audience is anyone that wants to understand transitions. One of the very reasons this theory came into being was that transitions became an object of study in their own right. Thus pillars aims to be a general theory of transitions. This suggests an unlimited scope but it is as well a statement about its limitations. On the one hand, ‘general’ is to be understood in the sense that it is not intended for transitions

### 1.3. ABOUT PILLAR THEORY

on a certain scale, or a certain sector. On the other hand, the theory explicitly limits itself to transitions in societal systems. It is not a grand theory of the social, nor of social change, nor of societal systems. It is not a general theory of change in societal systems. It deals with transitional change in societal systems, which still is a lot to deal with.

Still, transition researchers come in various forms. Some work by gathering statistical data to understand transition processes. Some use narratives on historical cases. Theorising is being done. Models are being made, conceptual, computational and mathematical ones. For all these researchers pillar theory aims to be instrumental, literally in that it aims to be used as a tool. Transition studies have rapidly grown into a fruitful research field with the ambition and possibilities of making its own transition to become a transition science. It is the author's strong conviction that scientific development and progress have in the past profited greatly by the triad of theoretical, empirical and modelling approaches reinforcing itself<sup>12</sup>. Thus, pillars aims to be a versatile theory, to be an instrument for various forms of transition research and their vehicle to transition science.

Much of the theory presented here is new and, if not, often brought together into one framework for the first time. This is not to say that it does not build from the work of transition pioneers like Rip, Rotmans, Geels, and several others. These scholars laid the foundations of transition science by letting integrated assessment (systems thinking, modelling, integral approaches to societal problems) and socio-technical transition studies (innovation focus, multi-level perspective) come together. In doing so they brought along a lot of the concepts, terminology and nomenclature that is used in this theory. However, although the overall interpretation of the concepts remains roughly the same, it often proved necessary to provide new definitions of concepts, either narrowing down or broadening their original intended use or scope. As research continues old assumptions will be proven false and explanatory devices will need to be adjusted to remain explanatory. Pillars aims to be a flexible theory, to accommodate the progress of transition studies, rather than only be a snapshot of the present body of knowledge. It is intended to be theory at work.

These being the ambitions of this endeavour: pillars to be a general, versatile, and flexible theory of societal transitions, an instrument for transition studies on its way to becoming transition science. How can one know that it is any good?

Theory is technology. Something human-made to get a grip on reality. It is good if it is useful. If it is to live up to the above demands it needs to be applied. In various contexts, in various ways, by various users. How it will hold up as transition studies progress cannot be foretold. In the meanwhile the thesis of which this book is a part contains three applications of pillar theory, or parts of it, as an indication that it works. It would be useful if interesting scientific questions can be asked with this theory. It would work if the answers found, even proving assumptions wrong, lead to improvement of the understanding of societal transitions and of the theory itself.

---

<sup>12</sup>This was exactly the aim of the 2007 workshop in the Lorentz Center of the Leiden University which led to a special issue on Computational and Mathematical Approaches to Societal Transitions (Timmermans & de Haan, 2008).

*CHAPTER 1. INTRODUCTION*

## Chapter 2

# Language



## CHAPTER 2. LANGUAGE

### 2.1 Beforeword

In this chapter the language of pillar theory will be presented, the words to discuss transitions with. As already mentioned, this is a constructed theory and many of the concepts it uses have entered via other disciplines or are widely used throughout science or daily life. This means that these words and concepts need to be redefined in the context of societal transitions. Even when the concepts are common to transition studies it is convenient to have a native pillar theory definition. In other cases it will prove necessary to introduce new words and concepts.

For ease of exposition and to do justice to the origins and originators, of the concepts transition speak is treated in two separate sections. The next section will straightforwardly present the definitions as they are used in pillar theory. The section following that one will wander and ponder on the origins and connotations of these concepts and their use in other research fields. The pillar theory definitions are printed in *italics* and gathered in a glossary at the end of this book.

### 2.2 The words

At this point the reader hopefully has an idea of what is meant in this book with a transition in a societal system. However, such concepts which are at the very basis of the theory deserve at least an attempt at a proper definition. Attempt? Attempt indeed because there is something problematic about definitions of fundamental concepts. A definition provides a unique description of a concept by relating it to other concepts that are either more fundamental or already defined on a similar level. Therefore a definition of a fundamental concept necessarily is circular, defining one with the other and *vice versa*. The other option would be to define these concepts in terms of even more fundamental concepts that in turn would require definitions, this would lead to infinite regress, which is even worse endless recursion of circularity.

Then to start with the beginning, some words on societal systems, the systems that go through these transitions that need to be understood better. With the noun 'system' the implication is made that the object of study can somehow be distinguished from its environment. This environment, judging by the adjective 'societal' is society<sup>1</sup>. Obviously people are an important aspect of society and since people have needs one can speak of societal needs. These needs can vary from the very basics, like food, clothes and shelter, to health care, education or for instance transportation. Let now a *societal system* be a part of society to which can be attributed a functioning and let *functioning* be the way a societal system meets a societal need.

For instance the societal need for health, as in physical well being, is met by health care systems. People make use of this system in various ways for various reasons. From buying aspirin for the occasional headache to undergoing a heart transplant to being submitted into a closed clinic for mental treatment, to going to a homeopathic practitioner or having ayurvedic massages. The societal sys-

---

<sup>1</sup>The reader will please be so kind as to forgive that *society* will go without a proper definition here and be content with the meanings generally associated with the word or take recourse to a dictionary.

## 2.2. THE WORDS

tem that one could refer to as the health care system is this interconnected whole of solutions with which society meets the societal need for health care. To put it more precisely, the ways in which the societal system does this, this meeting of the societal needs, is the functioning of the societal system and conversely, what produces this functioning is the societal system.

Now two questions emerge. The first: If a societal system is a part of society that apparently meets a societal need, then are there perhaps many societal systems meeting several societal needs? This can very well be the case and this leads to a picture of society as a patchwork of societal systems. It also leads to the picture that the environment of the societal system in which one is interested, consists of other societal systems. This patchwork of societal systems that forms the environment of one particular societal system is commonly referred to in transition studies as the *landscape*. It follows logically that the landscape, being composed of societal systems, can also be attributed with a functioning.

Consider now a mobility system, apparently there is a societal need for transportation. Both for goods and for people. The landscape, the environment of this societal system, is composed of other societal systems and in this case one can think of several production-consumption systems. An example could be the agricultural system dependent on the mobility system for the distribution of the goods it produces. Another clear example is the energy system, connected to the mobility system via fuel infrastructures and electricity production.

All these societal systems that form the landscape of, in this example, the mobility system are themselves enveloped in a landscape of societal systems. For this reason, more abstract influences on societal systems like global trends and environmental issues such as climate change are considered landscape influences as well, even though these influences sometimes cannot be properly attributed to a specific 'other' societal system. Quite often the term 'landscape influence' in transition studies is used as synonymous for 'exogenous'.

The second question emerging is the following: if a societal system meets a societal need, is it then conceivable that parts or facets of this societal need are met by certain societal *subsystems*? This is not unthinkable and such societal subsystems are referred to as *constellations*. In this manner one can think of the societal system as being composed of a number of constellations whose combined functioning produces the way the societal system meets a societal need. This links closely to Hout's (1969) definition of functioning where he speaks of "the contribution of any phenomenon to a larger system of which the phenomenon is part". In turn this also provides an interpretation of the power of a constellation. Power relative to another constellation, that is. Let the *power* of a constellation be the proportion of the total functioning that can be attributed to a constellation.

Some examples might be appropriate. If as a societal system one thinks of an agricultural system, then as constellations could be identified stock farming and agriculture itself as rather powerful constellations and, for instance, biological agriculture as a less powerful one. In a health care system one could distinguish between a 'care'-constellation (homecare, nursing, etc.) and a 'cure'-constellation (hospitals, physicians) and a financing constellation. A mobility system would in first world countries have a car-based personal mobility as well as public transportation constellations, and so on. There will always remain ambiguity in identifying both the constellations themselves as well as their boundaries. In any case the above examples sketch the idea of a societal system

## CHAPTER 2. LANGUAGE

being composed of constellations that all contribute to the functioning of the whole.

Constellations apparently have a lot in common with societal systems. They are societal subsystems and also are attributed functioning and there is in fact conceptually no real difference between the two. The difference lies in their application as the societal system defines the level and scale of analysis and constellations allow one to decompose the object of interest into smaller units of analysis. Constellations in turn could be further decomposed into subconstellations, though this might also cloud analysis when levels of description could be confused. The main point is, and this point will return a number of times, that what holds for societal systems *mutatis mutandis* holds for constellations and vice versa.

If the power of a constellation is the proportion of the total functioning of the societal system that can be attributed to it, then apparently ‘constellation’ is a scalable concept. On the one hand constellations are conceivable which contribute almost nothing to the total and thus are virtually powerless though having a distinct, perhaps very specialised function(ing). And on the other hand it is possible almost all functioning is brought forth by one very powerful constellation dominating the societal system to a large extent. This invites the idea of introducing a number of ideal types of constellations. Two ideal types are widely used: the regime and niches; and others are helpful additions that will be described here too. The concept transition itself will *en passant* be defined and redefined.

Such a powerful constellation as mentioned in the previous paragraph, dominating the functioning of the entire societal system is called the *regime*. It not only dominates the societal system, but is also typical thereof. As will be discussed further in the upcoming etymology section, the regime represents the mainstream, business as usual, the forces that be, conservatism and convention. If a societal system appears to be in equilibrium then it is the regime with which one identifies this equilibrium. It being representative for the societal system as a whole, and it being the most powerful constellation naturally implies that there be only one, and therefore one normally speaks of *the* regime. In colloquial transition speak the regime is often personified (as are niches) and regarded as the entity impeding and counteracting innovation.

This might appear a very monolithic interpretation of the regime concept, which it is. One could argue that even within the dominant societal practice there is room for deviancy and thus that innovation can spawn from the regime, this is for instance done by van der Brugge (2009). In pillar theory, however, such processes are treated separately, and although the same actors or institutions could be involved, the innovation is described with another constellation, precisely because it would have a different functioning.

It is not actually necessary that a societal system has a regime at all. It could very well be the case that a societal need is met by various constellations of which none is dominant, nor typical of the whole. For example, a mobility system where various modes of transport are available and used, such as some cities strive for. Another way in which a societal system can be without a regime is when it is actually *in* a transition, though this might also be considered the exception confirming the rule. In most of transition studies the existence of a regime is tacitly assumed, which has both historical reasons (again, await the etymology section) and practical analytical ones since it provides a clear

## 2.2. THE WORDS

perspective for the transition researcher. In most of this book the regime will be regarded as one constellation, but especially when the representative aspect is relevant it is sometimes convenient to refer to a cluster of closely related constellations as the regime, which is done for instance in chapter 1 of Part III.

If the regime embodies the powerful, yet conservative mainstream, then a *niche* is its innovative, avant-garde but not so powerful counterpart. At least this is the colloquial use. Reusing the constellation examples in the above one readily identifies biological agriculture as a niche and car-based mobility can in many cases be thought of as the regime. Niches are of particular interest to transition studies since one of the ways a transition could come about is that a niche becomes more powerful and either influences the functioning of the regime or assumes a more dominant role in the societal system itself. This process of a niche<sup>2</sup> gaining power is called *scaling-up*, like described in (van den Bosch & Taanman, 2006).

It appears that it is time to deal with the concept of transition itself. Since various descriptions, intuitions and connotations have already been discussed it is best to simply state a working definition here. Very well, let a *societal transition* be the radical changing of the functioning of a societal system from one state of apparent societal equilibrium to another. If this be a satisfactory definition then an equivalent — and very compact — one can be given: a *societal transition* is a radical change of regime. The adjective ‘radical’ is superfluous even, as will become clear in section 3.1.3 where is discussed how transitional change is necessarily radical.

What is still left open in these transition definitions is whether the regime changes, is changed or if another constellation becomes dominant. The latter invites the introduction of a new ideal-typical constellation. If novelty is to emerge in niches as suggested in the above and such a niche scales up and becomes a viable, competitive alternative for the regime then at a certain point it must lose its nicheness. It has become a *niche-regime*. In this view a niche-regime is a transient constellation, growing out of a niche and on its way to become the new regime. It is of course also possible that a niche-regime does not make it that far, or that it was never supposed to. In such a case a niche-regime could very well co-exist with a regime, think for instance of public transportation and car-based personal mobility.

Apart from these, other entities can be regarded as constellations as well, although perhaps less obviously so. Among these are transition arenas and transition experiments. Both concepts are in fact instruments in the governance framework called transition management<sup>3</sup>. As both concepts are instruments to bring about societal change they have an intentionally temporary nature and can be assumed to have little or no power, though this need not be an objection to regarding them as constellations.

A *transition arena* is a network formed by so-called frontrunners<sup>4</sup>, innovative minds with an aim to bring about societal change. In (Rotmans & Loorbach, 2009) transition arenas are discussed as creating space for niches, growing to

---

<sup>2</sup>Although later definitions for instance (van den Bosch & Rotmans, 2008) consider scaling-up a process concerning transition experiments, which are discussed shortly.

<sup>3</sup>For more on transition management see for example (Loorbach, 2007) or (Rotmans *et al.*, 2001).

<sup>4</sup>As distinguished from the peloton, the pack. Some transition frontrunners are sports aficionados.

## CHAPTER 2. LANGUAGE

become niches, and forming networks to connect niches and the key actors associated with them. The transition arena is conceptually akin to, but defined to contrast to, policy and political arenas. In these arenas a shared vision and language is to be developed with respect to the desired transition. Since bringing about, or propagating some societal innovation (Rotmans, 2005) could be regarded as the functioning of a transition arena, it appears not so far fetched to regard a transition arena as a constellation. In fact as Loorbach (2007) puts it: “the transition arena can be considered in its first phase as a niche; one could even define it as a policy niche.”

The transition management framework also envisions the transition arena initiating transition experiments, although these can very well take place outside the context of an arena. A *transition experiment*, according to van den Bosch & Rotmans (2008) is an “innovation project with a societal challenge as a starting point for learning aimed at contributing to a transition.” The important difference between a transition experiment and ‘normal’ innovation projects is that the latter usually focus on a certain technological innovation, whereas a transition experiment aims to contribute to a transition, with a societal innovation.

One could say that the experimentation conducted is about how a societal innovation, a novel form of functioning, holds in a protected but real-life situation. If successful, or should one say: satisfactory, the aim becomes to scale-up the transition experiment to become a proper niche. The learning referred to in the definition is actually cast in terms of learning about the structures, cultures, and practices emerging in the experiment, a triplet commonly used to describe functioning in transition studies<sup>5</sup>. This form of learning is referred to as *deepening*. Conducting a transition experiment in various contexts is referred to as *broadening*. In short, it seems not inappropriate to consider a transition experiment to be a constellation.

Both the previous constellations in one way or another revolved around the concept of societal innovation. It has been used here analogously to the use of a technological innovation in a socio-technical transition. But more than that, it was already referred to as being a form of functioning. Although the term was suggested in the context of societal transitions by Rotmans (2005) as a process, with innovation as a verb, it was and will be used here in its nounal form. Thus the last and most ephemeral ideal-type of constellation to be given here is the *societal innovation*. Let it be defined as a powerless constellation with a novel way to meet a societal need, societal functioning in its purest form.

### 2.3 Etymology

Obviously the conceptual language presented in the above has a history and many of the words used already have their specific meanings in certain scientific (sub)fields. This alone suggests that it is worthwhile to investigate the origins of the transition terminology somewhat, in other words some etymology seems appropriate.

It all begins with the concepts societal system, functioning and constellation. The roots of the first two concepts, however, have already been discussed so extensively in the introduction and the language section that further delving

---

<sup>5</sup>See section 3.1.2 on various approaches to describe functioning, among which this triplet.

### 2.3. ETYMOLOGY

would most probably lead to repetition and tedium of the reader. The constellation concept, though, is a different matter. Its application here, as a societal subsystem, is as far as the author is aware novel and original to the theory. Constellation in everyday speech mostly refers to a configuration of stars, which is faithful to its Latin roots. It is also used more metaphorically for groups of people or more general arrangements, which is also how constellation is to be understood in the context of this theory<sup>6</sup>, like its usage in (de Haan, 2007) where is spoken of a constellation of actors, structures and practices.

Much revolves around the concepts of ‘regime’ and ‘niche’, especially in combination with a ‘landscape’, with which they are at the basis of the so-called multi-level perspective (MLP). Although the use of these three concepts for the stratification of transitions goes back to the work of Rip & Kemp (1998), it is undoubtedly the prolific case-study work done by Geels (see (Geels, 2005*b*), as a good place to start) on various socio-technical transitions that has made the MLP into one of the preferred tools in the transition researcher’s conceptual toolkit. This work and various collaborations has led to several refinements and additions such as typologies of transition pathways (Geels & Schot, 2007) and multi-regime analyses (Geels, 2007; Raven & Verbong, 2007).

But obviously these terms have more history and are much more widely used than in the study of (socio-technical) transitions. In fact a review of the regime concept alone can easily fill a book in itself and will thus not be attempted here. Some things are relevant to point out however. In many applications, like the usage in studies of international relations (Krasner, 1983) and for example in urban regime theory (Mossberger & Stoker, 2001), the concept is used for its integrative qualities. The use of the term regime makes it possible to refer to a collection — constellation if you like — of actors, entities and processes as if it were a unity. This is the very quality of the term that made it so very useful in the context of transitions in societal systems.

If the regime concept is used in so many different contexts one might expect to encounter all sorts of definitions to suit these various areas of application. This appears, however, to not be the case. Take for instance the Rip & Kemp (1998) definition of technological regime:

A technological regime is the rule-set or grammar embedded in a complex of engineering practices, production process technologies, product characteristics, skills and procedures, ways of handling relevant product artifacts and persons, ways of defining problems — all of them embedded in institutions and infrastructures,

which is not at all that different from the Krasner (1982) definition of an international regime, that speaks of

sets of implicit or explicit principles, norms, rules, and decision-making procedures around which actors’ expectations converge in a given area of international relations.

That the regime concept in urban regime theory bears close resemblance to this is not surprising as Mossberger & Stoker (2001) themselves propose it as an elaboration of the international regime concept, though without *actually*

---

<sup>6</sup>Though the author will not deny being inspired by a certain Norwegian progressive Black Metal band (Arcturus, 1994).

## CHAPTER 2. LANGUAGE

redefining it for their application but accurately sketching the concept in several descriptions. The closest they come to a definition is when they speak of regimes as

the collaborative arrangements through which local governments and private actors assemble the capacity to govern.

Apparently there is a deeply felt need for a regime-like concept. This is illustrated by the fact that Dosi (1982) proposed the concept of a technological paradigm in analogy with Kuhn's (1962) definition which is very much comparable to that of the socio-technical regime as Schot & Geels (2007) also point out, especially considering the Nelson & Winter's (1977) definition by which the term entered innovation literature that also speaks of technological trajectories.

Roughly two interpretations of the word regime are used in everyday English<sup>7</sup>. One that corresponds with the usage in the definitions just mentioned for which also the word regimen is sometimes used, that of a fixed pattern or set of rules. And the other referring to a form of government, for example in the phrase 'a socialist regime'. This latter interpretation is perhaps even more widely used and the transition discourse consequently is prone to sound rather belligerent, with phrases like 'overthrowing the regime' not unheard of.

There is however another interpretation that seems far from ones used above. It is the usage of the term regime to refer to a dynamical system being in a certain qualitative state. This can refer to systems in general, though this usage is most common in the mathematical sciences, especially in the context of nonlinear systems<sup>8</sup>. One can for instance speak of systems being in their chaotic regime. Typically a system, though evolving over time, remains within a certain regime and it can be drawn into another by varying some parameter. In a certain sense this is of course the 'fixed pattern' interpretation again, though from a quite different angle. Nevertheless, as pillar theory takes on a systems view on transitions, this does appear relevant and since the regime concept now enters from several disciplinary directions it is comforting to know that no grave contradiction has arisen.

Similarly the term landscape has its interpretation deriving from socio-technical studies as the broader embedding of a regime in the technological aspects of society (see again Rip & Kemp, 1998). Which is close to the systems concept of environment (see von Bertalanffy, 1956, for instance) and the redefinition in pillar theory does not stray from that either. There is another use of the term landscape that is relevant here still, that of an adaptive landscape which is used in the study of evolution in its various manifestations. For instance in Kaufmann (1993) evolution is modelled by species wandering the adaptive landscape and climbing the hills therein, as how high uphill they are is a measure for their fitness. Although this might seem a remote interpretation from the socio-technical one, it does complement it in that it is an exogenous entity to which is to be adapted. Moreover, it is similar to the idea of societal systems adapting to outside influences to continue meeting societal needs.

The word niche also has a distinct evolutionary ring to it. Indeed it is from the study of ecosystems that this term originates. Of the early writings in which the niche concept is used the one by Elton (1927) appears most relevant. He defines niche simply as follows:

---

<sup>7</sup>See for example (Summers & Rundell, 1987).

<sup>8</sup>See for instance (Cross & Hohenberg, 1993) and (Strogatz, 1994).

### 2.3. ETYMOLOGY

the "niche" of an animal means its place in the biotic environment, its relations to food and enemies<sup>9</sup>.

This is an interpretation of a niche as a role, as a functional concept and as such very akin to pillar theory's interpretation of the niche as a constellation. At least one older ecological definition exists, for instance by Grinnell (1917) but there the niche is defined in terms of habitat and farther from the niche-as-constellation idea<sup>10</sup>.

Linguistically the conceptualisation of a niche in pillar theory is all wrong. See any English dictionary (e.g. Summers & Rundell, 1987) to learn that the word niche refers to a hollow place in a wall. From this the metaphorical use as a role obviously stems, and this must be the root for its use in ecology. The container has been replaced by what it is supposed to contain, so to say. This is no hanging matter, though it needed to be said.

The use of the niche concept in the study of technical and socio-technical change has been thoroughly investigated by Schot & Geels (2007) and it is certainly not the intention here to summarise, or attempt to improve upon, their work. From their excellent article emerges the common denominator that the niche concept is especially useful to address radical change. This in contrast with normal technological development along the lines set out by the regime or according to the technological paradigm. The emergence of a successful new technology is more or less equated to the process of speciation in biology. The niche then, is viewed as providing a space where the selection criteria of the regime do not hold, a place where an innovation can develop without having to conform to the trajectories of the technological paradigm.

This is applied by different scholars in different ways. Some view the niche as a place of isolation facilitating instances of punctuated equilibrium, thus accounting for the radicality of the innovation. Punctuated equilibrium was introduced by Eldredge and Gould in (Eldredge & Gould, 1972) as an alternative to the view that evolution always progresses in a slow, and even-tempered manner. Scholars like Levinthal (1998) take up this idea and combine it with the idea of technological innovation developing in niches. Others propagate ideas that resemble punctuated equilibria but do not explicitly use niches to make their point, like Tushman & Anderson (1986) with their technological discontinuities. This would all fall under the punctuated equilibrium paradigm as Gersick (1991) describes it.

Then there is the concept of transition itself. A detailed treatment about the nature of a societal transition has already been given in the previous chapter and what sets apart transitional change will be treated extensively in the next. Therefore the treatise here will restrict itself to how the *concept* came to be used as it is used in pillar theory.

One recalls that transition studies is more or less the result of the confluence of two rather different fields of research. The one being innovation studies and the study of socio-technical systems, and the other being that of integrated

---

<sup>9</sup>Which is followed by the badger-vicar analogy: "Thy ecologist should cultivate the habit of looking at animals from this point of view as well as from the ordinary standpoints of appearance, names, affinities, and past history . When an ecologist says 'there goes a badger' he should include in his thoughts some definite idea of the animal's place in the community to which it belongs, just as if he had said 'there goes the vicar.' "

<sup>10</sup>A good and systematic systems overview of the niche concept in ecology can be found in (Patten & Auble, 1980).



## CHAPTER 2. LANGUAGE

assessment. To start with the latter, in (Rotmans, 1998) transitions are referred to as a guiding notion within integrated assessment and a compact definition of a transition is given as well<sup>11</sup>, in the form of

a gradual, continuous shift in society from one “mode of operation” to another.

with the original quotation marks. Apparently inspired by the demographic transition (Davis, 1945), transitions are conceptualised as going through several phases. Three phases in (Rotmans, 1998): a take-off, an acceleration and a stabilisation phase, to which in (Rotmans *et al.*, 2001) a pre-development phase is added. That this is a broader conception of transition than just technological is clear from the description of the acceleration phase in which socio-cultural, economic, ecological and institutional changes are said to accumulate, and in the earlier description where the system is instable due to rapid technological, social and environmental change. In (Rotmans *et al.*, 2001) also the notorious s-shaped curve appears to illustrate the phases of a transition, linking it to population dynamics. It also established a link, though implicit, with innovation studies where such curves visualise the successful diffusion of an innovation (see e.g. Rogers, 2003). As a research field integrated assessment is partial to modelling approaches which accounts for its systems view on society and societal processes, which clearly found its way to the theory presented in this book.

The first occurrence of the transition concept in the technology studies literature appears to be in (Kemp, 1994). In this article the transition concept is employed in a broad sense to describe the kind of shifts necessary to come to an environmentally more sustainable energy system. Here one observes the gradual identification of the idea of a shift in regime with a transition, as a regime shift is sometimes seen as a means to bring about a transition and other times used interchangeably with it. It seems that the transition concept was actually narrowed down to suit socio-technical transitions rather than that it was broadened to encompass the idea of a societal transition.

In the following years the Dutch government became interested in the transition concept and the possibility of transition management as a way to address problems in societal systems, like energy and agriculture. This culminated in the ministry of environment (VROM<sup>12</sup>) commissioning a preparatory study for their 4th National Environmental Plan (NMP4<sup>13</sup>) from the ICIS institute, then led by Rotmans. In one of the three reports for that study Rotmans *et al.* (2000) use the multi-phase concept, as it came to be known, to investigate the possibilities for transition management for a transition to an emission-free energy supply system. In one of the background studies Geels & Kemp (2000) give a socio-technical perspective on transitions, applying the multi-level perspective to two historical cases. Interestingly, they define transitions as great changes in the way societal functions are fulfilled<sup>14</sup>. These societal functions are in their

---

<sup>11</sup>This definition in fact goes further back and can be traced back to Rotmans' earlier work described in the UN (1997) report.

<sup>12</sup>Actually, the *Ministerie van Volkshuisvesting Ruimtelijke Ordening en Milieubeheer*, the ministry of Housing, Spatial Planning and the Environment.

<sup>13</sup>The *Vierde Nationaal MilieubeleidsPlan*, the fourth national plan for environmental policy.

<sup>14</sup>The original definition is in Dutch and goes: “*Transities worden dus opgevat als grote veranderingen in de manier waarop maatschappelijke functies vervuld worden.*”

### 2.3. ETYMOLOGY

view fulfilled by socio-technical regimes<sup>15</sup>.

The commissioning of this preparatory study marked the confluence of the research field and the birth of the study of societal transitions for which the next chapters hope to provide a theoretical basis.

---

<sup>15</sup>First they use the term socio-technical system, but they switch to socio-technical regime to avoid confusion with the research field of large technical systems.

*CHAPTER 2. LANGUAGE*

## Chapter 3

# Pillars

### 3.1 Preliminaries

In the preceding chapter the reader was acquainted with societal transitions and with a language. In the following chapter the reader will, in essence, learn how to tell transition stories using that language; a theory is exposed that provides the elements to construct and reconstruct transition paths in various forms. A transition unfolding in time can be considered a story line, or several intertwined story lines. Set against some — not necessarily static — stage, the transition presents itself to an audience as a sequence of dramatic situations. The protagonists in a transition story are the constellations, their functioning their role. A description of the constellations in a societal system and thus the societal system itself, is called the societal systems composition. This theory describes the stage in terms of the *conditions* of the societal system and the dramatic situations are presented as a set of *patterns*. By concatenating patterns against a backdrop of the conditions, transition *paths* are created. These conditions, patterns and paths are therefore called the *pillars* of transitional change.

More concretely, the patterns are the different forms of transitional change that can occur in a societal system. The conditions describe the state of the system, specifically they are to be interpreted as abstract indicators for how prone the system is to transitional change. Since this theory is constructed rather than being the result of numerous iterations of the cycle of the scientific method, the hypotheses on how the conditions drive the patterns are somewhat speculative. Somewhat, for these hypotheses are obviously based on former theoretical work, empirical observations and common sense. However since these hypotheses only couple the conditions and the patterns, new insights can be incorporated easily. Rather than destroying it, falsification of an hypothesis would lead to an improvement of the theory.

If the theory provides the elements to construct and reconstruct any transition path and the coupling hypotheses show what leads to what, then one has a basis for dynamical models of transitions. Indeed, as indicated earlier, being a basis for computational and mathematical modelling approaches is an explicit goal of the theory. To elaborate this basis somewhat chapter 4 will elaborate a formalisation of the theory, casting the essence of the theory in a pseudo-mathematical symbolic language.

#### 3.1.1 Societal systems, functioning and constellations

In the chapter on language one of the central concepts was the societal system and its functioning — conveniently circularly defined as a system meeting some societal need and the way it does that. Even when narrowed down to one subject, one societal system, the societal need still has many facets and therefore so has the functioning. For example, the need for mobility has facets concerning long and short distance, comfort, appeal to a sense of freedom, efficiency, availability, etc. Not all facets are considered equally important, nor are they appreciated and used the same way over time. In the evolution of a societal system this leads to functional differentiation (Luhmann, 1982) into several societal *subsystems*: constellations.

Each constellation then can be considered to meet several facets of the societal need and consequently contributes a part to the total functioning of the societal system. In a mobility system, for instance, the societal need for gen-

### 3.1. PRELIMINARIES

erally accessible transport at low cost would be met by a constellation around public transportation, whereas short distance mobility could be met by constellations around slow modes of transport. The composition of a societal system — which constellations are present and how they meet which facets of the societal need — provides a detailed picture of its functioning. Some facets of the societal need can be met by more than one constellation simultaneously; think for example of car-based mobility and forms of public transport when it comes to commuter traffic. Depending on how the constellations evolved such a situation can give rise to a constellation taking over facets of functioning from others, competition or coexistence of different solutions.

The evolutionary dynamics of constellations, and consequently of the entire societal system, are driven by the interplay between the functioning and the societal need it meets — and the changes therein, since societal needs are obviously not constant. Evolution is generally thought of as giving rise to slow, incremental change and the evolution of a societal system is not considered to be different in that respect. Apparently transitional change is of a different nature, perhaps similar to punctuated equilibria in some theories on evolutionary processes. This seemingly fundamental difference between transitional change and ‘normal’ change in a societal system is worthy of some elaboration, which will be done in section 3.1.3, since it will have consequences for identifying the conditions for transitional change and the patterns in which it affects constellations and the societal system.

The composition of the societal system is the stage on which the transition story is performed, so describing and understanding a transition will inevitably involve painting a clear picture of this stage. This in turn requires the ability to describe constellations and all the facets of their functioning well. This, however, is no mean feat, as functioning, the way societal needs are met, is a many-headed dragon. The various facets of functioning will in general require different concepts and measures to characterise them. Moreover, there is a slight, but relevant difference between concepts to describe with and the description of these concepts. In other words, describing something using a conceptual language like ‘societal system’, ‘constellations’, ‘functioning’, etc. is something very different from the operational description of the concepts like ‘central to the functioning of the cure-constellation in a health care system is the approach of complaint-diagnose-treatment’. Theory works with concepts, the link with empirical reality is made through their operational descriptions.

#### 3.1.2 Describing constellations and functioning

Describing the composition of a societal system would involve identifying the constellations it consists of and giving operational descriptions of their respective functioning, which together gives the functioning of the societal system as a whole. The various facets of functioning will in general require different concepts and measures to characterise them. In a mobility system for instance, legislation is as important as the physical infrastructure as is the way people actually use it. These facets are conceptually not readily comparable, nor described in the same terms. Making such properties commensurable, or at least comparable is the realm of Integrated Assessment (Rotmans, 1998). Several transition-related integrated assessment models exist, like the TARGETS-model (Rotmans & de Vries, 1997), the SCENE-model (Grosskurth & Rotmans, 2005), the transi-

## CHAPTER 3. PILLARS

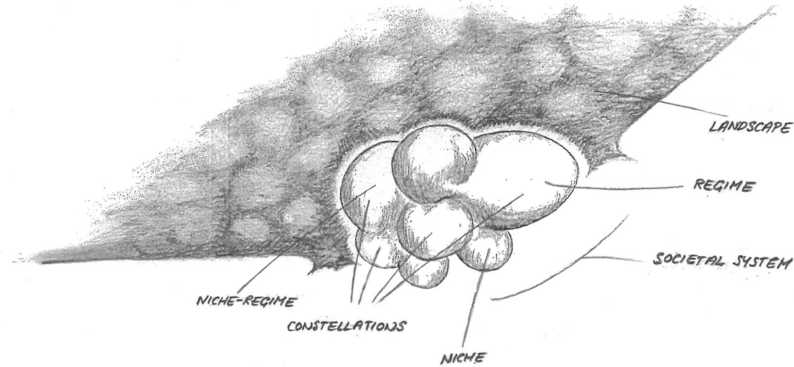


Figure 3.1: Composition of a societal system

tion monitoring framework in (Taanman, forthcoming in 2010) or Franzeskaki's (2009) Clover model. Casting any of this in a quantified form requires defining indicators, composite indicators, indexes and the like, on all of which more has been written than can be overviewed here in any sensible way.

### Structures, cultures and practices

One way of describing functioning that will be treated here is by means of the triplet *structures, cultures, practices*. Although still remaining much on the conceptual level and therefore arguably not an operational form of description, this triplet is very instrumental for narrative approaches. The triplet was proposed by Rotmans (Rotmans & Loorbach, 2009) and further elaborated by van Raak (2010). Definitions of these terms (taken from Grin *et al.*, 2010) are:

#### *Structures*

... physical infrastructure (physical stocks and flows), economic infrastructure (market, consumption, production) and institutions (rules, regulations, collective actors such as organizations and individual actors). Structure is recursive: it is both the result and means of acting.

#### *Cultures*

... the collective set of values, norms, perspective (in terms of coherent, shared orientation) and paradigm (in terms of way of defining problems and solutions). In our transition context, then, culture has a quite specific meaning, one that differs from the traditional sociological conceptualization of culture.

#### *Practices*

... the ensemble of production routines, behavior, ways of handling and implementation at the individual level, including self-reflection and reflexive dialogue.

### 3.1. PRELIMINARIES

In pillar theory however, practices are not explicitly used and functioning is considered a dual concept, comprised of *structures* and *cultures*. This goes back to an earlier duality introduced in de Haan (2007), where the same concepts are referred to as the interactive and reflective part of functioning. Thus, the working definition of structures and cultures that will be used in this book is:

#### *Structures*

The infrastructures and institutions related to the functioning of the societal system; *how it works*.

#### *Cultures*

The appraisal and legitimacy of the functioning of the societal system; *how it is perceived*.

Structures here refer to the physical infrastructures as well as the economic or legal aspects and cultures refer to cognitive, discursive and ideological aspects. That practices are not explicitly used is of course because of their link to the level of the individual which was also not explicitly used.

There is a straightforward argument why practices and the individual level need only implicitly be considered. Societal systems are complex systems with functioning as an emergent phenomenon. This implies that functioning, described by the structures and cultures, is robust with the processes and interactions that produce it on an underlying level. This underlying level is obviously the level of the individual and their (inter)actions. Robust means that the emergent phenomenon is decoupled from the underlying level in that the individual (inter)actions there do not affect the overall functioning<sup>1</sup>. Paradoxically the role of actors and practices is central in how functioning emerges in societal systems, but in the description of functioning one can abstain from it.

Now in societal systems the actors are aware of the functioning of the system and that they have a role in that. This role is an emergent property on the underlying level. The emergent property on the underlying level together with the emergent phenomenon on the higher level form the *conjugate* and in societal systems<sup>2</sup> such a conjugate can always be identified. A conjugate couples the emergent property on the underlying level with the higher level at which the emergent phenomenon resides. In this case the conjugate is practices, for example in the form of habits, routines and daily actions.

Obviously a societal system and its composing constellations are social structures in the sense meant by Merton (1968) and it might seem awkward to describe its functioning in terms of structures and cultures, all the more since culture is a concept with too many definitions already (see e.g. Kroeber & Kluckhohn, 1952), of which some are interpretations of culture as a social structure. The terms structures and cultures, however, do have strong intuitive appealing that make them suitable as concepts in describing functioning. Thus the term structures stems from the idea that there are physical and formal social structures that constrain and enable the practices via which the functioning is

---

<sup>1</sup>A framework for analysing emergence in this fashion can be found in (de Haan, 2006) as well as the terminology used here, e.g. robustness, conjugate, levels, etc.

<sup>2</sup>In fact all systems exhibiting type-III emergence. Again see (de Haan, 2006) or see the chapter "Consequences of Complexity" in the Foundations book for a more in depth treatment of emergence and its implications.



## CHAPTER 3. PILLARS

produced, almost in the sense of Giddens (1984). This all is not to deny that there are non-physical and informal social structures with similar properties, but exactly to be able to make this distinction the term cultures is introduced. Thus with cultures are meant the cognitive, discursive and normative aspects through which the societal system gives meaning to its functioning. This sense-making attribute to the concept of cultures brings it close to Geertz's (1973) and Bourdieu's (1977) interpretations of the term.

The added value of describing functioning in a dualism of structures and cultures is twofold. Firstly it does right to the fact that both are perceived as important aspects of societal reality. Moreover, the sense-making quality of cultures also hold for the researcher using them to describe societal systems in that they provide a perspective on the structural part of the description. The converse also holds since the presence of certain structures helps to understand the presence of certain cultural aspects. Secondly, in the cases that structures and cultures are not two sides of same coin exactly these discrepancies can provide insight into the condition of the societal system in the sense of whether the system is prone to transitional change.

### 3.1.3 Change and transitional change

The matter of what sets apart transitional change from other forms of change still needs to be resolved. This in fact goes back to the question of what a transition exactly is.

The *word* transition seems to imply qualitative change, that is, change not only in the sense of more or less of this or that but different *ways* to this or that. The dictionary definitions seem to second that. Here are the relevant lemmas for the word transition from the Oxford English Dictionary (Soanes & Stevenson, 2004) and the 'Dikke van Dale' (Geerts & Heestermans, 1992), the authoritative dictionary of the Dutch language<sup>3</sup>.

*Transition* (Concise Oxford English Dictionary)

**transition** ■ **n.** 1 the process or a period of changing from one state to another.

*Transitie* (Van Dale, Groot Woordenboek der Nederlandse Taal)

**transitie** (v.; -s), overgang;

Indeed a synonym is not always that helpful in providing a meaning. *Overgang* in turn (according to the same dictionary) translates to "the going from one state to another".

Both lemmas mention a change of state, and although it is true that the state of a system can be expressed quantitatively it is nevertheless a quality. The definition given in the chapter on transition language speaks of "... radical changing of the functioning of a societal system from one state of apparent societal equilibrium to another" or a radical regime change, which is synonymous<sup>4</sup>. The picture emerges of transitions as upheavals, tossing a societal system from one period of relative calm to another, changing it profoundly.

<sup>3</sup>Since much theorising on transitions is done by Dutch researchers in the English language, giving the Dutch definition of the word *transitie* does not seem superfluous

<sup>4</sup>For the regime was to represent a certain societal equilibrium to begin with.

### 3.1. PRELIMINARIES

So, what exactly does ‘radical’ mean here and how radical does it need to be? And moreover how far apart do the equilibria need to lie or how sharp needs to be their separation? The term radical itself has in contemporary English roughly two distinct uses, though sharing the same root: root<sup>5</sup>. The first use, which usually is also the first to spring to mind, is to denote extremity, rapidity or deviance from normality. This brings one back to where one started, questioning then the amount of radicality needed. It does, however, provide a clue that transitional change does not consist merely of the evolutionary processes by which constellations are able to continue their meeting of certain societal needs. Apparently then, one can also conclude that the conditions for transitional change are not to be found in the interplay between the functioning of constellations and the (changing) societal need. Where the conditions for transitional change *do* reside and what it *does* consist of will be treated in the respective sections of conditions (3.2) and patterns (3.2). First some more on what sets transitional change apart from normal societal change.

The normal development of a societal system was said to be evolutionary. In this sense the societal need could be considered the selection environment to which the system adapts. The processes through which constellations then emerge and develop can then be regarded as a form of speciation. Stretching this metaphor somewhat, one realises that a transition would be not so much the formation of a new species but more a jump in evolution like Eldredge & Gould (1972) suggested: punctuated equilibrium. The normal evolutionary development would then correspond to phyletic gradualism. Comparing transitions to punctuated equilibria in this manner places this theoretical treatment in the category of revolutionary change theories as described by Gersick (1991).

Indeed, another angle towards thinking on transitions in societal systems also leads one to arrive there: complex systems. Societal systems can be considered complex, adaptive systems. Conversely authors like Prigogine (e.g. in Nicolis & Prigogine, 1989) point out that a striking property of complex, adaptive systems is that they are systems with the intrinsic ability to go through transitions. They also point out that change in these systems is irreversible, which is also recognised as a key property of transitional change (Grin *et al.*, 2010) in societal systems. Complex, adaptive systems are also filed under revolutionary change theories by Gersick that according to her describe a revolutionary period — a period of transition — as a period where the ‘deep structure’ of the system is changed, as opposed to leisurely evolving according to it.

Thus one arrives at the other use of the term radical, more faithful to its etymology perhaps, which is to imply a relation with the origin or fundament of something. This is also close to the definitions of transition given by other authors who speak of fundamental shifts, structural changes, like in (Kemp, 1994; Rotmans *et al.*, 2001, etc.). So it appears that transitional change is intimately related to the composition of the societal system and consequently to the constellations of which it is composed. If this so be, then this would hold for transitional change as well as for what *drives* it. So what is to be found in the composition of a societal system?

Recapitulating: A societal system is a system functioning to meet some societal need. It is an open system, embedded in an environment called the landscape. It is a composite system, composed of societal subsystems called

---

<sup>5</sup>Deriving from the Latin *radix*: root

## CHAPTER 3. PILLARS

constellations. The way a societal system meets the societal need is through the joint functioning of the constellations that together bring forth the functioning of the whole.

Thus transitional change is either: non-evolutionary change in the functioning of constellations, especially the regime since it dominates the functioning of the entire societal system. Or: change in the way the constellations together bring forth the functioning of the societal system as a whole. For instance, by new constellations emerging and becoming more dominant or constellations influencing each other or merging. The former option for transitional change, the change in functioning of constellations, was to be of a non-evolutionary kind<sup>6</sup>. Again the same mechanisms come to mind: a constellation could change its functioning radically by being influenced by another constellation or merging with one. Although it is possible to promote the above processes to the status of ‘elementary mechanisms’, this might very well be constraining rather than enabling in treating transitional change. The reason for this will be given shortly. First a list of possible elementary mechanisms is given:

### *Creation*

It is not difficult to imagine that when there is a societal need unfulfilled, the societal system will respond by letting new functioning emerge. This could be compared to what is sometimes called filling a power vacuum (Avelino & Rotmans, 2009; Avelino, 2009), especially with the interpretation of power in terms of functioning discussed earlier in section 2.2. This filling of a power vacuum could result in the emergence of a new constellation or the scaling up of an existing one.

### *Co-evolution*

As in the ecological analogue it is possible that the development of one constellation is related to the development of another<sup>7</sup>. Especially when the constellations serve similar societal needs or depend on similar resources, it is not unlikely that the development of one shapes that of another. Whether this is rightfully to be called co-evolution is another matter<sup>8</sup>. In the context of transitions co-evolution is understood as “interaction between societal subsystems influences the dynamics of the individual societal subsystems, leading to irreversible patterns of change” (Grin *et al.*, 2010). Although the processes underlying transitional change ought necessarily to be of a non-evolutionary nature, co-evolution is another matter<sup>9</sup>. The functioning of a constellation can straightforwardly be interpreted as being a selection environment for the evolutionary processes in another, which then is affected more radically than it would be if it were only evolving according to its own selection environment, provided their functioning is ‘different’ enough.

---

<sup>6</sup>Since the normal, evolutionary development of constellations would change the system on too long a timescale. It is in other words too slow to be radical.

<sup>7</sup>Or others, co-evolution is not restricted to pairs.

<sup>8</sup>See van den Bergh & Stagl (2003) on sensible usage of the concept of co-evolution outside the ecological context and Kemp *et al.* (2007) for a discussion of co-evolution in the context of transitions.

<sup>9</sup>See also the chapter on “Consequences of complexity” in the first part of this book.

*Clustering*

Sometimes, for reasons of efficiency, expansion, or others, cooperation is a good means. One form of this cooperation is clustering, which would be the equivalent for constellations of a merger of companies. Think for example of the clustering of a distributed car rental system with a constellation around electric automobiles. The cluster is essentially a new constellation with the combined functioning of the constellations that formed it. This also means that the cluster can in principle count on the combined societal support of its composing parts.

Although it is worthwhile to pursue these elementary mechanisms in some more detail there is also the realisation that each of these processes in themselves would not amount to any appreciable transitional change. Therefore apparently there need to be several of these processes at work simultaneously or subsequently and it is one of the central assumptions of this theory that in transitions this happens in certain specific and identifiable *patterns*. This also provides a way in which the theory can be compatible with and accommodate other theories on transitional change: Once the patterns are identified<sup>10</sup>, processes or mechanisms found empirically or from other theoretical or modelling approaches simply become part of the pattern. Patterns can thus in a certain sense be seen as bundles of more elementary mechanisms.

But what triggers these patterns? When does a societal system ‘feel the need’ to go through a transition? More than machinery that is being switched on by a certain event, patterns of transitional change must be seen as just that: patterns in the whole of processes that make up transitional change. Following this line of thought one tends to look at broader circumstances than events or incidents, in other words one looks for conditions that are drivers more than triggers. That such conditions were to be found in the composition of the societal system, where after all the transitional change takes effect, was already acknowledged earlier here; what these conditions could be is explored and elaborated in the next section. That incidents, events and such are more to be regarded as the consequences of a societal system prone to transition than the other way around is discussed in the section on persistent problems and symptoms of unsustainability.

---

<sup>10</sup>They will be deduced from a more general scheme, in fact.

## 3.2 Pillar I: Conditions

...are the misalignments, dysfunctions and impediments in a societal system that make it prone to transitional change.

Three conditions are recognised: Tension with the landscape, stress within constellations and pressure from constellations.

Tension, stress and pressure are necessary but not necessarily sufficient conditions for transitional change.

### 3.2. PILLAR I: CONDITIONS

In the previous section transitional change was alluded to as *not* being driven by the interplay between societal needs and the functioning of constellations. Instead, the suggestion was made that instead the *conditions* for transitional change could be found in the very composition of the societal system: the functioning of constellations in relation to the landscape, to each other and of the different facets of functioning within a constellation. The idea behind this is that in the absence of conditions for transitional change the societal system evolves, gradually, incrementally to meet the changing societal needs. This in contrast to what is considered to be transitional change where systems change fundamentally and radically. If transitions change the very fundamentals of the functioning of societal systems, then it appears that the conditions for transitional change too might have their origins in the composition of the societal system.

This in mind and turning to the protagonists of transition stories, constellations, one wonders what forces these endure that could be interpreted as the conditions for transitional change. The functioning of constellations can be compromised in many ways. If this has its origin in the relation between the functioning of the constellation and its environment, the landscape, this will be denoted as *tension*. If the functioning of a constellation is compromised because of different facets of functioning within a constellation are at cross with each other, one speaks of *stress*. Then there still is the possibility of the functioning of a constellation being compromised by the functioning of one or several other constellations. For instance if another constellation becomes a serious alternative. If this is the case one says a constellation experiences *pressure*. These three, tension, stress and pressure are considered necessary but not necessarily sufficient conditions for transitional change to occur. The three conditions for transitional change are summarised and illustrated with some examples in table 3.1 and treated in some more depth in sections 3.2.1, 3.2.2 and 3.2.3.

---

#### Conditions for transitional change

---

##### **Tension** — *between a constellation and the landscape*

Adverse functioning of a constellation in relation to its environment, the landscape.

*Symptoms:* Problematic resource or waste flows across the boundaries of the societal system, depletion and overproduction. Issues concerning perception, acceptance and legitimisation of how a constellation meets a societal need.

*Examples:* Difficult political climate, negative public opinion and environmental awareness, climate change, resource depletion, pollution and waste disposal problems, structurally rising costs of publicly financed systems.

Continued on next page...

---

Conditions for transitional change — *continued*

---

**Stress** — *within a constellation*

Internally adverse functioning of a constellation.

*Symptoms:* Conflicts of interest, perverse mechanisms, means becoming goals, contradictory self-perception. Negative consequences of optimisations. Functioning impeding itself.

*Examples:* Free market philosophy in publicly financed systems (health care, for instance), policy stimulated overproduction, regional systems perceiving themselves to serve an agricultural function whilst effectively being a residential area subsidising on a forensic community. Subsidies or quota harming rather than protecting fish populations. Livestock epidemics as a consequence of intensification of agricultural production. Traffic congestion.

**Pressure** — *from other constellations*

Adverse functioning of a constellation with respect to another constellation.

*Symptoms:* Considerable alternatives emerging for the constellations functioning. Negative effects of competition, functioning becoming obsolete or made so by innovative solutions of other constellations. Functioning impeded by constellations laying claim on similar resources or infrastructures, having incompatible institutional settings, or otherwise.

*Examples:* Constellations emerging around a new technological solution (electric mobility pressuring the internal combustion based regime) or with entirely different approaches to the societal need (decentralised sustainably generated electricity putting pressure on the fossil fuelled central regime).

---

Table 3.1: The conditions for transitional change

Obviously, if an inconsiderable, powerless constellation is under severe conditions for transitional change this is still not likely to have transitional consequences for the societal system at large. For this reason it is convenient to assume that, if no constellation is specified, the conditions refer to the *regime*, which, being the most powerful constellation, is representative for the functioning of the entire societal system. This is a consequence of constellations being similar entities as societal systems, both simply defined as having a functioning for fulfilling some societal need. Thus in general, what holds for a societal system *mutatis mutandis* holds for constellations and *vice versa* as well.

Finding operational descriptions for the conditions, in other words indicators, indices, etc. for obvious reasons comes with the same difficulties as with the operational description of constellations and their functioning. Thus for the same reasons this will not be treated here. However, if the triplet structures, cultures and practices is used to describe functioning one can add some refine-

## 3.2. PILLAR I: CONDITIONS

ment to the conditions for transitional change. Different facets of functioning now fall either under the ‘structural’ or ‘cultural’ side of things which helps to identify some typical manifestations of tension, stress and pressure.

### 3.2.1 Tension

In many ways constellations, and consequently societal systems, depend on their relation with the environment: the landscape. Constellations are open systems and sustain themselves, their functioning and their boundaries, by maintaining an appropriate influx of resources, be they energy, goods, financial capital, labour or information. The converse holds of course for what leaves in the form of waste or resources for other systems. These resource and waste flows are, however, only part of how a constellation needs to fit its environment and to avoid running the risk of focussing too much on these underlying processes it is appropriate to describe it on the higher level of its functioning. Via practices a constellation sustains the structures and cultures of its functioning and how these structures and cultures are related to the landscape can indicate if the constellation suffers from tensions.

Since functioning itself could be described in terms of structures and cultures, one is straightforwardly led to the idea of structural and cultural tensions. Structural tension would amount to a misalignment of the functioning of a constellation and the landscape, with regards to its physical, infrastructural, economic, formal and legal aspects. This is also referred to as interactive tension for its tangible characteristics, this in contrast to cultural tension, which therefore is also referred to as reflective tension.

Cultural tension amounts to the misalignment of cognitive, discursive, normative, ideological aspects of functioning with respect to the landscape. Structural tension is a mismatch of the workings of a constellation with respect to its environment, whereas cultural tension is a mismatch in the making sense of the workings of a constellation with respect to the environment.

So, typical signs of structural tension are: resource depletion, pollution and waste

disposal problems, as well as the functioning of one constellation compromising the functioning of another in another societal system<sup>11</sup>. Specifically, one can think of an energy system, or a mobility system exhibiting structural tension when depleting the fossil fuels they depend on. However a bio fuels based regime can suffer structural tension if it compromises the functioning of another societal system like agriculture or food-logistics systems. Apparent examples are water management systems suffering structural tensions in the most literal sense, facing the consequences of climate change, and collectively financed health care systems that become too costly to sustain.

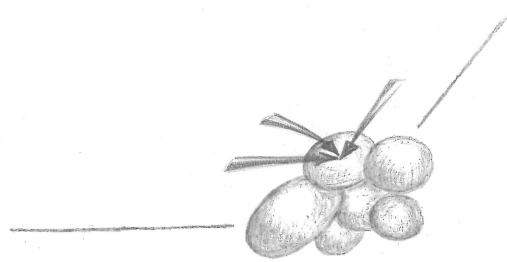


Figure 3.2: Tension

<sup>11</sup>Note that this is not *pressure*, tension concerns influences stretching *across* the boundaries of the societal system.



## CHAPTER 3. PILLARS

Typical signs of cultural tension are a changed political climate, or awareness of problems formerly not noticed, different patterns of consumer behaviour. As is to be expected cultural tension is in many senses the pendant of structural tension. For instance, the awareness that the climate is changing is a cultural tension felt by many industries that might be unaffected by the structural tension induced by climate change, think again of an energy system. Another example is a political climate that favours the free market ideal, which produces a cultural tension for societal systems providing public services, like again health care, or social security.

### 3.2.2 Stress

The functioning of a constellation is the result of the interplay of different facets of that functioning. How a mobility system functions depends not only on different mobility services offered by the different modes and their infrastructures but also on how these different facets are interlinked. The condition of stress refers to some facets of functioning not working well together or being a hindrance to each other. An example for the car-based individual mobility constellation would be that the facet of providing a sense of personal freedom is at odds with being collectively stuck in the consequential gridlock. This holds for societal systems as much as for individual constellations. Again in the following the focus will be on the regime-constellation as it is representative for the societal system at large.

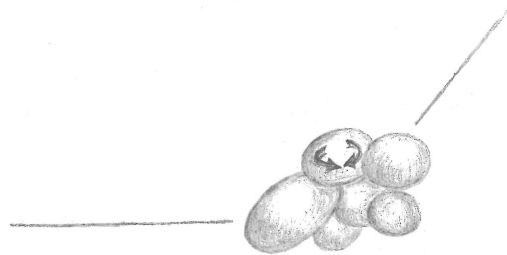


Figure 3.3: Stress

If the regime-constellation is described in terms of structures and cultures, one expects the two to be in a certain balance since they are two aspects emanating from the same processes of emergence. However, functioning is the emergent result of several emergent phenomena, resulting in several facets. As a result discrepancies can creep in. The two extremes of this are that either the structures and cultures within the regime become misaligned, resulting in a certain sense in the system not practising what it preaches, or that keeping structures and cultures aligned and consistent has become a goal in itself, in which case the practising is self-interpreted by the system as fulfilling the societal need. Let the former case be referred to as systemic hypocrisy and the latter as perversity.

Many examples of systemic hypocrisy, the misalignment of structures and culture in the regime, can be found. For instance, the reorganisation of the Dutch health care system in recent years has resulted in a culture of free market thinking where health care is firstly thought of as a product. The structures, however, are still based on a system of organised solidarity, providing health care as a right<sup>12</sup>. Another example would be that of a region — like the Dutch province Zeeland, see (Henneman *et al.*, 2006) — that on the structural side serves the societal function of a residential area with a services economy,

<sup>12</sup>Compare this example of stress in a health care system with the tension example.

## 3.2. PILLAR I: CONDITIONS

whereas, perhaps for historical reasons, the cultural side still regards itself as an agricultural region.

As regards societal systems exhibiting perversity many examples are available as well. Any system where means become goals are prone to this symptom of unsustainability. An example of perversity in societal systems is the Dutch agricultural system of a couple of decades ago. In the post-war Netherlands the agricultural system was government regulated with the aim of dampening the oscillation between overproduction and shortages<sup>13</sup>. This was done via minimum prices and guaranteed buying of the produced agricultural products. This led to the perversion that farmers began to produce more and more against prices that had no connection with the demand, which resulted in butter-mountains and milk-seas and Dutch consumers paying double the price of foreign consumers.

### 3.2.3 Pressure

The last type of condition for transitional change has to do with the presence of an alternative to the functioning of a constellation in a societal system. Consider the possibility that a new constellation emerges, for instance with new technology, another approach in general or providing a way to meet a societal need that is essentially new and as such leaves the old functioning without reason for existence. Although it is in principle possible that two constellations providing functioning for similar aspects of a societal need co-exist in a stable manner, like for instance car-based mobility and public transport, it is likely that a constellation experiences pressure from such an upcoming alternative. If this concerns the regime, which will be assumed tacitly in most of the following, such an upcoming alternative constellation is referred to as a niche-regime.

If the regime is considered a complex adaptive system, one realises that it needs to maintain a constant flux of material, energy and information to sustain its functioning and its boundary. If an alternative emerges for the regime, this can basically affect it in two ways. The first is that the alternative constellation, the niche-regime, makes use of the same flows of resources, in other words: competition. The second is that the functioning of the niche-regime makes the functioning of the regime in a way obsolete, which could be named obsolescence. In the following paragraphs these symptoms of pressure are typified

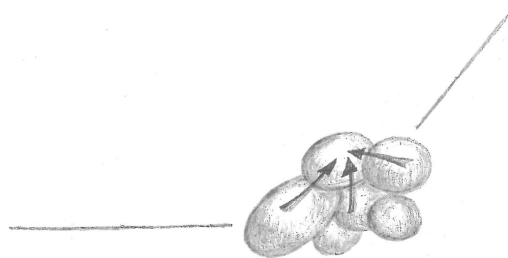


Figure 3.4: Pressure

in terms of structures and cultures again to keep the systemic perspective and refraining from the actual describing of the possible resource flows and such.

Although competition has clear intuitive connotations it still is worthwhile to view it from an emergence perspective in terms of structures and cultures. Obviously suppliers of sustainable energy need to compete with the regime of fossil fuel burning energy suppliers in terms of market share, which is a straightforward, albeit narrow measure of meeting the societal need, but there is more.

<sup>13</sup>See (Vijver, 2005), for an overview.

## CHAPTER 3. PILLARS

On the structural side there is existing legislation and infrastructure for which an upcoming alternative competes as well. Something similar holds for the cultural part, which becomes prone to changes in existing perspectives on how needs are met, as well as in the awareness of the shortcomings of the current functioning offered by the regime.

In the case of obsolescence the pressure is felt differently. This is where the regime finds that other structures and cultures have found ways to meet needs the regime was used to meeting, although there perhaps was no tension or stress. An example can be that of the traditional record industry that experiences considerable pressure from the constellation around digital music carriers, downloading and the likes. Another example, although on a smaller technological scale, is found in the work of Geels (2005*b*) where is described how the constellation that provided the supply of ice to houses became obsolete after the swift introduction of the electric refrigerator. The difference, although the two types of pressure always overlap of course, resides in that the new structures and cultures do not simply replace the old ones but have a new place of their own rendering the old ones obsolete.

### 3.2.4 Persistent problems and symptoms of unsustainability

Sustainability and *persistent problems* are concepts that appear frequently in literature on societal transitions, especially when it is concerned with the influencing or management thereof (see e.g. Rotmans, 2005). It is therefore appropriate to discuss how these concepts are related to conditions of transitional change.

The latter, persistent problems, are complex problems deeply rooted in the underlying structure of the systems that suffer from them. Their varying definitions are somewhat elusive (Rotmans, 2005; Dirven *et al.*, 2002) — which appears to be very suitable to the concept. Persistent problems are referred to by Rotmans (2005) as a superlative to Rittel & Webber's (1973) wicked problems. Perhaps superlative is too strongly put but persistent problems can be regarded to societal systems what wicked problems are to policy and planning. Because of their being rooted in the underlying structure of the systems that suffer from them, transitions are suggested to be needed to overcome them. This appears to be very much parallel to how Kuhn (1962) uses the term anomalies to denote the unresolvable problems that come with each scientific paradigm. Coincidentally, Dosi (1982) introduces his concept of technological paradigm in explicit analogy to Kuhn's paradigms and, as discussed in the etymology of the regime concept, technological paradigms have been an influential concept in transition thinking.

The former, sustainability or rather sustainable development is notoriously ill-defined. Weaver & Rotmans (2006) argue that this is a reflection of the concept being intrinsically normative, subjective and ambiguous and a generic definition therefore is a useless attempt. The Brundtland definition (1987) of it being development that “meets the needs of the present without compromising the ability of future generations to meet their own needs”, nevertheless is commonly used<sup>14</sup>. Sustainability is often used in the context of transitions, either

---

<sup>14</sup>And easily recast in terms of societal systems since it also refers to meeting needs. A

### 3.2. PILLAR I: CONDITIONS

as a ‘goal’ or as a framework in envisioning processes (Loorbach, 2007). Moreover, in the transition context the problems involved in achieving sustainable development are referred to as wicked (Kemp & Loorbach, 2006).

Apparently dealing with persistent problems and sustainability are in a sense similar issues and both have been associated with transitions as a way to address them. In the light of this it appears appropriate to investigate whether persistent problems can be described using the conditions for transitional change or *vice versa*. Also, it might be interesting to follow the thought that a sustainable societal system is a system that is *not* prone to transitional change, and thus does not exhibit conditions for transitional change.

Although a generic definition of sustainability is beyond reach according to Weaver & Rotmans (2006), they claim that commonalities can be found to make the concept operational but that doing justice to the integrative nature of sustainability would involve the use of concepts that cut through time, space and domain. In this sense, then, the concepts of tension, stress and pressure are suitable. This would amount to equating conditions for transitional change with symptoms of unsustainability.

That this makes sense is clear from the following: Tension, stress and pressure are necessary, though not necessarily sufficient conditions for transitional change. Transitional change fundamentally alters the way a societal system meets societal needs. If conditions are present that make the system prone to such processes that is a sign that the current functioning cannot be sustained. Hence, conditions for transitional change can be interpreted as symptoms of unsustainability.

Although the conditions for transitional change also carry their normative and subjective load this is of a different nature than that of sustainability. The normative, subjective and ambiguous aspects are intrinsic to the concept of sustainable development, since it implies a choice of what is worthwhile to sustain. With the conditions for transitional change these are the result of the system demarcation and the way functioning is described. Since adverse functioning, which is the key concept for the conditions for transitional change, is expressed in terms of functioning *itself*, no new normativeness, subjectivity, or ambiguity is introduced in identifying conditions for transitional change.

Regarding the concept of persistent problems, the argument is more straightforward. The definition of a persistent problem, as being rooted in the underlying structure of the system that suffers it, makes it a sign for a condition for transitional change. This is because the conditions per definition have their origin in the functioning, and thus the underlying structure, of the societal system.

Persistent problems themselves, however, do not drive transitional change. They can merely be regarded as signs that conditions for transitional change are present. Conversely, transitions can not be said to be solutions to persistent problems, however, a transition would lead a societal system to a radically different way of functioning that possibly does not produce the persistent problem.

Some more on the conditions for transitional change vis-à-vis persistent problems, perhaps. Although both concepts are abstract measures for issues, anomalies, or misalignments rooted in the very make-up of the societal system some

---

redefinition could be for instance: A sustainable societal system functions in such a way that it meets the current societal needs, without compromising the functioning of present, nor future societal systems.

## CHAPTER 3. PILLARS

differences can be identified. On the one hand, the level of abstraction, for instance, is higher when speaking in terms of conditions: a persistent problem can be pointed out in the societal system, as it were. On the other hand, speaking in terms of conditions is more specific about the nature and location of the issue in relation to the systemic structure, that is, within or between constellations or between a constellation and the landscape.

### 3.2.5 Connected conditions

In some cases some *apparent* ambiguity or subjectivity might arise in whether to regard something as a tension, as stress or as pressure. This is caused by the fact that some processes and mechanisms can lead to different adverse effects and thus different conditions for transitional change, moreover different conditions might affect the same constellation in a similar way. Following are examples here involving multiple related conditions for transitional change.

#### *Intensive agriculture*

Intensive agriculture can be characterised as agriculture that tends to maximise the meat or crop production per acre. The very same processes and mechanisms that make this possible, monocultures and factory farming, also have adverse consequences like plant diseases and animal epidemics that impede the functioning of the same agricultural societal systems. Thus, here animal epidemics and plant disease are symptoms of the persistent problems of intensive agriculture, and signs of stress.

However, the same functioning also can have adverse effects to *other* societal systems, the landscape, which would indicate tension. For instance the environmental pollution or the negative image public opinion about bio industry would be signs of tension.

#### *Mobility*

Some ambiguity might arise when a transition away from fossil fuel is considered. Such transitional change might just as well be driven by tension, for instance, the awareness of a link to climate change or simply oil depletions, as it might be by the pressure of upcoming alternatives like electric mobility.

#### *Energy supply and mobility*

One pathological example might be the following. What if the fossil fuel based regime experiences a condition for transitional change because society decided that the remaining fossil fuels are to be used for energy supply only? If energy supply and mobility are considered constellations in a larger 'energy use' societal system then this would be pressure. Pressure from the fossil fuel using energy supply constellation upon the mobility constellation, both claiming access to the same resources. However, if energy supply and mobility are considered constellations in *different* societal systems, then energy supply would be part of the *landscape* of the mobility system. The condition would then be tension.

### 3.3 Pillar II: Patterns

... are the ways transitional change takes place in a societal system.

Patterns describe how constellations are affected in transitional change. Three patterns are recognised: Reconstellation, a constellation emerging or gaining power from outside the societal system; empowerment, a constellation emerging or gaining power from within and adaptation, a constellation incorporating alternative functioning.

Patterns are formed when conditions are present. Societal transitions are sequences of patterns.

### CHAPTER 3. PILLARS

What actually happens when transitional change happens to a societal system? From the previous, one already guesses that transitional change is a turmoil of processes that alters the composition of a societal system. After a transition a societal system does different things, differently. What also can be concluded from the previous is that transitional change is more than just slight adjustments to the functioning of the constellations, rather it is change that affects the societal system at large. If one is to investigate how transitions work one could attempt to map out the myriads of processes that would underly the transitional change. Although this is presumably the preferable approach scientifically speaking it has some downsides as well. To begin with each ‘discovered’ mechanism or process will only exist in the form of a to be tested hypothesis, which is in itself no problem. However, there remains the matter of how these underlying processes or mechanisms add up and interlock to produce the system level transitional change that one is interested in. So one runs the risk of ending up with an incomplete puzzle, probably with some pieces wrong, that doesn’t yet show the bigger picture.

Another approach is possible too. Think of transitional change as an emergent phenomenon produced by these, as yet, unknown underlying processes and mechanisms. Then take recourse to what *is* known about transitional change, per definition. Transitional change alters the underlying structure of the societal system, by which its functioning is radically changed. In the pillar-theoretical language this is to say that the composition of the societal system is altered: after transitional change the relative power and functioning of the constellations in the societal system are different. Following this line of reasoning one arrives at the idea of simply identifying the things that can happen to a constellation in terms of power and functioning, things to do with constellations emerging, growing, adapting, and becoming alternatives for incumbent regimes — leaving aside for now the matter of what causes or drives this. These things will be called the *patterns* of transitional change.

First, distinguish between a pattern describing transitional adaptation of constellations and patterns that are concerned with their emergence and gain of power. A pattern describing transitional adaptation, which will be called simply *adaptation*, inevitably will entail more than the incremental, evolutionary change that constellations go through to keep meeting their societal needs. Rather, one can expect that this pattern tells of constellations absorbing innovations, possibly even merging their functioning with other constellations, or constellations co-evolving with others. Think for instance of biological products becoming part of the standard product range of regular supermarkets, which could lead to incorporation of this niche in the incumbent regime. This would result in a transitional adaptation of the food production and logistics regime, while retaining much of its original structural and cultural basis. Another example could be the cure regime in a health care system co-evolving with the niche of complementary medicine, transitionally changing both constellations, for instance a shift from an illness-centred to a patient-centred approach in the cure regime and a move towards evidence-based working in complementary medicine.

Second, among the patterns concerned with the emergence and gain of power of constellations, distinguish between those that describe these processes as working from outside the societal system upon a constellation and those that do so as working from within the societal system. To clarify this distinction somewhat consider the difference between how natural gas became the dominant

### 3.3. PILLAR II: PATTERNS

source for household heating in the Dutch energy system (see e.g. Rotmans *et al.*, 2001), a process which was very much government led, and how the automobile came to be the dominant mode of personal transportation, growing out of a small niche of racing enthusiasts (Geels, 2005 *a*). The former pattern, the outside-in or top-down one, will be denoted as *reconstellation*, whereas the latter, the inside-out or bottom-up one, will be denoted as *empowerment*. The patterns are tabulated and illustrated with some examples in table 3.2.

---

#### Patterns of transitional change

---

##### **Reconstellation** — *Top-down constellation change*

A new constellation emerges or an existing one gains power from *outside* the societal system.

*Examples:* Government led structural changes typically are reconstellation processes, be they inducing a switch from coal to natural gas as the main source for heating or centralising or privatising the financing of public services. Implementing European policies on a national level would also fall under this moniker as well as processes involved in socio-economic and political revolutions. Think also of constellations finding new societal functions, for instance railway companies using their infrastructures to offer telecommunication services.

##### **Empowerment** — *Bottom-up constellation change*

A new constellation emerges or an existing one gains power from *within* the societal system.

*Examples:* A new technology finding a market niche. Medical specialists unite to acquire a protected status for their profession. Lifestyle niches clustering with technological niches, like living ecologically aware and electric mobility merging to an eco-chique niche. Medication replacing psychotherapy as preferred treatment for mentally ill.

##### **Adaptation** — *Internally induced constellation change*

A constellation incorporates alternative functioning.

*Examples:* Oil companies moving into green energy sources. Incorporation of alternative health care practice in regular insurance schemes. Public transportation systems offering taxi services.

---

Table 3.2: The patterns of transitional change

Obviously, describing transitional change in this way, with patterns formulated like this, is a stylisation. They are not really ideal types but rather categories of transitional change. These are not empirically discovered patterns, they are deduced using certain criteria of distinction. This however, does mean that



## CHAPTER 3. PILLARS

they are *complete*. There are three and only three patterns<sup>15</sup>. The completeness argument is easily shown. A constellation undergoing transitional change means that it emerges, gains power, or it incorporates alternative functioning. When it incorporates alternative functioning it is adaptation. When emerges or gains power it — the power, or the constellation emerging — can come either from within or from outside the societal system<sup>16</sup>. Empowerment comes from within and reconstellation from outside. No other cases are to be found<sup>17</sup>.

The description of each pattern in the above was still rather sketchy, but in their separate sections 3.3.1, 3.3.2 and 3.3.3 they will be treated in some more depth. The processes and mechanisms that typically can be considered part of each of these patterns will be discussed there too.

Since the patterns are a complete set of descriptions of transitional change, this implies that any transition path can be described with them. A transition path is essentially the sequence of patterns, or perhaps several intertwined sequences, that trace the transitional change over time. This trace can be generated by a simulation model, a narrative, or even a mere list of indicators changing over time. The idea is that every such trace can be represented as a sequence of patterns of transitional change. This idea will be elaborated extensively in section 3.4 on paths. That patterns are building blocks for transition paths and are presented in a verbal form here, seems to imply a designated utilisation in narrative approaches or story lines. This is, however, only partly true. The systematic nature of the patterns, a simple consequence of their being deduced<sup>18</sup>, gives straightforward opportunity for formalisations. And with that the patterns also aim to be a stepping stone for computational and mathematical modelling approaches. One move in that direction will be made in chapter 4 where a formal form of the theory is given, a formulaic recast of the conditions, patterns and paths in pseudo-mathematical notation.

If one is to construct a dynamical model of a societal transition using the patterns it will not suffice that all transitional change can be described with them. One also needs to know which pattern will occur when. In other words, a dynamical model needs a way to decide which pattern to employ next in paving its modelled transition path. This is of course the difficult part, and it is also where the theory becomes speculative. But to even become a theory proper a gap needs to be bridged. A gap between the state of the societal system, or rather the conditions for transitional change, and the patterns which are expected to occur as a consequence thereof. A good theory<sup>19</sup> is capable of prediction in some form, or at least allows hypotheses thereon to be formulated and tested. For a dynamical model, as well, one needs to know which patterns are driven by what conditions. As said this coupling is speculative, with little

---

<sup>15</sup>Other patterns, belonging to other sets, can of course be found, just as other sets can be deduced.

<sup>16</sup>When there is only one constellation in the societal system, that is when the constellation *is* the societal system, the three patterns collapse into two: adaptation and reconstellation. This, because empowerment would mean emerging or gaining power from within the societal system, which is the constellation itself. The constellation exists and therefore needs not to emerge. If it would merge with another constellation, this other one would necessarily come from outside the societal system, making it reconstellation if it gains power or adaptation if it incorporates its alternative functioning, or both.

<sup>17</sup>Though sub-patterns might be identified, obviously.

<sup>18</sup>One could also say: postulated

<sup>19</sup>Good theory is discussed in Book I, chapter 1, section 2: Explaining theory.

### 3.3. PILLAR II: PATTERNS

more basis than the definitions and descriptions of the patterns themselves, other theories on transitions and societal change, scarce empirical knowledge, and common sense.

That pillars is a constructed theory is in this sense both a weakness and a strength. Weakness, as indicated, for the hypotheses are inevitably still to be tested and their falsification could render the theory defunct. Strength since it allows a construction where the hypotheses are, as it were, modules in the theory. For pillars this means that the coupling of conditions and patterns, the hypotheses on what patterns are driven by which conditions, needs to be separated from the conditions and patterns themselves. How this is done and the suggestions for hypotheses will therefore be treated separately in section 3.3.4. This manner of separating hypotheses from the core of the theory hopefully allows for its falsifiability in a nondestructive way.

The transitional change of patterns can apply to any constellation. Any constellation, be they niches, niche-regimes or the regime itself can undergo transitional change described by patterns. However, the regime, being the most powerful constellation in a societal system, is in a sense the main character in a transition. Therefore in the following sections that elaborate the patterns, the focus will be mostly on the regime or on how viable alternatives for it emerge and gain power. Consequently when is spoken of a pattern working on the societal system, rather than a specific constellation, what is meant is working with respect to the regime. This is much the same as in the treatment of the conditions for transitional change where the conditions could be those of any constellation, but if none is specified the regime is assumed as it dominates the functioning of the entire societal system.

#### 3.3.1 Reconstellation

The processes and mechanisms that are part of the reconstellation pattern all have in common that they contribute to an alternative constellation coming into being, whilst not being part of what they are forming an alternative for. If in a transition, a societal system experiences reconstellation, an alternative for the regime is formed by forces from outside that system. Since the dynamics governing the development of the functioning of a constellation was considered to be evolutionary and thus incremental, a process or mechanism is needed that changes the system from the outside in, so to speak. Either that, or the system harbours some nucleus of alternative functioning: a niche, which grows, spreading through and turning over the regime. Which is exactly the sort of process that is part of the complementary *empowerment* pattern.

Typically, the skeleton for a new or different form of societal functioning is erected, for instance, by building new infrastructures or setting up institutions and legislation. The new functioning's take off is backed by the powers that made it possible to erect it in the first place, confidence in the new constellation comes from advance studies or proven functioning in other contexts.

#### Reconstellation processes

Central to the idea of reconstellation is that the transitional change is imposed on the constellation or societal system. Typical processes involved are:

- *Reformative legislation*

## CHAPTER 3. PILLARS

Laws enforcing structural or cultural change, e.g. environmental law. Modifying, rather than codifying.

- *Installation of infrastructure*

Reforming the functioning of a constellation or a societal system by changing or creating its infrastructures.

- *Regional reorientation*

Land reform, re-parcelation, imposing new functions on certain regions like deforestation for agriculture, touristic exploitation of cities.

### 3.3.2 Empowerment

The processes and mechanisms that are part of the empowerment pattern contribute to an alternative constellation coming into being whilst it is already functioning within the societal system. If in a transition, a societal system sees a constellation becoming empowered, an alternative for the regime functioning rises by forces from within that system. As indicated under the complementary pattern of *reconstellation* this is another way transitional change can affect a societal system, where the alternative emerges under the eye of the incumbent regime.

Typically a new or different form of functioning is already in action in a niche where it develops alongside the regime. The niche constellation grows and expands its functioning by providing better fitting functioning than the regime or by merging and forming clusters with other niches. It can also gain influence by anchoring itself in higher institutional levels, getting patents, contracts arranging a preferential status, denominations or by becoming official in several senses.

#### Empowerment processes

The pattern of empowerment concerns constellations bringing transitional change to a societal system by becoming a more dominant factor in its functioning. Typical processes involved are:

- *Acquiring recognition*

Patents, emergence of standards, certification for products as well as for professions or education.

- *Forming unions*

Workers, trade or other unions, cartels, collusions, lobby and pressure groups.

- *Professionalisation*

Specialisation, education and learning, research and development, acquiring expertise.

### 3.3.3 Adaptation

The processes and mechanisms that are part of the adaptation pattern contribute to an existing constellation incorporating alternative functioning. In this sense it is a counterintuitive pattern of transitional change in that the existing constellation referred to is typically — although not necessarily — the regime, and the incorporation of alternative functioning into it would continue or expand the incumbent regime's dominance in the societal system. This does not readily appeal to ones idea of radical change and yet in many senses it is. The alternative functioning is incorporated exactly because the normal gradual evolution would not be enough for a constellation adapting under conditions for transitional change.

There are several ways alternative functioning can be incorporated in a constellation. A constellation can absorb functioning that is already reasonably well developed in a niche — possibly even as a societal innovation emerging in, and fostered by it. The functioning can also have its origins from outside the societal system and absorbing it would mean applying it in a different context<sup>20</sup>. Another possibility is that a constellation co-evolves with another constellation, whereby it takes over facets of its functioning (and vice versa) without incorporating the whole of it, leaving it to function on its own so to say.

#### Adaptation processes

In adaptation a constellation typically incorporates new or different functioning. Typical processes involve:

- *Re-positioning*  
Orienting towards new target groups or markets, assuming a new role or identity.
- *Re-organising*  
Structurally growing or shrinking, forming or merging groups, divisions or specialities.
- *Innovation*  
Mechanisation, automation, modernising infrastructure, employing new or different technology, knowledge or expertise.

### 3.3.4 Coupling hypotheses

Transitional change in this theory takes the form of patterns that form under certain conditions. The matter of which conditions lead to which patterns, however, appears to have been carefully avoided thus far. But, as mentioned earlier, these are the sort of statements one would like to be able to make with a theory like this one. Moreover, if it is to be a basis for modelling approaches it is a necessity.

If the theory would be built on such hypotheses the risk is that a falsification would render it useless. Still at this point in the development of the research

---

<sup>20</sup>Like in health care where psychiatry at a certain point begins treating patients in beds, by analogy with hospitals.

### CHAPTER 3. PILLARS

field it is unavoidable to work with tentative hypotheses on the causal factors in transitional change. The approach taken in pillars is to construct the theoretical framework in such a way that these hypotheses are treated as separate modules. As research furthers and insight in the dynamics grows they can be adjusted, replaced or discarded without having to discard large parts of the theory.

Why then even attempt tentative *coupling hypotheses* one might ask? Entire schools of thought even doubt if such causal statements can be made about processes in complex, societal systems. Apart from this theory taking a more positivist stance in that there is also another matter. Theory development profits greatly from interaction with empirical research and modelling approaches. For the latter again, causal hypotheses are a necessity. In a dynamical model setting the hypotheses can be studied as to how they influence the behaviour of societal systems in transition and how this compares to empirical data. This amounts to an approach to theorising as if it were an experimental science. In this scientific practice the model is the theoretician's laboratory.

In the following, tentative coupling hypotheses will be proposed as to which conditions drive which patterns. The hypotheses are proposed on strictly theoretical grounds, which at this point seems appropriate. The basis of available empirical data, which only scarcely has been analysed within this theoretical framework, is too weak to formulate sensible hypotheses. Something similar holds for utilising other theoretical work, which itself is often still speculative and would moreover need to be adapted to fit into the pillars<sup>21</sup>. The hypotheses will also have a somewhat open character, in that they give an idea of the sort of dynamics to expect without making them operational to the point that they already comprise a dynamical model.

Since at this point the effect of the patterns has not been cast in quantified form the hypotheses too will only speak of driving or not. A condition driving a pattern is to be interpreted as the presence of a certain condition stimulating the pattern to manifest itself. If the condition is stronger the influence is stronger and the pattern will form more clearly and or rapidly. There is no condition that has a negative influence on a pattern for this would imply that transitional change would be reversible, which as was already argued in section 3.1.3, it is not. The proposed scheme of influences is given in table 3.3.

---

<sup>21</sup>Which would not be impossible, though.

### 3.3. PILLAR II: PATTERNS

| Condition | Pattern           | Consequence  |
|-----------|-------------------|--|
| Tension   | → Reconstellation | { A constellation emerges or gains power from outside the societal system                              |
|           | ↘                 |  |
|           | ↗                 |  |
| Stress    | → Adaptation      | { A constellation incorporates alternative functioning from outside or from within the societal system |
|           | ↘                 |  |
|           | ↗                 |  |
| Pressure  | → Empowerment     | { A constellation emerges or gains power from within the societal system                               |

Table 3.3: Hypotheses on which conditions drive which pattern

The general idea behind this coupling scheme is the following: If for instance reconstellation is the pattern of transitional change working from the outside into a societal system, for instance through government imposed legislation, then it is not far fetched to assume that the drivers behind such actions also couple to what's outside the system, which implies tension as a driver. Similarly reconstellation could also be invited, as it were, if the driver for transitional change is some internal issue which the constellation itself is not able to resolve, this would be a case of stress drives reconstellation.

Why then not pressure driven reconstellation, one might ask? This would imply that the presence of a viable alternative for a constellation drives an influence from outside the societal system to impose transitional change on it. Although in itself this is conceivable it would then be *empowerment* of the alternative constellation driven by the pressure it exerts or *adaptation* of the constellation experiencing this pressure. If one wants to do some theoretical nit-picking on this and insist on reconstellation then it is still the case that the influence of the alternative constellation (the 'pressure') must find a way to do its influencing from *outside* the system, rendering the influence a *tension*.

Applying symmetry to these arguments it is clear that the empowerment pattern can be driven only by pressure and stress and the only pattern driven by all conditions for transitional change would be adaptation. This makes sense for reasons other than symmetry as well, since adaptation is the pattern where the constellation changes from within<sup>22</sup>, making it susceptible for influences from all 'directions'.

This scheme leaves open several things. Conditions can be identified for any constellation; any constellation can be under tension, stress or pressure. Moreover, the consequences mention only *a* constellation for the pattern to affect. Also there is no mention of what functioning a constellation that emerges

<sup>22</sup>Alright, a symmetry argument as well if you like...

## CHAPTER 3. PILLARS

or gains power will have, nor what functioning will be incorporated through adaptation. Thus one only gets an idea of what sort of dynamics is to be expected under the identified conditions. Of course this can be made operational to a much farther extent, which would amount to making a dynamical model. For such an attempt many additional assumptions need to explicitly be made, which has been done in Part III, chapter 2, which elaborates the theory in a computer model.

For a transition analysis one could, for instance, make the scheme more operational by assuming that all conditions are those of the regime. Since the regime dominates the functioning of the societal system, this implies that its conditions also dominate the dynamics that they produce. Then one can proceed to scan for the patterns that are implied to form, according to the scheme. One could also track each constellation over a period of time and identify the patterns and conditions that occurred. In this fashion a possibly complex history can be unravelled as several intertwined developments. This has been done in Book III, chapter 1, which elaborates a historical case study on Dutch health care.

It might be appropriate and about time to look somewhat closer at the various forms of transitional change that can occur following this scheme. Apparently the three patterns and their three driving conditions are coupled in such a way that seven principal forms of transitional change are possible. In the following, those forms will each be very briefly introduced and followed by two examples, with the first ones for each form coming from the health care case study mentioned above.

### **Reconstellation driven by tension**

Problems or issues with the environment of a constellation or societal system drive transitional change coming from outside it:

- The (partial) market liberalisation of health care financing to face the rising health care expenditure.
- The Dutch *Deltaplan* for protection against storms and floods from the North-Sea. A large infrastructure of dykes was installed on government instigation.

### **Reconstellation driven by stress**

Internal issues or problems of a constellation or societal system drive transitional change coming from outside it:

- The 1865 'Thorbecke'-laws arrange and codify many aspects of health care and its practicing, requiring the passing of state exams to become a physician, amongst many other things.
- The post WWII emergence of the West-German state around a structure instated by the western allied forces.

### **Adaptation driven by tension**

Problems or issues with the environment of a constellation or societal system drive transitional change coming from within it:

### 3.3. PILLAR II: PATTERNS

- The advances in medical science and increasing demand for specialist treatment in polyclinics lead to an increase in the number of specialist physicians, driving re-organisation in the professional association (NMG) and the position of the specialist with respect to e.g. the general practitioner.
- Increasing environmental awareness and market opportunism lead large industrial companies to invest in and market 'green' options like the Toyota Prius and Shell hydrogen and bio fuels.

#### **Adaptation driven by stress**

Internal issues or problems of a constellation or societal system drive transitional change coming from within it:

- Hygiene, a- and antisepsis, hygiene and anaesthetics enter the hospitals changing these institutions into places of non-lethal treatment and surgery.
- Re-orientation of the Rotterdam harbour towards container transshipment.

#### **Adaptation driven by pressure**

The presence of a viable alternative for the functioning of a constellation or societal system drives transitional change coming from within it:

- The rise of the specialists within the medical community leads to the hospitals becoming more and more places of specialist treatment and surgery, as opposed to general care.
- Telecommunication systems changing their role and functioning in accommodating the immense growth of mobile communication.

#### **Empowerment driven by stress**

Internal issues or problems of a constellation or societal system drive transitional change by letting an alternative constellation arise:

- General practitioners unite *within* the medical association into what later becomes the NVH (Dutch Association of General Practitioners) as they feel their position threatened by the upcoming group of specialists. Similarly the pharmacists unite (into the NMP, the Dutch pharmacists association) to secure their position when it is called into question when they remain unrepresented in a state-committee. In both cases emancipation of the professional groups within the medical community resulted.
- *Coups d'état, golpes d'estado, putsche, etc.* where a part of the societal system, typically the military, assumes power in times of crises or turmoil. As opposed to revolutions, where the overthrowing constellation comes from *outside* the societal system classifying them as reconstellations (driven by tension or stress, depending).



## CHAPTER 3. PILLARS

### Empowerment driven by pressure

The presence of a viable alternative for the functioning of a constellation or societal system drives transitional change coming from an upcoming alternative constellation:

- During the rise of the cross-organisations<sup>23</sup>, setting the standards for nursing and its education, they take a central and rather autonomous role in public hygiene and preventive health care.
- The rise of the electric guitar to a central position on the stage of modern music.

The attentive reader undoubtedly has noticed that phenomena like businesses or entire sectors growing through factors like increasing demand, economies-of-scale, mass-production, etc. have not found a place as examples in the above. This is not because these processes are not important in transitional change, on the contrary, but rather that they can fall under different categories in different circumstances. For instance, in the health care case there is the example of the pharmaceutical industry growing rapidly after the World Wars, because of industrialisation, new technological and medical possibilities and growing demands. This changed the constellation of pharmacy tremendously, but for the societal system of health care as a whole it was a process within the pattern of adaptation driven by tension. Conversely the self-propelling rapid growth of mobile communication was a process within the pattern of empowerment of the niche of mobile telephony, driven by pressure<sup>24</sup>.

---

<sup>23</sup>In the Netherlands, home and community care organisations of religious denominations.

<sup>24</sup>Compare to the example of *Adaptation driven by pressure*, also involving mobile communication.

### 3.4 Pillar III: Paths

...are the trace of transitional change over time.

Paths are concatenations of patterns, sequences describing how the societal system evolves during the course of a transition. Any path can be represented as such a sequence and the possible dominance of a certain pattern yields families of ideal types.

Four families of paths are recognised: Top-down paths where reconstellation is dominant. Bottom-up paths dominated by empowerment. Squeezed paths where both reconstellation and empowerment are of large influence, and transformation paths dominated by adaptation only.

### CHAPTER 3. PILLARS

A transition *path* is how a transition unfolds itself, the actual manifestations of transitional change in the course of time. To begin to understand transitions one must at least be able to describe these paths well. A systematic way to tell transition tales is one step in the direction of making narrative case studies of transitions comparable, which in turn gives greater value to the conclusions that can be drawn from empirical data as well as making it easier to generalise and form hypotheses with a broader scope. This section is about transition paths, how they can be described using the language and theoretical apparatus developed so far, how transition paths can be made comparable and how several ideal types can be distinguished.

One could consider a transition path the story of transitional change happening to a societal system. A play telling of the rise and fall of constellations in their efforts of meeting societal needs in times of change. Or, adopting a more mechanistic world-view for a moment, the time evolution of a societal dynamical system. A trajectory through the phase space of the societal system, connecting the points describing the system state in terms of a number of chosen variables, as it goes through the transition. This already illustrates that there are many ways of looking at transition paths: some focussing on the subsequent states of the system, some on the processes of change themselves, some tracking a single indicator throughout and some continuously sketching the bigger picture. Whichever view is taken, the challenge here will be to describe this view with the carefully chosen words of the previous sections.

The difficulties with giving a proper description of a transition path are in a sense the difficulties encountered with describing the state of a societal system (its composition and its condition<sup>25</sup>) compounded with those of describing transitional change (the patterns). It probably is possible to describe the course of a transition by describing how the composition of the system develops over time, or how the conditions change as the transition takes place. Just as it appears possible to only list the patterns that have subsequently altered the system state. However, each of these descriptions on its own probably would not provide a satisfactory overview of the transition path, so a combined approach seems desirable. To warrant the comparability and the possibility to distinguish ideal types that was desired as well, this approach also needs to be sufficiently systematic.

Whereas the composition and condition are rather specific to the societal system at hand, the patterns are more generic. It therefore seems appropriate to use the sequence of patterns as the skeleton of the description of a transition path. This will also be instrumental in comparing transitions because one can make statements like: “Unlike this transition which has a lot of reconstellation, this one is mostly a sequence of empowerment and adaptation patterns.” It is also a key to identifying ideal types: “This transition was reconstellation all the way.” To make this more specific to the transition path one is describing, before each pattern the composition and condition of the societal system needs to be given. To employ the theatrical metaphor once more: Each pattern is a dramatic situation<sup>26</sup>, influencing the plot of the transition play. The composition and conditions provide the stage description and the motivation of the protagonists.

---

<sup>25</sup>On the plural and singular use of ‘condition’: the conditions describe the condition of the societal system.

<sup>26</sup>This draws from Polti’s (1868) idea that there is a limited number (36 in his case) of distinct dramatic situations that can occur in any given work of drama, story, play, etc.

### 3.4.1 Chains of change

A transition story line tells a tale of a societal system going through transitional changes. Typically this involves describing the state of the system prior to the transition and whether there are signs that might indicate if the system is prone to change. Then one describes how the system changes and in what state it is after that. Depending on what is desired from the story line these descriptions are iterated or not and are superficial or highly detailed. This can be stylised in a straightforward manner using the tools from this pillar: first the system state is described in terms of its composition and condition then the transitional change is described by the relevant pattern and the new system state is given. This is iterated until the transition path has been run through, producing a chain of patterns changing the system state.

For example, consider a societal system, not further specified, and describe its composition. The composition would be an inventory of the constellations and their relative power and the landscape, or, using the constellation ideal types, this would be a list describing the regime, relevant niches, present niche-regimes and the landscape. Then describe the condition of the societal system — by indicating if there is tension, stress or pressure — that possibly drives transitional change. The composition and condition now provide a description of the system state. Next, give the pattern that changes the system state and give the result of the working of the pattern by describing the new composition and conditions of the societal system. This, a sequence of system state — pattern — new system state, is the smallest chain of transitional change possible. Obviously from the new system state one can obtain a following system state by concatenating another pattern to the chain. This procedure is repeated until the transition path is adequately described.

A systematic description as meant above can be easily cast in schematic form with the transition path brought back to its essential building blocks. Such a scheme would consist of naming the constellations that are about to be changed by the pattern that follows and the present conditions for transitional change, then the pattern followed by the new system state, and so on like in table 3.4 where reconstellation and subsequent adaptation leads to a transition.

|                           |   |   |   |                 |   |
|---------------------------|---|---|---|-----------------|---|
| Composition               | / | Condition   | → | Pattern         | → |
| Regime                    | / | Tension and stress on regime                                  | → | Reconstellation | → |
| → Regime and niche-regime | / | Tension, stress and pressure on regime                        | → | Adaptation      | → |
| Post transition:          |   | New regime from niche-regime, old regime perished or subsumed |   |                 |   |

Table 3.4: Transition path featuring the reconstellation and adaptation patterns

### CHAPTER 3. PILLARS

Table 3.5 gives a similar example of a simple chain where empowerment, followed by adaptation pulls a societal system through a transition.

| Composition               | / | Condition   | → | Pattern     | → |
|---------------------------|---|---|---|-------------|---|
| Regime, two niches        | / | Stress present on the regime                                  | → | Empowerment | → |
| → Regime and niche-regime | / | Stress and pressure on regime                                 | → | Adaptation  | → |
| Post transition:          |   | New regime from niche-regime, old regime perished or subsumed |   |             |   |

Table 3.5: Transition path featuring the empowerment and adaptation patterns

A hypothetical — and only slightly idealistic — transition story corresponding to the transition path in table 3.5 could be the following. Consider an agriculture and food-supply system where the incumbent regime consists of intensive agriculture with distribution to supermarkets. The system suffers from stress, of which regular outbreaks of animal diseases are symptomatic. Two niches are present in the system, one distributes biological products via a subscription and home-delivery system and the other is an innovative form of mixed, non-intensive agriculture where crops and meat are produced on small scales, in the region where these will be consumed.

The stress present on the system empowers these two niches and allows them to merge, forming a complete, decentralised, production and distribution system based on small farms. The growing presence of this new niche-regime puts pressure on the regime, while it also still suffers from the stress mentioned above. This eventually leads to adaptation of the intensive agriculture and supermarket regime, where intensive agriculture is gradually phased out and supermarkets focus on other products. Then the niche-regime effectively becomes the new regime and a transition would have taken place.

In this way, in principle any transition path can be described with a sequence of patterns as backbone. To say exactly ‘where’ the transition occurred remains a matter of interpretation, what counts is to be able to tell the transition story. Whether only a limited amount of paths are actually relevant is in essence an empirical question. It is, however, possible to identify ideal-typical paths in which certain patterns dominate the course of the transition, more on this in the next section. A real example of using sequences of patterns to unravel transition paths is done in book III, chapter 1, which is an elaborate case study about the history of Dutch health care.

#### 3.4.2 Typical paths

The multi-pattern approach to societal transitions can be exploited in another way as well. Instead of using the patterns as building blocks for transition

### 3.4. PILLAR III: PATHS

paths like was elaborated in the preceding part, one can also classify transition paths according to the dominance one pattern or another has had in the course of a transition. In other words: one can use patterns to produce a typology of transition paths. This would be a deductive typology as opposed to an inductive one, in which the criteria of distinction would be abstracted from a collection of documented transitions.

The typology presented in the following has as its first criterion of distinction the pattern that dominates the course of the transition. That gives, at a first glance, three main types: one for reconstellation, one for empowerment and one for adaptation dominated transition paths. The reconstellation pattern suggests transition paths unfolding ‘top-down’, in contrast to the empowerment pattern that would have a ‘bottom-up’ character. Domination of the adaptation pattern would imply a sort of internal transformation and although it can even be disputed whether it would still be a transition if the regime adapts without the influence of other societal subsystem, ‘transformation’ paths are included here.

Whereas it is disputable whether adaptation can lead to a proper transition, the above paths completely exclude the possibility that a transition is the result of being ‘squeezed’ between ‘top-down’ and ‘bottom-up’ influences. This idea of a transition resulting from an undercurrent connecting to landscape developments, however, is central to some angles to transition thinking and especially in the context of governance and policy making (e.g. Rotmans & Loorbach, 2006). Hence squeezed transition paths, where reconstellation and empowerment both shape the course of the transition are included in this typology. Thus four families of transition paths are obtained: top-down, bottom-up, squeezed and transformation paths.

The second dimension of the typology is the influence of the incumbent regime on the unfolding of the transition. Or, which is in a sense equivalent, the extent to which the regime steers, manages or coordinates processes in the transition. This obviously corresponds to the role of the adaptation pattern in the course of the transition<sup>27</sup>. Additionally there is the realisation that transitions sometimes, maybe even most of the time, leave the system unaffected, get stuck or fail in other senses. Thus for each main type a failed transition path is included in the typology as well. This produces a total of eleven transition paths, ideal types of transitions tabulated in 3.6, 3.7, 3.8 and 3.9, below which a small discussion of the typology with respect to existing typologies is given.

---

<sup>27</sup> Earlier versions of this typology referred to this as much/little regime influencing. Calling this by its name: adaptation, simplifies the typology in that it bases itself even more on the pattern influences. It also increases its symmetry in that all ‘families’ of paths have a role for the adaptation pattern and the families are deduced by which pattern dominates the transition in addition to the presence of more or less adaptation.

<sup>28</sup> Goldstone proposes a broad and contemporary definition of revolution, combining elements of many earlier theories and definitions by: “an effort to transform the political institutions and the justifications for political authority in a society, accompanied by formal or informal mass mobilization and non institutionalised actions that undermine existing authorities.”

---

**Top-down paths** — *Reconstellation dominated*

---

- Radical reform (with adaptation)

Here the regime is reformed according to the cultures and structures of some outside constellation. Examples are the framework directives of the EU forcing local governments to reform their institutional structures. Indeed many institutional transitions appear to be of this type (North, 1990). The description by Frantzeskaki *et al.* (2007), after Lane (1997), of public sector reforms, places them in this category as well. If the viewpoint of Ovin (2001) is taken then also the democratic transitions of former East-block countries can be classified as this type.

- Revolution (without adaptation)

This is the case when a constellation outside the regime invades the societal system and replaces the incumbent regime. Political and socio-economic revolutions can be considered of this type (for instance if one takes Goldstone's (2001) recent definition<sup>28</sup>). Also characteristics of this type are found when governments (not to be confused with the regime here) provide new structures in sectors that were formerly self-regulating or locally arranged, centralising health care financing, for example.

- 
- Collapse (failed transition)

A reconstellation dominated path does not necessarily lead to a new stable societal system. Think for instance of failed revolutions leaving states in chaos for years to come or of foreign forces failing to rebuild a nation after an invasion. Also radical reform transitions like privatisations of entire sectors can lead to deterioration of the societal functioning of such a sector, like the privatisation of the railways in New Zealand that are currently being undone. Similarly large scale collectivisations have been known to lead to badly functioning systems, like the collectivisation of Cuban agriculture.

---

Table 3.6: Top-down transition paths (reconstellation dominated)

---

**Bottom-up paths — Empowerment dominated**


---

- Reconfiguration (with adaptation)

One or several niches are empowered and become a niche-regime which in turn takes on the role of the dominant constellation, thus becoming the new societal regime.

Basically this path is similar to Geels & Schot (2007) socio-technical reconfiguration path, although their re-alignment/de-alignment pathway can also easily be classified as this type since in essence it is about niches scaling up and becoming empowered *by* the regime and not *in spite of it* like in the following ‘substitution’ path.

- Substitution (without adaptation)

Like reconfiguration with the difference that the regime is *not* involved in the empowerment process. The niche or niches scale-up on their own force or merits and the subsequently emerging niche-regime successfully competes with the regime and takes its place.

Obviously technological substitutions are of this type, but one also thinks of new popular political parties invading the political landscape.

—

- Backlash (failed transition)

Backlash refers to niches initially gaining power or popularity, for instance in the form of a hype, while their novel or deviant way to meet societal needs still fails to become the new mainstream. This can happen if the demand for certain functioning grows too rapidly and the niche is unable to consolidate this and consequently cannot cope with, for instance, swings in demand. Another possible path to a backlash can be that some novel functioning is initially adopted by many until some unforeseen risk or problem becomes apparent and the niche is abandoned. Backlash was described as a possible transition path by van der Brugge & Rotmans (2006).

---

Table 3.7: Bottom-up transition paths (empowerment dominated)



- Teleological (with adaptation)

These transitions are the result of a regime adapting to changed circumstances not by reforming itself but allowing outside influences to reconstellate structures and cultures and simultaneously incorporating novel functioning in these processes. If the regime is actively adapting and steering in transition processes, connecting niches with landscape developments, one can almost speak of a managed transition. This led to the idea of transition management (e.g. Rotmans & Loorbach, 2006; Loorbach, 2007) which is a governance framework to propagate this type of transition. This framework is currently being implemented in, amongst others, the context of the Dutch energy transition (Loorbach, 2007) and a transition in long term Dutch health care (van Raak, 2010).

This obviously corresponds largely to Smith *et al.* (2005) “purposive transition” characterised by high co-ordination and much influence from outside the regime. In Rotmans’ ‘cartwheel’ typology (Rotmans, 2005) of transitions they appear as one of three dimensions along which one can classify transitions, the others being co-ordination like in (Smith *et al.*, 2005) and level of aggregation adding a sense of scale.

- Emergent (without adaptation)

A transition where niche functioning and influences from outside the societal system somehow team up to a transition without active influence of the incumbent regime appears counter-intuitive. An example, however, could be the communication revolution of the past decade(s). Niche functioning like mobile telephony and internet use, empowered by technological developments and reconstellated through liberalisation of communication markets, led to a transition from the then reigning communication regime, roughly speaking characterised by phone, fax and postal services.

This transition path would correspond roughly to what Smith *et al.* (2005) describe as “emergent transformations”, where low co-ordination is combined with influences and novelty from outside the regime. This also corresponds to the emergent transitions in Rotmans’ cartwheel (Rotmans, 2005), the pendant of teleological or targeted transitions there.

Continued on next page...

---

**Transformation paths** *Adaptation dominated*


---

- Transformation

The regime essentially transforms on its own, successfully adapting to the point that societal needs are again met adequately and no tensions or stress plague it. Adaptation can have taken the form of absorbing niche functioning or co-evolution with it.

This transformation path is similar to the socio-technical one of Geels & Schot (2007). If interpreted as possible unfoldings of transitions, the transition contexts of Smith *et al.* (2005): “reorientation of trajectories” and “endogenous renewal”, are of this type as well, because of their focus on the “internal adaptation” of the regime.

—

- System breakdown (failed transition)

Attempts of the regime to adapt do not lead to a new societal equilibrium and the regime cannot meet the societal needs adequately any more. This failed transition path was already identified by van der Brugge & Rotmans (2006). Various forms of societal collapse can be classified under this type.

---

Table 3.9: Transformation paths (adaptation dominated)

---

**Squeezed paths** — *continued*


---

—

- Lock-in (failed transition)

This can be the case if, stimulated by government schemes or international pressure, the regime incorporates and stimulates a different way to fulfil a societal need, say a new energy production technology or liberalisation of internal markets, and this novel way is not sufficiently developed or too incompatible with current structures or demands. This would then lead to the situation of an alternative functioning in the margin of the mainstream or in a troublesome manner.

Van der Brugge & Rotmans (2006) describe this as a transition path where an innovation does gain influence in the societal system but fails to completely replace the regime co-existing with it in a locked in state.

---

Table 3.8: Squeezed paths (reconstellation and empowerment of similar influence)

### CHAPTER 3. PILLARS

As one can already see from the above tables, several other typologies of transitions, be they societal, socio-technical or institutional, fit in this typology. Comparing typologies is a cumbersome matter if one cannot simply compare the dimensions. In the transition literature some typologies use dimensions that cannot easily be compared to the one presented above. For instance Smith *et al.* (2005) use “resource locus” versus “steering of adaptive response”, and van der Brugge & Rotmans (2006) who use the shape of transition curves to classify different types of failed transitions, and Rotmans’ cartwheel (2005) which uses an ‘aggregation’ dimension, in addition to an emergent vs. targeted and a high/low coordination dimension that are straightforwardly interpreted in this typology. Another difficulty arises if the typology has an inductive character basing itself on many historical transitions (like that of Geels & Schot, 2007). In all these cases an attempt has been made to locate the specific types in those typologies in this one. As stated, in most of the cases this was possible and the types were incorporated in the preceding tables.

## Chapter 4

# Formalisation

## 4.1 Beforemalisation

Formalisations of theories about social phenomena are not quite common, let alone of their entire theoretical framework. Therefore some considerations before embarking to cast pillar theory in a formal form.

A good theory should be able to describe, explain and predict<sup>1</sup> phenomena within its range of application. For this, it is not necessary for a theory to have a formal form. Moreover, if the audience for a theory will, at least for the larger part, consist of scholars with no mathematical background, why would it be desirable to pursue a formal form?

For various reasons. When it concerns description, explanation and prediction one can maintain that a theory in a formal form leaves less space for the ambiguity natural language tends to bring along. True, formal representation might not capture all the aspects that a intuitively appealing verbal description might, but any description will in some sense come short and a formal representation will at least be more strict. This holds as well for explanation and prediction. If the theoretical framework itself is cast in an unambiguous symbolic language, then its explanatory devices can be employed in a more systematic way, as mechanisms in the Bunge (2003) sense<sup>2</sup>.

One a more general level, not all formalisations are necessarily quantifications. The formalisation proposed in this chapter, however is. Although kept in the abstract, that is, no indicators or indexes are actually employed, the core concepts are represented by numbers, vectors and the like. This implies that it is *in principle* possible to have a quantitative representation and *interpretation* of concepts like functioning, this, the practical difficulties in finding an operational form thereof notwithstanding. Even if such an operational form would be infeasible then it still serves to make the structure of the theory transparent. This brings up an even more pressing reason to pursue a formalisation: Recasting a theory in a formal form exposes it in its skeletal form, its bare essence. The explanatory mechanisms can be treated *as such*, causes can be separated even more clearly from consequences, assumptions from conclusions, and once consistent in its formal form a theory thus has analytical rigour.

Then there is a third, more practical reason to pursue formalisation. The ambition to be a basis for computational and mathematical modelling approaches. Such approaches obviously *are* formalisations and supplementing the theory with a formal form can in this sense be seen as a proof of principle. As presented here the formalisation still does not provide a dynamical model in itself but rather is a rigorous version of the rest of this book. Based on this formalisation, however, a computer model was made and elaborated in Part III, chapter 2: Building a model on pillars. It must be noted that the formalisation presented here is not the only possible formal representation of the theory. Just as many models can be made for the same phenomenon, several formalisations in principle are possible, with the symbolic language presented here obviously reflecting the predilections of the author. The leading idea, however, was to remain as close to the theory as possible, representing it in its essential form.

---

<sup>1</sup>Prediction is here understood as a statement that in principle can be tested.

<sup>2</sup>See also Book I, chapter 1: Understanding explanation.

## 4.2 Functioning and constellations

### 4.2.1 Parametrising functioning

Many parametrisations of functioning are possible just as many ways are possible to describe a societal need and the way a societal system would meet it<sup>3</sup>. For example in (Schilperoord *et al.*, 2008) the functioning of regimes and niches is a point in a  $n$ -dimensional ‘practice’ space and in (de Haan, 2008a, which is also chapter 3 in Part III) where functioning is a real valued function on a domain in  $\mathbb{R}$ . The parametrisation here is suggested by the idea that functioning has several facets. For example if a health care system is considered then one facet could be specialist treatment and another one could be insurance, different aspects of the societal system with different associated actors and institutions. In an energy supply system, providing the infrastructure for electricity could be a facet, as could be the policy regulating the energy market.

Describing the functioning of a societal system then comes down to producing a list of these facets until a satisfactory picture emerges<sup>4</sup>. Some facets will be much more prominent in the system’s functioning than others obviously. To accommodate this the facets will need to be weighted in some manner. Then why not simply let them be a number, say, a real one. With the interpretation of a negative one being that of inhibition of the facet in question, for example a technological solution actively not incorporated in the system (like the situation for nuclear energy in the Netherlands in the 1990s) or a service explicitly not offered (like services for private enterprise in socialist systems). Assume that the functioning of a societal system consists of  $n$  facets, which can be denoted as  $\varphi_i$  with  $i$  running from 1 to  $n$  and  $\varphi_i \in \mathbb{R}$ . The functioning of an entire system then corresponds to some vector  $\vec{\varphi}$  in  $\mathbb{R}^n$  with components  $\varphi_i$ . Thus, some notation:

Let functioning be defined as

$$\vec{\varphi} = (\varphi_1, \dots, \varphi_n). \quad (4.1)$$

With each  $\varphi_i$  corresponding to a facet;

$$\varphi_i \in \mathbb{R}, \text{ and consequently } \vec{\varphi} \in \mathbb{R}^n. \quad (4.2)$$

Variations upon this theme are very well possible. One can group several facets to represent various aspects of societal functioning. For instance a total of, say 27, facets grouped in three sets of nine facets each representing a ‘P’ in a triple-P sustainability model, or the sets representing structures, cultures and practices respectively. If extra clarity is desired one can always make such a cluster a separate vector,  $\vec{\varphi}$ ,  $\vec{\psi}$ , etc. Pushing this idea somewhat further for a dualism in functioning, like in the theory presented in this book, the functioning vector could be complex<sup>5</sup>, that is  $\varphi \in \mathbb{C}^n$ , and have the real parts correspond to the structural side and the imaginary parts to the cultural side of each facet. Since the complex structure of  $\mathbb{C}^n$  will not be used in this chapter the functioning vectors are taken as elements of  $\mathbb{R}^n$ . All can be straightforwardly generalised to complex vectors, so no potential generality is lost here. From here on bold face will be used to denote functioning vectors, i.e.  $\boldsymbol{\varphi} \in \mathbb{R}^n$ .

<sup>3</sup>See also 3.1.2 for more on describing functioning.

<sup>4</sup>What constitutes a satisfactory picture or if it is at all possible to attain such a thing is an entirely different matter.

<sup>5</sup>With boldface distinguishing a complex vector from a real one.

### 4.2.2 Societal systems and constellations

In the previous section the notation concerning functioning was expounded. With this it is now possible to formally discuss societal systems. Let a societal system, not further specified, be denoted with the symbol  $\mathbb{S}$  and, identifying the system with its functioning,  $\mathbb{S} \in \mathbb{R}^n$ .

Already mentioned earlier was the notion that functioning is the way a societal system meets a societal need. If this be so then the societal need can be described in similar terms, as a sort of desired functioning for instance. A welcome chance to introduce some suggestive notation. Let the societal need met by a societal system  $\mathbb{S}$  be denoted by  $\bar{\mathbb{S}}$ , again represented by some vector in  $\mathbb{R}^n$ .

The functioning of a societal system was considered to be composed of that of several constellations. This implies that each constellation itself has a functioning, represented by a vector in the same linear space as that of the entire societal system. So, if an arbitrary constellation be denoted by  $\mathbf{c}$  with  $\mathbf{c} \in \mathbb{R}^n$ , then summing over all  $N$  constellations present gives the functioning of the societal system as a whole<sup>6</sup>:

$$\mathbb{S} = \sum_{k=1}^N \mathbf{c}_k. \quad (4.3)$$

Let not a *constellation*  $\mathbf{c}$  be confused with a *facet*  $\varphi_i$ , here. A constellation is a subsystem of a societal system, like personal mobility in the case of a mobility system, and is therefore a vector having facets as components, which in turn could correspond to the technological options (car, bicycle) or infrastructural aspects (petrol stations, roads) they entail. Since facets are numbers this provides a way to give nuance to the functioning of a constellation (e.g. personal mobility has emphasis on fossil fuel facets and more so than, for instance, public transportation).

The societal system was said to be composed of constellations. This means it is also a sort of a collection, a *set*. To avoid introducing new symbols or adding diacritics,  $\mathbb{S}$  will play a double role here. One as a vector, being the sum of the vectors of the constellations, and one as a set, being the *composition* of the constellations. That is,  $\mathbb{S}$  also denotes:

$$\mathbb{S} = \bigcup_{k=1}^N \mathbf{c}_k. \quad (4.4)$$

Whether the set or the vector is meant will be clear from context.

### 4.2.3 Comparing functioning and power

When are two constellations alike, one could ask. When their functioning is, would be the obvious answer. In short, there is need of a way to compare functioning. First, it would be convenient to have a measure for the ‘size’ of a constellation, a measure of the proportion of functioning it represents with respect to that of the entire system. Such a measure was earlier already interpreted as the power of a constellation. One expects the power of a constellation

<sup>6</sup>To further avoid confusion  $k$  will be used to index the constellations (from 1 to  $N$ ) and  $i$  as the index for the facets, the components of the constellations.

## 4.2. FUNCTIONING AND CONSTELLATIONS

to be greater if it contributes more to the total functioning of the system. All other things being equal this holds per facet, that is, a constellation that has a facet  $i$  twice the value of that of some other constellation, can be considered twice as powerful regarding that facet. There is no need to compare these to the facets of the entire system, since they were the sums of those of the constellations already. So, for a constellation, all the facets added up would give a measure of its power. Not a very good one though, since the contribution of one facet could be levelled out by a negative contribution of another. The sum of the squares of the facets, or its square root will do nicely however. Choosing the symbol  $\pi$  for the power of a constellation<sup>7</sup>,

$$\pi^2 = \sum_{i=1}^n c_i^2, \quad (4.5)$$

is obtained, which is simply the absolute value of the functioning-vector of the constellation. The square of the total power to be distributed in the societal system is then

$$\Pi^2 = \sum_{i=1}^n \mathbb{S}_i^2. \quad (4.6)$$

Note that the total power is *not* equal to the sum of the power of the constellations.

Now, if two constellations are equally powerful, they can differ only in how they have their power distributed over their facets, internally, and they will be more alike the more their corresponding facets have values close together. A constellation is most alike to itself — and most unlike its negative, both of which have the same power. So, if for instance the sum of the products of the corresponding facets of the constellations is taken as a measure of ‘aliqueness’, this measure can vary between plus or minus the square of their power. In fact, the constellations need not equally powerful to for this measure to be sensible as it will vary between plus or minus the product of their powers. This invites the use of the sum of the products of the facets to compare functioning of constellations.

Since functioning was represented by vectors the above simply says that two vectors of equal length differ only in direction. The measure proposed is simply an inner product on the functioning space, with an interpretation of ‘alignment’ of constellations as well size or power. The inner product of two functioning vectors  $\varphi$  and  $\psi$  is then

$$\varphi \cdot \psi \equiv \sum_{i=1}^n \varphi_i \psi_i. \quad (4.7)$$

A concept like the power of a constellation  $\mathbf{c}$  is now straightforwardly redefined as

$$\pi_{\mathbf{c}} = \sqrt{\mathbf{c} \cdot \mathbf{c}} = \sqrt{\mathbf{c}^2}, \quad (4.8)$$

its absolute value. Alignment of two constellations  $\mathbf{c}_1$  and  $\mathbf{c}_2$  can be measured in terms of the angle  $\theta$  between them;

$$\cos \theta = \frac{\mathbf{c}_1 \cdot \mathbf{c}_2}{\pi_1 \pi_2}, \quad (4.9)$$

---

<sup>7</sup>It is mandatory to use a Greek letter if not utterly impossible. Besides that, the letter  $P$  is to be claimed shortly.



## CHAPTER 4. FORMALISATION

which can range from 0 for constellations that are the same (except for a possible difference in power) to  $\pm\pi$  for constellations that are their opposites.

### 4.2.4 The ideal-typical constellations

In the treatise of the theory some ideal-typical constellations were discussed: the regime, niches and niche-regimes. Since these are constellations they need no further introduction, suffice to suggest some notation like  $\mathbf{R}$  for the regime,  $\mathbf{n}$  for a niche, and a niche-regime can be denoted as  $\mathbf{r}$ . The landscape is not a constellation proper, but could be attributed a functioning and so can be represented similarly with a vector, let the symbol  $\mathbf{L}$  be used for it. This might be the place to explore a bit how the notation and formalism can be applied using these ideal types.

#### The regime

Depending on the reading of the theory, the regime is either the most powerful constellation in the societal system, the constellation that is representative for it or a small collection of closely related constellations that is<sup>8</sup>. This latter reading stems from the idea that a regime can always be identified in a societal system. The most powerful constellation pretty much speaks for itself:

$$\pi_R > \pi_c \quad \forall \quad \mathbf{c} \in \mathbb{S} \setminus \mathbf{R}, \quad (4.10)$$

if one wants to get formal about it<sup>9</sup>.

Being representative as a constellation for the entire system amounts to being more or less in line with it, as well as having considerable power. In other words the angle between the regime and the societal system is expected to be small. This goes well with the reading of most powerful constellation since this angle per definition is smaller if the power of the regime is greater. The regime then is not necessarily the constellation that is best aligned with the societal system, but rather that it is a consequence of its great power that it be so and it is still very well possible for other constellations to be more in line with the societal system as a whole.

#### Niches

Niches were constellations of little power, at least compared to that of the regime, with a novel or deviant functioning. It is actually quite simple to define deviancy. Taking the regime and the societal system as reference one could say that any constellation which has an angle with the functioning of the societal system *greater* than that of the regime is deviant. The angle itself could be called the deviancy. Moreover if the angle lies between that of the regime with the societal system, the constellation could be called conforming. Let the deviancy  $\delta$  of a constellation  $\mathbf{c}$  be defined by its angle with the societal system as follows:

$$\cos \delta_c \equiv \frac{\mathbf{c} \cdot \mathbb{S}}{\pi_c \Pi}. \quad (4.11)$$

Then a constellation is deviant if  $\delta_c > \delta_R$  and conforming if  $\delta_c < \delta_R$ .

<sup>8</sup>Like in the health care case study, Book III, chapter 1: Transitions in Dutch Healthcare

<sup>9</sup>Well, that is what this chapter is all about isn't it?

### Niche-regimes

Niche-regimes are, as their name suggests, in between niches and regimes. They are constellations of considerable power and viable alternatives to the incumbent regime. Niche-regimes as concepts are typically employed as the transient constellations to become the new regime during a transition. Although it usually does carry this connotation, it therefore introduces some normativity if one is to say that a niche-regime is also a deviant constellation<sup>10</sup>, apart from alternative. If a niche-regime is of considerable power and thus  $\pi_r \approx \pi_R$ , this implies also that their deviancies are close, which however does not imply that they are alike.

In any case, deviancy was easily enough defined, but ‘viable alternativivity’ appears another matter. It is apparently related to power as well as to deviancy and whereas the alternativivity appears straightforward, the term viable eludes definition. For this the concept of stress needs to be introduced, which will be done shortly.

### The landscape

The landscape is the envelope of the societal system, its environment. It is not a constellation proper and although it can be attributed functioning and in some respects treated in a similar manner it is conceptually a rather different entity. Viewing society as a patchwork of societal systems the idea comes to mind that the landscape is something like the complement of a societal system. The landscape has a broader societal scope than that of a societal system and its functioning thus will have a higher number of dimensions. That is to say, if there is a set of facets to choose from to describe a societal system, then the remaining facets would describe the landscape. If the idea of embedding or a patchwork is pursued then the landscape ideally has low or zero values for the facets that describe the societal system, whereas the functioning vector of the societal system would not even contain most of the facets the landscaped does<sup>11</sup>.

Where the landscape and the societal system *do* share facets they obviously influence each other. Regarding landscape influence it makes sense to speak of its power. A measure for landscape influence or the power of the landscape relevant to a societal system would be the power obtained by restricting the functioning vector of the landscape to the facets it shares with the societal system. A more relevant measure of landscape influence will be defined shortly in the guise of tension though.

## 4.3 Conditions

Now that the conceptual foundations are recast in this *pseudomath* the rest of the pillars themselves can be (re)formulated. Beginning with the conditions for transitional change: tension, stress and pressure. These conditions drive patterns that in turn transitionally change the societal system. Thus they can be used as indicators of how prone a system is to such change, indicators for the condition of the system, so to say. Since whatever holds for societal system

<sup>10</sup>Regardless if the transition at hand is considered desirable or not.

<sup>11</sup>One could always include them and define them to be zero, thus pre-embedding them in the higher dimensional functioning space of the landscape.

## CHAPTER 4. FORMALISATION

*mutatis mutandis* holds for societal *sub*systems, one can speak of the tension, stress and pressure of a societal system as well as of a constellation.

### 4.3.1 Tension

Tension was defined as the misalignment with the landscape. Which appeals to the idea that a societal system or constellation that is functioning well, is in sync with its environment. It is rather a negative definition though, since tension as a condition for transitional change might actually be a positive force. In any case it is precisely where the societal system is in contact with the landscape, which is where they share facets, that it can experience its influence in the form of tension. By the definition of the landscape a very compact definition of tension appears possible, letting the tension  $T$  experienced by a constellation  $c$  simply be defined by their inner product,

$$T = \mathbf{c} \cdot \mathbf{L}. \quad (4.12)$$

Thus the constellation or societal system experiences no tension if its functioning precisely ‘fills up’ that which is left open for it in the landscape and tension increases if its functioning interferes more. However, this definition would not take into account that a small constellation will feel the influence of a tension more than a larger one. This can be solved easily by normalising the above tension by dividing by the power of the constellation at hand, obtaining

$$T \equiv \frac{\mathbf{c} \cdot \mathbf{L}}{\pi_c}, \quad (4.13)$$

which still accounts for the relative influence of the each overlapping facet on the constellation, but corrects for its power as a whole. For the tension of the entire system

$$T \equiv \frac{\mathbf{S} \cdot \mathbf{L}}{\Pi} \quad (4.14)$$

directly follows.

### 4.3.2 Stress

Stress was a measure of the issues internal to a constellation or a societal system. If a constellation is functioning in such a way that it suffers from it, it can be said to experience stress. It can also refer to a constellation not living up to its own expectation or the part of society it associates itself with. This could be translated as a constellation or societal system not meeting the societal need it is supposed to address. This invites using an inner product with the societal need vector  $\bar{\mathbf{S}}$  which will be greater if the constellation or societal system is more aligned with the societal need. But, as they say, there can be too much of a good thing. Or too little for that matter. Suppose a constellation is in line with the societal need, but the some facets exceed the desired value. Think simply of overproduction for instance. The following expression seems appropriate, for the stress that a societal system could suffer<sup>12</sup>:

$$S = (\bar{\mathbf{S}} - \mathbf{S})^2. \quad (4.15)$$

---

<sup>12</sup>As well as for inviting an extremely hissing mnemonic device.

### 4.3. CONDITIONS

It measures, per facet, the difference between the societal need and the functioning that meets it and expresses this difference in terms of power squared. With this expression both a shortcoming and an excess in functioning regarding a specific facet increase the stress.

However, as was the case for the definition of tension, this definition leads to the stress suffered by a societal system to be proportional to the total power of present in that system. This, especially for such an ‘internal’ condition, appears nonsensical. Therefore the same cure is applied by scaling out the power of the societal system in the definition:

$$S \equiv \frac{(\bar{\mathbb{S}} - \mathbb{S})^2}{\Pi}. \quad (4.16)$$

If the stress of a constellation is to be calculated in a similar manner, the vector  $\bar{\mathbb{S}}$  needs to be restricted to the facets it shares with that of the constellation and the rest set to zero, or the constellation’s functioning vector needs to be expanded to include the remaining facets with the values of  $\bar{\mathbb{S}}$ . The first is readily done by introducing a projection matrix<sup>13</sup>  $\hat{\mathbf{P}}_c$ , with the expected properties ( $\hat{\mathbf{P}}_c^2 = \hat{\mathbf{P}}_c$  and such). Then the stress of a constellation becomes

$$S = (\hat{\mathbf{P}}_c \bar{\mathbb{S}} - \mathbf{c})^2, \quad (4.17)$$

which again is in need of scaling out the power of the stressed constellation, giving:

$$S \equiv \frac{(\hat{\mathbf{P}}_c \bar{\mathbb{S}} - \mathbf{c})^2}{\pi_c}. \quad (4.18)$$

#### 4.3.3 Pressure

Pressure can be seen as the effect of the presence of other constellations upon a certain one. This can be interpreted as competitive pressure in an economical sense but also issues concerning standing and status. The pressure a constellation exerts on another would obviously be felt where facets overlap, and be more felt if a constellation has more power. The inner product can be employed to measure the pressure a constellation exerts on another, like

$$P = \mathbf{c}_1 \cdot \mathbf{c}_2. \quad (4.19)$$

This would imply that the pressure felt by a constellation is not only proportional to the power of the other but also to its own, so this is divided out again to obtain

$$P \equiv \frac{\mathbf{c}_1 \cdot \mathbf{c}_2}{\pi_1}, \quad (4.20)$$

as the pressure exerted by constellation 2 on constellation 1.

The pressure exerted on a constellation by multiple other constellation would be the sum of their individually exerted pressures, leading to

$$P = \sum_{i=1}^N \frac{\mathbf{c} \cdot \mathbf{c}_i}{\pi_c}, \quad (4.21)$$

---

<sup>13</sup>Using a hat and bold face to signify a matrix or operator and distinguish it from the pressure  $P$  to be introduced after this. One always runs out of symbols.

## CHAPTER 4. FORMALISATION

for the pressure felt by this constellation through the presence of constellation 1 to  $N$ . And for the total pressure on a constellation due to all others present in the system

$$P \equiv \frac{\mathbf{c} \cdot (\mathbb{S} - \mathbf{c})}{\pi_c}, \quad (4.22)$$

is obtained.

### 4.3.4 More on power and the dimension of the conditions

The conditions that were described here are the drivers of the patterns that will be described in the next section. Patterns are the societal mechanisms that bring forth transitional change and if conditions are to drive these then apparently conditions are some sort of societal force. Although units for concepts like functioning are deliberately not specified here, it is interesting to have a closer look at the dimension some of the quantities would have.

The basic quantity clearly is functioning, upon which most, if not all, other quantities are defined. For instance deviancy which was defined as the angle between two functioning vectors, and as such itself a dimensionless quantity. The other quantity is power. Although power is defined, as a length in fact, via functioning and would share the same unit, it has a distinct conceptual status and interpretation. Even more strongly put it is more likely that functioning would be expressed in terms of power than the other way around.

The way power is defined in this formalism still leaves open whether power is a conserved quantity for the entire system or that power can be created or disappear over the course of time. If power is chosen to be a conserved quantity, this is called a zero-sum interpretation of power, like discussed in Parsons's (1963) work. There are good reasons to have power be a conserved quantity just as there are for it not being conserved and this is left for the convenience of the researcher, theoretician or modeller to decide.

Returning to the conditions of transitional change then, it is satisfying to notice that all conditions have the dimension of power. Which makes sense as power is understood as the ability, or that which gives ability, to bring about or impede societal change. In other words any sensible concept of a driver of transitional change would necessarily have the dimension of power.

### 4.3.5 Notes on notation and convention

So far every condition was defined as if felt by a constellation. It was also noted that what holds for a constellation, holds for the societal system as a whole. Or rather the other way around, but nevertheless. To avoid unnecessary confusion it is appropriate to introduce some conceptual and notational conventions here. If a condition is referred to without further specifying which constellation, then one is to assume that the societal system as a whole is meant. So tension without further ado would mean the tension as defined in equation (4.14) and stress, without reference to a constellation in specific, as (4.16). To distinguish the total or system-wide tension and stress from their respective constellation specific counterparts let the latter be notated as  $T_c$  and  $S_c$  for a constellation  $c$ .

Pressure however is a different matter, since the system as a whole does not feel pressure. If it would, it would be caused by another societal system rendering it a *tension*. However, in transition studies' colloquial speech, the

societal system is often equated with the regime. And for good reason since it is per definition the constellation that is representative for societal system, in the sense that it dominates its functioning. And it does make sense to speak of the pressure on the regime and to interpret that as a pressure on the societal system as whole. Thus let the conceptual and notational convention be adopted that the total tension  $T$  and total stress  $S$  are identified with those of the regime:  $T_R$  and  $S_R$ . This identification is a good approximation when  $\pi_R$  is close to  $\Pi$ , which it ought to be if it is a regime in its intended sense. Thus from this point on a condition without reference to a specific constellation refers to the regime:

$$\begin{aligned} T &= T_R \\ S &= S_R \\ P &= P_R \end{aligned} \quad (4.23)$$

## 4.4 Patterns

Tension, stress and pressure are the conditions that drive transitional change and transitional change comes in patterns. As mentioned in the above, when the patterns were first introduced they were referred to as affecting either constellations or entire societal systems, and which was meant would be clear from context. One of the charms of formalisation is that such ambiguity stands out directly and therefore will be got rid of here. The implicit regime convention will be used here as well so when is spoken of a certain pattern without reference to a specific constellation the regime is assumed. This allows the patterns to be defined on a ‘general’ constellation, before any further explorations.

If a pattern is something that affects a constellation with transitional change, and constellations are represented by vectors in a functioning space, then apparently patterns are functions mapping functioning vectors to functioning vectors, or  $Pattern : \mathbb{R}^n \rightarrow \mathbb{R}^n$ . Or:

$$c \mapsto \mathbf{Pattern}(c). \quad (4.24)$$

Three patterns were distinguished: reconstellation, adaptation and empowerment. Let these be denoted straightforwardly as:

Reconstellation:

$$\mathbf{Rec}(c) \quad (4.25)$$

Adaptation:

$$\mathbf{Ada}(c) \quad (4.26)$$

Empowerment:

$$\mathbf{Emp}(c) \quad (4.27)$$

Since the conditions drive the patterns they are parameters of these functions. For tension driving some reconstellation pattern this would be denoted as  $\mathbf{Rec}(c; T)$ . But it might be more convenient to use a notation where the conditions are subscripted, both for the societal system or constellation that might suffer from them, and for the patterns driven by them. Thus a societal system under pressure is

$$\mathbb{S}_P, \quad (4.28)$$

## CHAPTER 4. FORMALISATION

and when this would drive an empowerment pattern affecting constellation  $\mathbf{c}$  this would be:

$$\mathbf{Emp}_P(\mathbf{c}). \quad (4.29)$$

If a pattern is supposed to affect several constellations simultaneously, for instance when it concerns some interaction or process of merging, the notation would look something like

$$\mathbf{Emp}_S(\mathbf{c}_1, \mathbf{c}_2), \quad (4.30)$$

in the case of stress leading to an empowerment process involving constellations  $\mathbf{c}_1$  and  $\mathbf{c}_2$ , for example their merger.

But what do these patterns actually do to constellations? It would obviously go too far to pretend that an actual formulation of the functional form of a pattern could be given, this would depend on the societal system in question, the choice of facets to describe it, and many other things. A description of what a pattern does as a function ought to be possible though.

The definitions as given in table 3.2 are in terms of the origins of the processes that produce the change — from within or from outside the societal system — *and* the effect they have on a constellation. However, if a pattern is cast in the form of a mathematical function description, then the origin of the process is hardly of importance and only the effect counts.

*Nota bene:* it is important here to distinguish between the processes that produce transitional change, the mechanisms within the patterns, so to speak, and what drives them, the conditions. On the one hand, for example, a reconstellation pattern typically works through processes and mechanisms from outside the societal system, in a health care system a government measure would be a reconstellation process for as much as government is not considered a part of the health care system. On the other hand this reconstellation pattern could very well be *driven* by a condition internal to the constellation, i.e. stress. This all might appear unnecessarily complicated theoretical nit-picking, but this theory takes pride in clearly separating causes and consequences in its framework. Which amounts to separating what produces transitional change (the processes and mechanisms: *patterns*) and what causes this transitional change to be produced (the *conditions*). The only ambiguity — intentionally — retained here is that the patterns, as the processes and mechanisms *producing* transitional change, also *describe* it.

Returning to what the patterns do, one observes that both empowerment and reconstellation explicitly refer to the emergence or gaining of power of a constellation. These are obviously patterns of growth. The adaptation pattern speaks of alteration of functioning, and in the description of the processes further on in the text it becomes clear that this could involve increasing, decreasing, or maintaining the constellation's power. It should be noted that non-transitional growth or decline of a constellation, for instance by the gradual expansion of activities or decrease in demand, is not described with patterns<sup>14</sup>. Furthermore the adaptation and empowerment pattern involve the merging or clustering of constellations, or absorption if one is much more powerful than the other.

If the formalisation is to be taken any further some sacrifices of subtlety will be inevitable. The ambiguity referred to a moment ago will be the first

---

<sup>14</sup>This is a matter of timescales, since, for constellations to change in any appreciable way via non-transitional processes would take several times as much time.

victim. If a pattern is cast in some formal form, computer code, an equation or formula, then such a formula will be simply a procedure describing what happens under what circumstances. No room is left for accounting for the origin of the processes and mechanisms that produced the change that happened since the patterns now *are* these very processes and mechanisms. Consequentially, to some, formalisation will appear caricaturisation, which is actually not that far from the truth although it is no laughing matter.

Thus the picture emerges that the reconstellation pattern produces the growth of a constellation without affecting its functioning much, that the empowerment pattern increases the power of constellations with an effect on their functioning as they merge, and that the adaptation pattern alters the functioning of constellations, also through merging or absorption, while having little, or even a negative, effect on their power.

Thus, a pattern working on a constellation  $\mathbf{c}$  takes it to a state  $\mathbf{c}'$ , as in  $Pattern(\mathbf{c}) = \mathbf{c}'$  and so far this affects them in the following manner:

Reconstellation:

$$\mathbf{Rec}(\mathbf{c}) = \mathbf{c}' \quad (4.31)$$

with  $\pi_{\mathbf{c}'} > \pi_{\mathbf{c}}$  and  $\theta_{\mathbf{c}\mathbf{c}'} \approx 0$ .

Empowerment:

$$\mathbf{Emp}(\mathbf{c}_1, \mathbf{c}_2) = \mathbf{c}' \quad (4.32)$$

with  $\pi_{\mathbf{c}'} > \pi_{\mathbf{c}_1} + \pi_{\mathbf{c}_2}$  and  $\theta_{\mathbf{c}_1\mathbf{c}'} \vee \theta_{\mathbf{c}_2\mathbf{c}'} > 0$ .

Adaptation:

$$\mathbf{Ada}(\mathbf{c}_1, \mathbf{c}_2) = \mathbf{c}' \quad (4.33)$$

with  $\pi_{\mathbf{c}'} \approx \pi_{\mathbf{c}_1} + \pi_{\mathbf{c}_2}$  and  $\theta_{\mathbf{c}_1\mathbf{c}'} \vee \theta_{\mathbf{c}_2\mathbf{c}'} > 0$ .

There seems to be something problematic with the above formulations of empowerment and adaptation. Cannot a constellation adapt or get empowered without it merging with or absorbing another? This appears counterintuitive. Transitional change was considered to be radical change, but there must be some way a constellation can innovate its societal functioning without interaction with another constellation. For this the concept of a societal innovation as a constellation is employed, like introduced in the language section. Let a societal innovation be denoted with the symbol<sup>15</sup>  $\mathbf{i}$ . Then, if for example the regime is adapting under tension without absorbing another constellation the notation would be:

$$\mathbf{Ada}_T(\mathbf{R}, \mathbf{i}) = \mathbf{R}'. \quad (4.34)$$

If it is clear from context what led to the change in functioning in an empowerment or adaptation process the societal innovation can be implicit, but the use of the  $\mathbf{i}$  constellation at least allows one to specify it.

Some examples would be:

A niche, becoming empowered with a societal innovation to a niche-regime by stress on the regime:

$$\mathbf{Emp}_S(\mathbf{n}, \mathbf{i}) = \mathbf{r} \quad (4.35)$$

<sup>15</sup>In bold face to distinguish a societal innovation  $\mathbf{i}$  from the imaginary unit  $i$ .



## CHAPTER 4. FORMALISATION

The regime adapts under tension (implicit societal innovation):

$$\mathbf{Ada}_T(\mathbf{R}) = \mathbf{R}' \quad (4.36)$$

A niche gains power through reconstellation driven by the tension on a niche-regime:

$$\mathbf{Rec}_{T_r}(\mathbf{n}) = \mathbf{n}' \quad (4.37)$$

Two niches cluster to a niche-regime under pressure

$$\mathbf{Emp}_P(\mathbf{n}_1, \mathbf{n}_1) = \mathbf{r} \quad (4.38)$$

To conclude, the emergence of a new constellation around some societal innovation, under tension, stress or pressure could be denoted as follows:

$$\begin{aligned} \mathbf{Rec}_{T, S \text{ or } P}(\mathbf{i}) = \mathbf{c} \\ \text{or} \end{aligned} \quad (4.39)$$

$$\mathbf{Emp}_{T, S \text{ or } P}(\mathbf{i}) = \mathbf{c},$$

where the empowerment version is preferable only if one really wants to stress that the constellation emerges from within the system.

## 4.5 Paths

The formalism presented here can in this form already be used to describe transition paths, which will be treated with some examples in the next section. To make the step from this pseudo mathematical formal language to actual dynamical models is a different matter. The step towards a modelling implementation based on this formalism is not attempted in this book, but in the next book this is done in chapter 2. What will be done here, however, is explore what such an operational form would entail.

### 4.5.1 Describing paths

Just as in the narrative form of section 3.4.1, patterns can be concatenated to form chains of change. In fact the above formula form could serve as an even more schematic and compact way to do so and as such might serve as complement to a transition story line. As an example take the transition path as schematically described in table 3.5. It starts with a regime and two niches, so the system could be represented as  $\mathbb{S} = \{\mathbf{R}, \mathbf{n}_1, \mathbf{n}_2\}$ . There is mention of stress on the regime leading to an empowerment pattern in which the two niches cluster and become a niche-regime:  $\mathbf{Emp}_S(\mathbf{n}_1, \mathbf{n}_1) = \mathbf{r}$ . The system now consists of a regime and a niche regime  $\mathbb{S} = \{\mathbf{R}, \mathbf{r}\}$ . The stress remained and the regime now also experiences pressure from the upcoming alternative functioning of the niche-regime. The adaptation pattern induced by this further weakens the regime, to the point that the niche-regime becomes the new regime and the regime needs to assume the role of a niche-regime  $\mathbf{Ada}_{S, P}(\mathbf{R}, \mathbf{r}) = \mathbf{r}', \mathbf{R}'$ . Or,

to give this transition in one chain:

$$\begin{aligned}
& \{\mathbf{R}, \mathbf{n}_1, \mathbf{n}_2\}_S \mapsto \mathbf{Emp}_S(\mathbf{n}_1, \mathbf{n}_1) = \mathbf{r} \\
& \rightarrow \{\mathbf{R}, \mathbf{r}\}_{S,P} \mapsto \mathbf{Ada}_{S,P}(\mathbf{R}, \mathbf{r}) = \mathbf{r}', \mathbf{R}' \\
& \rightarrow \{\mathbf{r}', \mathbf{R}'\}.
\end{aligned} \tag{4.40}$$

If one prefers to tell the story of chain (4.40) from a different viewpoint — say from that of an actor associated with the regime — one might prefer not to say that the regime was weakened while adapting to the upcoming alternative, but rather that the niche-regime underwent a reconstellation, since the government issued some new legislation, and in that way the niche-regime took over:

$$\begin{aligned}
& \{\mathbf{R}, \mathbf{n}_1, \mathbf{n}_2\}_S \mapsto \mathbf{Emp}_S(\mathbf{n}_1, \mathbf{n}_1) = \mathbf{r} \\
& \rightarrow \{\mathbf{R}, \mathbf{r}\}_S \rightarrow \mathbf{Rec}_S(\mathbf{r}) = \mathbf{R}' \\
& \rightarrow \{\mathbf{r}', \mathbf{R}'\}.
\end{aligned} \tag{4.41}$$

It is now up to the transition researcher to investigate whether pressure was a relevant factor near the end of the transition and to ascertain which version is the better representation of the actual cause of events.

## 4.6 Modelling

One can imagine that the descriptions of paths like in the previous section could also be produced by a simulation model, using some operational form of the patterns running on a computer. If such a computer model was written to simulate an actual transition case its output could be compared straightforwardly to a path description in such a form produced by a transition researcher interpreting the available data on the case in traditional ways. In such a fashion the research methods could complement each other and one would gain deeper understanding than one normally would by doing a narrative analysis and making a model. As mentioned several times this is one of the aims for which this theory was constructed: computational and mathematical approaches to societal transitions.

There has been some debate (e.g. Timmermans *et al.*, 2008; Schilperoord *et al.*, 2008) whether something like a ‘transition model’ exists. Or in a similar sense if modelling transitions needs a different way of modelling that is able to capture the radical nature of transitional change. Obviously modelling aspects of transitions is readily done in regular modelling approaches, but capturing transition dynamics as such might be another matter. What the characteristics of a transition model are is difficult to ascertain. Schilperoord *et al.* (2008) argue that such models should move away from ‘equilibriumist’ approaches and be equipped with endogenous transformation processes. In this sense then pillars would be instrumental in transition models because it treats transitional change as different from other change, explicitly, by employing patterns.

If the step to a modelling attempt is taken from this formalism, the patterns need to be given some operational form, like an actual mathematical function

## CHAPTER 4. FORMALISATION

working on the functioning vectors of the constellations or a procedure in a computer program doing that. This in principle ought to not provide grave difficulties since the key concepts, like power, tension, stress, etc. now have been defined mathematically. Given a societal system as a set of constellations and their functioning vectors, the conditions for transitional change are readily calculated. And with an appropriate set of hypotheses on how the patterns respond to the conditions a running model can be produced straightforwardly, as has been done in Book III, chapter 2: Building a model on pillars.

As already aluded to in section 4.1 the formalisation here is in itself not yet a dynamical model. For this extra assumptions and hypotheses need to be introduced, for instance about the way the coupling hypothesis are translated and how the model is structured to accommodate the workings of the patterns. As an example think of the hypothesis of reconstellation being driven by tension and stress. In what proportion and order is a matter of choice and calibration, but what constellation will be reconstellated? How will the system choose? In an *ex ante* narrative analysis this is clear and when a scenario or story line is being made, a great deal of information comes from context. In a model these choices have to be made *a priori* and this means introducing more, *ad hoc*, hypotheses.

Obviously such choices are made as well when treating transitions in narratives. In a model however, at least one is forced to make explicit choices<sup>16</sup> that can be checked and in some cases even be used as parameters so they can become research questions in themselves. Apart from this, modelling introduces its own methodological difficulties as any modelling approach will come with its peculiarities. Equations might be too complicated to solve or have undesired chaotic behaviour where quasi-equilibrium was expected. Computer programs might have runaway behaviour for crucial variables, etc. Approximations will be needed and simplifications implemented to keep things tractable and to be able to interpret results. These are technical-methodological problems that need not to be addressed here.

The modelling process in itself, regardless of the model output is, just as is the case with experiments in a laboratory, of great value in the development of knowledge and understanding of the subject at hand. Maybe a typical trait of transition modelling is this explorational nature, with the model as a laboratory for the transition scientist. Not the predictions or data from the simulation are most important, but rather the rigorous thought experimenting that was required, it is in this spirit that pillar theory was constructed.

---

<sup>16</sup>Unfortunately not always to make choices explicit, though.

# Literature

- ARCTURUS 1994. Constellation. Nocturnal Art (MCD/MLP).
- AVELINO, F. 2009. Empowerment and the challenge of applying transition management to ongoing projects. *Policy Sciences*, 42(4): 369–390.
- AVELINO, F. & J. ROTMANS 2009. Power in transition: An interdisciplinary framework to study power in relation to structural change. *European Journal of Social Theory*, 12(4): 543–569.
- BAR-YAM, Y. 1997. *Dynamics of Complex Systems*. Studies in Nonlinearity. Westview Press, Boulder, Colorado.
- BEDAU, M. 2002. Downward causation and the autonomy of weak emergence. *Principia*, 6: 5–50.
- BEDAU, M. A. 1997. Weak emergence. *Nous*, 31(Supplement: Philosophical Perspectives, 11, Mind, Causation, and World): 375–399.
- VAN DEN BERGH, J. C. & S. STAGL 2003. Co-evolution of economic behaviour and institutions: Towards a theory of institutional change. *Journal of Evolutionary Economics*, 13(3).
- VON BERTALANFFY, L. 1956. General systems theory. *General Systems*, 1: 1–10.
- VAN DEN BOSCH, S. & J. ROTMANS 2008. *Deepening, Broadening and Scaling up: A Framework for Steering Transition Experiments*. 02. Knowledge Centre for Sustainable System Innovations and Transitions (KCT).
- VAN DEN BOSCH, S. & M. TAANMAN 2006. How innovation impacts society. patterns and mechanisms through which innovation projects contribute to transitions.
- BOURDIEU, P. 1977. *Outline of a Theory of Practice*. Cambridge University Press.
- VAN DER BRUGGE, R. 2009. *Transition Dynamics in Social-Ecological Systems: The Case of Dutch Water Management*. Ph.D. thesis, Dutch Research Institute for Transitions, Faculty of Social Sciences, Erasmus University Rotterdam.
- VAN DER BRUGGE, R. & J. ROTMANS 2006. Towards transition management of european water resources. *Water Resources Management*, 21(1): 249–267.

## LITERATURE

- BRUNDTLAND, G. H. 1987. Our common future. Technical report, World Commission on Environment and Development (WCED).
- BUNGE, M. 2003. *Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge*. Toronto studies in philosophy. University of Toronto Press Incorporated, Toronto.
- BYRNE, D. 1998. *Complexity Theory and the Social Sciences*. Routledge, London.
- COLEMAN, J. 1990. *Foundations of Social Theory*. The Belknap Press of Harvard University Press, Cambridge, Massachusetts and London, England.
- CROSS, M. C. & P. C. HOHENBERG 1993. Pattern formation outside of equilibrium. *Reviews of Modern Physics*, 65(3): 851–1112.
- CRUTCHFIELD, J. 1994. Is anything ever new? In *Integrative Themes*, edited by G. Cowan, D. Pines, & D. Melzner, volume XIX. Addison-Wesley, Reading, MA, Santa Fe.
- DAVIS, K. 1945. The world demographic transition. *The Annals of the American Academy of Political and Social Science*, 235: 1–11.
- DE WOLF, T. & T. HOLVOET 2005. Emergence versus self-organisation: Different concepts but promising when combined. In *Engineering Self-Organising Systems: Methodologies and Applications*, edited by S. A. Brueckner, G. D. M. Serugendo, A. Karageorgos, & R. Nagpal, Engineering Self-Organising Systems, pages 1–15. Springer-Verlag, Berlin.
- DIRVEN, J., J. ROTMANS, & A. VERKAIK 2002. Samenleving in transitie: Een vernieuwend gezichtspunt. Essay, Innovatienetwerk Agrocluster en Groene Ruimte.
- DOSI, G. 1982. Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Research Policy*, 11: 147–162.
- DUTCH MINISTRY OF ENVIRONMENT 2001. Nationaal milieubeleidsplan 4 (nmp4). nota, Ministerie van Volksgezondheid, Ruimtelijke Ordening en Milieubeheer.
- ELDREDGE, N. & S. GOULD 1972. *Punctuated Equilibrium: An Alternative to Phyletic Gradualism*. Models in paleobiology. Cooper & Co, San Francisco.
- ELTON, C. 1927. *Animal Ecology*. Sidgwick and Jackson, London.
- EMMECHE, C., S. KOPPE, & F. STJERNFELT 1997. Explaining emergence: Towards an ontology of levels. *Journal for General Philosophy of Science*, 28: 83–119.
- FRANTZESKAKI, N. & J. DE HAAN 2009. Transitions: Two steps from theory to policy. *Futures*, 41(9).
- FRANTZESKAKI, N., M. D. JONG, & P. BOTS 2007. The sour and sweet grapes of an institutional transition, impacts of institutional change on public policy design in water management sector in the netherlands.

## LITERATURE

- GEELS, F. 2005*a*. The dynamics of transitions in socio-technical systems: A multi-level analysis of the transition pathway from horse-drawn carriages to automobiles (1860-1930). *Technology Analysis & Strategic Management*, 17(4): 445–476.
- 2005*b*. *Technological Transitions and System Innovations: A Co-Evolutionary and Socio-Technical Analysis*. Edward Elgar, Cheltenham, UK; Northampton, MA, USA.
- GEELS, F. W. 2007. Analysing the breakthrough of rock ‘n’ roll (1930-1970) multi-regime interaction and reconfiguration in the multi-level perspective. *Technological Forecasting and Social Change*, 74(8): 1411–1431.
- GEELS, F. W. & R. KEMP 2000. Transitie vanuit sociotechnisch perspectief. Technical report, TUE & MERIT, commissioned by the Dutch Ministry of Environment (VROM), The Hague.
- GEELS, F. W. & J. SCHOT 2007. Typology of sociotechnical transition pathways. *Research Policy*, 36: 399–417.
- GEERTS, G. & H. HEESTERMANS, editors 1992, 1995. *Groot Woordenboek der Nederlandse Taal*. Van Dale, 12th edition.
- GEERTZ, C. 1973. *Interpretation of Cultures*. Basic Books.
- GELL-MANN, M. 1994. Complex adaptive systems. In *Systems Thinking*, edited by G. Midgley, volume IV. Sage publications.
- GERSICK, C. J. 1991. Revolutionary change theories: A multilevel exploration of the punctuated equilibrium paradigm. *Academy of Management Review*, 16(1): 10–36.
- GIDDENS, A. 1984. *The Constitution of Society*. University of California Press, Berkeley and Los Angeles.
- GILBERT, N. 2008. *Agent-Based Models*. Sage Publications, London.
- GILBERT, N. & R. CONTE 1995. *Artificial Societies: The computer simulation of social life*. UCL Press Limited, London.
- GILBERT, N. & J. DORAN 1994. *Simulating Societies*. UCL Press, London.
- GOLDSTONE, J. A. 2001. Toward a fourth generation of revolutionary theory. *Annual Review of Political Science*, 4(1): 139–187.
- GRIN, J., J. ROTMANS, & J. SCHOT 2010. *Transitions to Sustainable Development Development: New Directions in the Study of Long Term Transformative Change*. Routledge.
- GRINNELL, J. 1917. The niche-relationships of the california thrasher. *The Auk*, 34(4): 427–433.
- GROSSKURTH, J. & J. ROTMANS 2005. The scene model: Getting a grip on sustainable development in policy making. *Environment, Development and Sustainability*, 7(1): 135–151.

## LITERATURE

- DE HAAN, J. 2006. How emergence arises. *Ecological Complexity*, 3(4): 293–301.
- 2007. Pillars of change: A theoretical framework for transition models, presented at the esee 2007 conference: “integrating natural and social sciences for sustainability”.
- 2008a. The dynamics of functioning: Investigating societal transitions with partial differential equations. *Computational & Mathematical Organization Theory*, 14: 302–319.
- 2008b. Symptoms of unsustainability: Diagnosing societal systems with complexity.
- HALL, A. & R. FAGEN 1956. Definition of system. *General Systems*, 1: 18–28. Reprinted in "Systems Thinking" Sage publications, edited by Gerald Midgley, 2003.
- HEDSTRÖM, P. 2005. *Dissecting the Social*. Cambridge University Press, Cambridge, New York, Melbourne.
- HENNEMAN, P., J. DE HAAN, J. ROTMANS, & M. MINNESMA 2006. Zeeland op een kantelpunt: Moet, kan en wil zeeland veranderen? Technical report, DRIFT, commissioned by: Zeeuwse Milieu Federatie/Hogeschool Zeeland.
- HOLLAND, J. 1995. *Hidden Order: How Adaptation Builds Complexity*. Ulam Lectures Series. Helix books / Perseus books, Cambridge, Massachusetts.
- HOULT, T. F. 1969. *Dictionary of Modern Sociology*. Littlefield, Adams & Co, Totowa, New Jersey.
- KAUFMANN, S. A. 1993. *The Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, New York, Oxford.
- KEMP, R. 1994. Technology and the transition to environmental sustainability. the problem of technological regime shifts. *Futures*, 26(10): 1023–1046.
- KEMP, R. & D. LOORBACH 2006. *Transition Management: A Reflexive Governance Approach*, chapter 5, pages 103–131. Edgar Elgar.
- KEMP, R., D. LOORBACH, & J. ROTMANS 2007. Transition management as a model for managing processes of co-evolution towards sustainable development. *The International Journal of Sustainable Development and World Ecology*, 14: 1–15.
- KRASNER, S. 1982. Structural causes and regime consequences: regimes as intervening variables. *International Organization*, 36(2): 185–205.
- KRASNER, S. D. 1983. *International Regimes*. Cornell University Press, New York.
- KROEBER, A. L. & C. KLUCKHOHN 1952. *Culture: A Critical Review of Concepts and Definitions*. Peabody Museum of Archaeology and Ethnology, Harvard University, Cambridge, Massachusetts.
- KUHN, T. S. 1962, 3rd edition. *The structure of scientific revolutions*. University of Chicago Press, London, Chicago, second edition edition.

## LITERATURE

- LANE, J. 1997. *Public sector reform, Rationale, trends and problems*. Sage publications, London.
- LEVINTHAL, D. A. 1998. The slow pace of rapid technological change: Gradualism and punctuation in technological change. *Industrial and Corporate Change*, 7: 217–247.
- LOORBACH, D. 2007. *Transition Management: New Mode of Governance for Sustainable Development*. International Books, Utrecht.
- LUHMANN, N. 1982. *The Differentiation of Society*. Columbia University Press, New York.
- 1984. *Social Systems*. Stanford University Press, Stanford, California.
- McKELVEY, B. 1999. Complexity theory in organization science: Seizing the promise or becoming a fad? *Emergence*, 1(1): 5.
- MERTON, R. K. 1968. *Social Theory and Social Structure*. The Free Press, New York.
- MOSSBERGER, K. & G. STOKER 2001. The evolution of urban regime theory: The challenge of conceptualization. *Urban Affairs Review*, 36(6): 810–834.
- NELSON, R. R. & S. G. WINTER 1977. In search of a useful theory of innovation. *Research Policy*, 6: 36–76.
- NICOLIS, G. & I. PRIGOGINE 1989. *Exploring Complexity*. W.H. Freeman and Company, New York.
- NORTH, D. 1990. *Institutions, institutional change and economic performance*. Cambridge University Press, Cambridge.
- OVIN, R. 2001. The nature of institutional change in transition. *Post-communist economies*, 13(2).
- PARSONS, T. 1963. On the concept of political power. *Proceedings of the American Philosophical Society*, 107(3): 232–262.
- PATTEN, B. C. & G. T. AUBLE 1980. Systems approach to the concept of niche. *Synthese*, 43: 155–181.
- PETTIGREW, A. M. 1997. What is a processual analysis. *Scandinavian Journal of Management*, 13(4): 337–348.
- POLTI, G. 1868. *The Thirty-Six Dramatic Situations (Translation: Lucile Ray)*. James Knapp Reeve, Franklin Ohio.
- VAN RAAK, R. 2010, forthcoming. Re-defining health systems from the perspective of transition theory. In *Transitions in health systems*, edited by J. Broers & J. Bunders.
- RAVEN, R. P. & G. P. VERBONG 2007. Multi-regime interactions in the dutch energy sector: The case of combined heat and power in the netherlands 1970–2000. *Technology Analysis & Strategic Management*, 19(4): 491–507.



## LITERATURE

- RIP, A. & R. KEMP 1998. *Technological Change*, volume Volume 2, pages 327–399. Battelle Press, Columbus, Ohio.
- RITTEL, H. W. J. & M. M. WEBBER 1973. Dilemmas in a general theory of planning. *Policy Sciences*, 4: 155–169.
- ROGERS, E. M. 2003. *Diffusion of Innovations*. Free Press, New York, 5th edition.
- ROTMANS, J. 1998. Methods for ia: The challenges and opportunities ahead. *Environmental Modeling & Assessment*, 3(3): 155–179.
- 2005. *Societal Innovation: between dream and reality lies complexity*. Inaugural Addresses Research in Management Series. ERIM, Erasmus Research Institute of Management, Rotterdam.
- ROTMANS, J., R. KEMP, & M. VAN ASSELT 2001. More evolution than revolution: Transition management in public policy. *Foresight*, 3(1): 17.
- ROTMANS, J., R. KEMP, M. VAN ASSELT, F. GEELS, G. VERBONG, & K. MOLENDIJK 2000. Transities & transitie management: De casus van een emissiearme energievoorziening. Technical report, ICIS & MERIT, commissioned by the Dutch Ministry of Environment (VROM), The Hague.
- ROTMANS, J. & D. LOORBACH 2006. Transition management: reflexive governance of societal complexity through searching, learning and experimenting. In *Reflexive Governance for Sustainable Development*, edited by J. Voss, D. Bauknecht, & R. Kemp, page 53. Edward Elgar, Cheltenham.
- 2009. Complexity and transition management. *Journal of Industrial Ecology*, Vol. 13, Special Issue on Complexity and Industrial Ecology(2): 184–196.
- ROTMANS, J. & B. DE VRIES 1997. *Perspectives on Global Change: The TARGETS Approach*. Cambridge University Press.
- SCHILPEROORD, M., J. ROTMANS, & N. BERGMAN 2008. Modelling societal transitions with agent transformation. *Computational & Mathematical Organization Theory*, 14(4): 283–301.
- SCHOT, J. & F. W. GEELS 2007. Niches in evolutionary theories of technical change: A critical survey of the literature. *Journal of Evolutionary Economics*, 17(5): 605–622.
- SENGE, P. M. 1990. *The Fifth Discipline: The Art & Practice of The Learning Organization*. Random House, London.
- SHALIZI, C. R. 2001. *Causal Architecture, Complexity and Self-Organization in Time Series and Cellular Automata*. Ph.D. thesis, University of Wisconsin-Madison.
- SMITH, A., A. STIRLING, & F. BERKHOUT 2005. The governance of sustainable socio-technical transitions. *Research Policy*, 34(10): 1491–1510.
- SOANES, C. & A. STEVENSON, editors 2004, 2006. *Concise Oxford English Dictionary*. Oxford University Press, 11th edition.

## LITERATURE

- STACEY, R. 1996. *Complexity and creativity in organizations*. Betterr-Koehler Publishers, San Francisco.
- STROGATZ, S. H. 1994. *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry and Engineering*. Studies in Nonlinearity. Perseus, Cambridge, MA.
- SUMMERS, D. & M. RUNDELL, editors 1987. *Longman Dictionary of Contemporary English, New Edition*. Longman Group UK Limited.
- TAANMAN, M. forthcoming in 2010. *Transition Monitoring*. Ph.D. thesis, Dutch Research Institute for Transitions, Erasmus University Rotterdam.
- TIMMERMANS, J. & J. DE HAAN 2008. Special issue on computational and mathematical approaches to societal transitions. *Computational & Mathematical Organization Theory*, 14(4): 263–265.
- TIMMERMANS, J., J. DE HAAN, & F. SQUAZZONI 2008. Computational and mathematical approaches to societal transitions. *Computational & Mathematical Organization Theory*, 14(4): 391–414.
- TURNER, J. H. 2003. *The Structure of Sociological Theory*. Wadsworth/Thomson Learning, Belmont, seventh edition.
- TUSHMAN, M. L. & P. ANDERSON 1986. Technological discontinuities and organizational environments. *Administrative Science Quarterly*, 31: 439–465.
- UN 1997. Global change and sustainable development: critical trends. Technical report, United Nations, Economic and Social Council, Commission on Sustainable Development.
- VAN DE VEN, A. H. 1992. Suggestions for studying strategy process: A research note. *Strategic Management Journal*, 13(Special Issue: Strategy Process: Managing Corporate Self-Renewal): 169–191.
- VAN DE VEN, A. H. & M. S. POOLE 2005. Alternative approaches for studying organizational change. *Organization Studies*, 26(9): 1377–1404.
- VIJVER, M. 2005. *Protein Politics*. Ph.d., Twente University. ISBN: 90-365-2181-5.
- WEAVER, P. M. & J. ROTMANS 2006. Integrated sustainability assessment: what is it, why do it and how? *International Journal of Innovation and Sustainable Development*, 1(4): 284–303.

*LITERATURE*

# Glossary

## **Adaptation**

One of the three patterns of transitional change, describing a constellation altering its functioning either through interacting or merging with other constellations within or from outside the societal system. *Internally induced constellation change.*

## **Broadening**

Conducting similar transition experiments in various societal contexts. *Adapted from van den Bosch & Rotmans (2008).*

## **Conditions**

More accurately, *conditions for transitional change*, the misalignments, dysfunctions and impediments in a societal system that make it prone to transitional change. The first pillar of transitional change.

## **Constellation**

Societal subsystem with a part of the functioning of the societal system it is part of, meeting parts or aspects of the societal need associated with the whole societal system.

## **Coupling hypotheses**

Set of hypotheses coupling the conditions and patterns pillars. Causal assumptions about which conditions drive which patterns.

## **Cultures**

The appraisal and legitimacy of the functioning of the societal system; *how it is perceived*. Either part of the duality structures and cultures, or the triplet structures, cultures, and practices, used to describe constellations and their functioning.

## **Deepening**

Learning in and about a transition experiment in a specific context, with the aim of gaining understanding of its societal functioning. *Adapted from van den Bosch & Rotmans (2008).*

## Glossary

### **Empowerment**

One of the three patterns of transitional change, describing a new constellation emerging, or an existing one gaining power, either on itself or through interacting or merging with other constellations *within* the societal system. *Bottom-up constellation change.*

### **Functioning**

The way a societal system meets a societal need.

### **Landscape**

The environment of a societal system.

### **Niche**

Constellation with, in comparison to the regime, less power and a novel or deviant functioning.

### **Niche-regime**

Viable, competitive alternative to the regime with appreciable power, though less powerful than the regime.

### **Paths**

More accurately, *transition paths*, the trace of transitional change over time. Paths are concatenations of patterns, sequences describing how the societal system evolves during the course of a transition. The third, and last, pillar of transitional change.

### **Patterns**

More accurately, *patterns of transitional change*, the ways transitional change takes place in a societal system. Patterns describe how constellations are affected in transitional change. Patterns are formed when conditions are present. Societal transitions are sequences of patterns. The second pillar of transitional change.

### **Persistent problems**

Complex problems deeply rooted in the underlying structure of the systems that suffer from them.

### **Pillars**

More accurately, *pillars of transitional change*, division of the theory in three semi-independent parts, dealing exclusively with the system state, change, or its possible evolutions.

### **Power**

Of a constellation: the proportion of the total functioning of the societal system that can be attributed to it.

**Practices**

The ensemble of production routines, behavior, ways of handling and implementation at the individual level, including self-reflection and reflexive dialogue. *Taken from Grin et al. (2010)*. Part of the triplet structures, cultures, and practices used to describe constellations and their functioning.

**Pressure**

Condition for transitional change that occurs when the functioning of a constellation, generally the regime, is being compromised by the functioning of one or several other constellations.

**Reconstellation**

One of the three patterns of transitional change, describing a new constellation emerging, or an existing one gaining power by influences from *outside* the societal system. *Top-down constellation change*.

**Regime**

Most powerful constellation in the societal system and typical of it.

**Scaling-up**

Process in which a constellation, typically a niche, gains power.

**Societal innovation**

Powerless constellation with a novel way to meet a societal need; bare societal functioning.

**Societal system**

A part of society to which can be attributed a functioning.

**Societal transition (I)**

The radical changing of the functioning of a societal system from one state of apparent societal equilibrium to another.

**Societal transition (II)**

Radical change of regime.

**Stress**

Condition for transitional change arising when the different facets of the functioning of a constellation, generally the regime, do not combine well.

**Structures**

The infrastructures and institutions related to the functioning of the societal system; *how it works*. Either part of the duality structures and cultures, or the triplet structures, cultures, and practices, used to describe constellations and their functioning.

*Glossary*

**Tension**

Condition for transitional change having its origin in a compromised relation between a constellation, generally the regime, and the landscape.

**Transition arena**

Innovation network formed by frontrunners, strategic thinkers from relevant different backgrounds addressing a transition as a societal challenge and outlining paths to meet it. *Adapted from Loorbach (2007).*

**Transition experiment**

Innovation project with a societal challenge as a starting point for learning aimed at contributing to a transition. *Taken from van den Bosch & Rotmans (2008).*

# Index

Adaptation, 56  
Broadening, 30  
Conditions, 47  
Constellation, 27  
Coupling hypotheses, 62  
Cultures, 40  
Deepening, 30  
Empowerment, 57  
Functioning, 26  
Landscape, 27  
Niche, 29  
Niche-regime, 29  
Paths, 68  
Patterns, 56  
Persistent problems, 52  
Pillars, 38  
Power, 27  
Practices, 40  
Pressure, 47  
Reconstellation, 57  
Regime, 28  
Scaling-up, 29  
Societal innovation, 30  
Societal system, 26  
Societal transition, 29  
Stress, 47  
Structures, 40  
Tension, 47  
Transition arena, 29  
Transition experiment, 30





Book III  
Buildings



# Table of Contents

|      |   |           |
|------|---|-----------|
| 1    | “Transitions in Dutch Health Care” . . . . .                    | <b>7</b>  |
|      | <b>BUILDING NARRATIVES — A HISTORICAL CASE STUDY . . . . .</b>  | <b>7</b>  |
| 1.1  | Introduction . . . . .  | 8         |
| 1.2  | Methodology: applied pillar theory . . . . .                    | 9         |
| 1.3  | Dutch health care as a societal system . . . . .                | 11        |
| 1.4  | The landscape . . . . .   | 14        |
| 1.5  | Transition overview and questions raised . . . . .              | 15        |
| 1.6  | Patterns in health care . . . . .                               | 17        |
| 1.7  | A closer look: addressing the questions . . . . .               | 38        |
| 1.8  | A closer look on the big picture . . . . .                      | 43        |
| 1.9  | Conclusion and discussion: ironies and paradoxes . . . . .      | 44        |
| 1.10 | Appendix: the pattern matrix . . . . .                          | 47        |
|      | References . . . . .  | 54        |
| 2    | “Building a Model on Pillars” . . . . .                         | <b>57</b> |
|      | <b>BUILDING MODELS — RIGOROUS THOUGHT EXPERIMENTS . . . . .</b> | <b>57</b> |
| 2.1  | Introduction . . . . .  | 58        |
| 2.2  | Outline of the model . . . . .                                  | 58        |
| 2.3  | Model behaviour . . . . .                                       | 62        |
| 2.4  | Conclusions and discussion . . . . .                            | 65        |
| 2.5  | Plots and graphs . . . . .                                      | 67        |
|      | References . . . . .  | 72        |
| 3    | “The Dynamics of Functioning” . . . . .                         | <b>75</b> |
| 3.1  | Introduction . . . . .  | 76        |
| 3.2  | Formalism . . . . .   | 77        |
| 3.3  | Elaborations . . . . .  | 83        |
| 3.4  | Conclusions and discussion . . . . .                            | 88        |
|      | References . . . . .  | 91        |



# About these buildings

This book is built on the pillars from the preceding book. It is a bundle of elaborations of the theory that was presented in book II. Each elaboration stands on itself and can be read independently of the other chapters of this book.

The bundle of chapters that make up this book, however, can also be considered a proof of point. Because the theory presented in the preceding book did not come into being the ‘ordinary’ way, that is, via inductive cycles and the scientific method, its validity and use needs to be shown in a somewhat extraordinary manner as well. The theory is a construct, something built from components to serve a certain goal, and rather than try to prove or argue that it is *right* this book will try to convince the reader that it is built *consistently* and that it *works*.

One of the claims the theory put forth was that it would be able to describe and explain societal transitions in the broadest sense. That it should be able to serve as well in a narrative case-study, as in computational and mathematical approaches. Therefore this book features three elaborations, one building narratives and two building models.

The narrative elaboration is a historical case study of transitions in Dutch healthcare. In it, the patterns are used to reconstruct the historical story lines and understand the transitions the Dutch healthcare system went through and under what conditions this occurred.

The second building is a computational elaboration. The entire theoretical structure has been cast in the form of a computer model as directly as reasonably possible. Since the model is a more or less direct representation of the theory it allows rigorous thought experimenting with the theory itself and to study the consequences of certain theoretical presuppositions.

The third building is a mathematical elaboration. Some of the central transition concepts such as constellations and functioning are modelled using partial differential equations and studied both analytically and with numerical simulations for some simple stylised examples. These examples focus on the interactions between the regime and niches and the possibilities for the latter to scale-up.

*TABLE OF CONTENTS*

## Building 1

# Narrative historical case study: “Transitions in Dutch Health Care”<sup>1</sup>

This chapter presents a historical case study of the developments in Dutch health care, from the early 19<sup>th</sup> to the end of the 20<sup>th</sup> century. The transition perspective was taken under the assumption that it was appropriate to describe and explain dynamics over such a long time-span on a large scale. In that time-span the Dutch health care system developed from almost nothing to a complex, privately and publicly organised system. The sources used for this case study were secondary, and sometimes primary, historical literature. This data was interpreted in the framework of the theory developed in Book II of this thesis. The Dutch health care system is described as a societal system in terms of constellations and their structures, cultures and practices. The evolution of this system and the transitions it went through have been described using the patterns and conditions as put forward in the theory developed in this thesis. Using this approach it was possible to formulate and address a number of research questions concerning the Dutch health care system. Apart from this an important aim of this chapter is to present a narrative elaboration of the theory which led to a step-by-step plan to apply the theoretical framework to a historical case.

---

<sup>1</sup>This chapter is based on the case study work the author has done with Roel van Raak.



## 1.1 Introduction

This chapter builds from the theory as developed in chapter 3 of Book II. A narrative historical case study is presented of Dutch health care and the transitions it went through over — roughly — the last two hundred years. It is an example of how transition paths can be constructed and reconstructed using patterns and conditions, like suggested in chapter 3.4.1 of book II.

In the last two centuries health care in the Netherlands evolved from a very limited, often charity based system providing mainly care, to a complex, privately and publicly organised system focussing on cure. The history of Dutch health care is a rich research subject and various in depth studies of parts of it can be found, for instance Houwaart's (1991) standard work on the hygienist movement or Juch's (1997) detailed account on the rise of the medical specialists. A study of the developments leading to the current Dutch health care system as a whole seems to be absent, however. Such an overarching study and analysis obviously runs the risk of paying for broadness with shallowness or losing overview in striving for completeness.

However, the idea that Dutch health care is about to go through major transitions is voiced by many. This alone would be an indication that the time is right for an investigation of how this system took shape and which transitions it has already gone through. A number of other reasons why such an integral study is relevant can be given as well, for example:

How did it come about that health care is so much about the opposite: the curing of illnesses? What is the reason for the apparent subordinate position of care to cure? Why do integral approaches appear to be giving way to specialised ones time and again?

The Dutch proverb that it is better to prevent than to cure<sup>2</sup> would lead one to think that prevention should have an important part in any health care system. Historically in the Dutch case it had, what happened to it?

Can patterns be identified in the development of Dutch health care that might give insight in current developments? Do certain patterns typically occur under certain conditions? Is there any logic in the historical development? For example how to account for the apparent 'stress' between the imperative to cure and limited financial resources or how to grasp what is at the root of the 'tension' in wanting to be able to care for an ever growing group of elderly people. Or how to understand why more abstract notions related to health care concerning well-being, quality of life and the human factor are difficult matters in the current system?

For all of these reasons the transitions in the Dutch health care system are the object of study of this chapter, more specifically the period ranging from roughly the beginning of the nineteenth century until the end of the twentieth, for reasons exposed later on. An interest in the transitions this system went through, and not just the historical development as such, invites the perspective of transition studies, viewing Dutch health care as a societal system.

Transitions in Dutch health care are an interesting case, albeit a very demanding one. Since the historical development of a health care system is more than just the history of medicine or medical inventions, the approach needs to be broader than just historical or innovation focussed. And since it concerns

---

<sup>2</sup>*Beter voorkomen dan genezen.*

## 1.2. METHODOLOGY: APPLIED PILLAR THEORY

changes on the level of the practices of the individual patient, the level of medical culture, as well as the level of government structures, a truly multi-level or rather a scalable conceptual framework is a necessity. Bearing in mind the questions raised earlier, some additional demands on the theory are made. The theoretical approach needs to be able to untangle the various intertwined historical story lines into patterns that can be interpreted. Furthermore, it must be equipped with concepts to address in a sensible way what conditions are that cause transitional change to occur.

Whether the theory will offer all of the above is a research question in itself and one that also will be addressed. Applying these fresh conceptual tools to the demanding case at hand will be a tough test and it is fair to say that this chapter serves a twofold cause, first understanding the transitions in Dutch health care and second the evaluation of the theory used.

Something more on the structure of this chapter. After this introduction the reader will be presented with the methodology, which is as much the outcome as it was the approach to this case study. After giving the general idea it will be cast in the form of a step-by-step plan. After that, Dutch health care and its development will be described as a societal system. First its demarcations and how it is composed of certain constellations, followed by a description of these constellations and their historical developments. This provides a bird's-eye view of the overall transition and the larger developments therein which allows a reframing of the central questions to be answered, both in a methodological as well as a case-related sense.

Once the questions are clear the overall transition, the grand scheme, is presented as a complex of patterns. Chains of these patterns form the transition paths that underly this overall transition; they are like threads in the larger story line of the transition. Subsequently some specific patterns will be elaborated in more depth and detail to address the questions raised. The outcome of this will not so much be straightforward conclusions but rather a discussion and evaluation of the historical, methodological and theoretical sides of this case study.

## 1.2 Methodology: applied pillar theory

Given the theory presented in book II of this thesis, the matter at hand is to now apply this to the health care system in the Netherlands. The overarching research question for this chapter is what the transitions were that the Dutch health care system went through in the period ranging roughly from the end of the 19<sup>th</sup> to the end of the 20<sup>th</sup> century. With the theoretical background in mind this means that we must first identify which constellations make up the societal system of Dutch health care as it was functioning at the end of the 20<sup>th</sup> century. This also presupposes that a demarcation be made of what this societal system is. This will be done by answering the question: "how is the societal need for health care met?", which leads to an idea of what the functioning of the system is, although it would change over time. Once a clearer picture is established of the societal system and its boundaries we commence to chart what influences from outside the system are to be taken into account, that is, we sketch the landscape.

Next comes the larger part of the empirical research for this chapter which is

## BUILDING 1. “TRANSITIONS IN DUTCH HEALTH CARE”

to describe the current composition of the societal system, or more exactly, the situation at the end of the twentieth century. And subsequently the historical developments that led to it. This will be done by describing each constellation in terms of their structures, cultures and practices and a historical narrative describing how they came to what they are. These time lines will then be combined and give rise to an analysis in which we will identify the patterns of transitional change that describe how these constellations emerged, merged, split, changed or disappeared altogether and the conditions under which this occurred. From this analysis we attempt to draw some conclusions at the system level, by which is meant that we try to see if for instance phases in the overall development can be distinguished.

This can be summarised in the following ‘step-by-step’ plan to a historical transition pattern analysis with some guiding questions for every step:

### 1. SYSTEM COMPOSITION

#### (a) Demarcating functioning and time-span

How is the societal need met by the system; what is its functioning? What will be explicitly taken as functional parts of the societal system? What is the temporal demarcation, the considered time span?

#### (b) Identifying constellations

What are the constellations that make up the composition of the societal system at the end of the considered time line? What is their functioning in terms of structures, cultures and practices?

#### (c) Describing the landscape

What landscape influences are to be taken into account. What is not part of the system but provides boundary conditions for its development? What are national or global developments that have influenced its history? What influences from other systems are to be considered?

### ***Phrasing questions* — What does one want explained?**

### 2. SYSTEM EVOLUTION (UNRAVELLING)

#### (a) Constellation and landscape developments

How and when did these constellations emerge, how did they change, split, merge, or disappear to come to the current situation? What is the rough development of the landscape in which all this occurred? Also, what constellations were present in the beginning of the considered time span that disappeared although being influential or noteworthy in the development of the current system?

#### (b) Constellations then — constellations now

What was the functioning of the relevant constellations at the beginning and at the end, and maybe in between, of the considered time-span? What were their structures, cultures and practices?

#### (c) Identifying patterns and conditions

What were the conditions and patterns of transitional change behind these developments?

### 1.3. DUTCH HEALTH CARE AS A SOCIETAL SYSTEM

***Rephrasing questions* — What is that which one wants explained in terms of patterns and conditions?**

#### 3. SYSTEM ANALYSIS (RECONSTRUCTING)

##### (a) Tracing patterns and conditions

What is the overall transition in terms of the transition paths of the individual constellations and the sequence of patterns that underly it? How did these paths influence each other; where there patterns that influenced multiple constellations?

##### (b) Addressing questions

What was the chain of patterns that led to or describes the phenomenon to be explained? Which patterns influenced the development of multiple constellations? What were the relevant conditions that drove these patterns?

The above plan has been used as a guideline to structure the remainder of this chapter. Some additional notes on temporal demarcation in general perhaps: One of the criticisms of transition analyses is: why let the transition ‘start’ here and ‘end’ there? And is not a transition in this time-span unimportant in the larger time-span? Temporal demarcation is to a certain extent arbitrary, one can look as far back as one wants or as available data permits. In any case the multi-pattern approach advocated here allows one to analyse any sufficiently long period from a transition perspective, regardless if an overall transition path is recognisable or if phases can be distinguished.

## 1.3 Dutch health care as a societal system

### 1.3.1 Demarcations

This roughly corresponds to the first two steps in the step-by-step plan and a bit of step three. The system we will be studying is the Dutch health care system from halfway the 19<sup>th</sup> to the end of the 20<sup>th</sup> century: roughly one and half century of Dutch health care. The demarcation of what is Dutch is regarded as self-explanatory. Health care is somewhat more difficult to grasp as a system. For this analysis we use the working definition of a health care system as a system meeting the needs of the inhabitants of the Netherlands concerning their physical and mental well being in the broadest sense, without going as far as taking into account their happiness, self development and education or fulfilment of life’s desires. In slightly simpler terms it is the system concerned with delivering health care services in its various forms. The temporal demarcation is chosen to begin in the first decades of the 19<sup>th</sup> century. If one insists on more exact numbers the years 1815 to 1818 could be appropriate starting points, the birth of the Dutch state and the date the first health care law became effective. It is the time following the defeat of Napoleon ending the French occupation that amongst other things dismantled the guild system. It is also the time that a health care system was still in a nascent form and medical science was at the verge of many breakthrough insights and innovations. The end of the period is chosen at the end of the twentieth century since at that time a somewhat stable

## BUILDING 1. “TRANSITIONS IN DUTCH HEALTH CARE”

health care regime began to see large scale, government initiated reforms which are still too fresh to appreciate with appropriate distance.

### 1.3.2 Constellations

The health care system is by no means a monolith and its functioning is distributed over various societal subsystems — *constellations* — each associated with typical health care services. Although sometimes certain actors or professions (e.g. the pharmacist) or institutions (e.g. hospitals) are typical or even iconic for the constellations they are associated with, the societal subsystem is always what is meant, even if at times the actor or institution is used to refer to the constellation, which will prove to be almost inevitable. The point of view is that of the constellation, in other words a systemic view, actors are taken into account, though often implicitly.

Since the societal system of Dutch health care has grown more complex over time<sup>3</sup> it appears sensible to first consider the composition of the health care system at the height of its complexity. Thus beginning at the end, eight constellations were identified that together compose the societal system of Dutch health care near the end of the twentieth century. The constellations around:

#### General practice

*Functioning:*

The providing of non-specific health care to all categories of patients for all complaints in principle. Gate-keeping for the health care system and specialist treatment in specific and prescribing of medicine.

*Associated actors and institutions:*

The general practitioner, his or her practice.

#### Mental health

*Functioning:*

Health care related to the mind.

*Associated actors and institutions:*

Psychiatrists, mental institutions.

#### Pharmacy

*Functioning:*

Relating to the preparation, production and selling of drugs.

*Associated actors and institutions:*

The pharmacists, the pharmacy and the pharmaceutical industry.

#### Specialist health care

*Functioning:*

The providing of specialised health care, either specialised for specific illnesses, bodily functions, types of patients or with relation to certain technology.

---

<sup>3</sup>This can be safely said in this case, be it a general phenomenon for complex societal systems or not.

### 1.3. DUTCH HEALTH CARE AS A SOCIETAL SYSTEM

*Associated actors and institutions:*

Physicians, surgeons.

#### **Hospital based care**

*Functioning:*

Related to care requiring patients to stay under close supervision of medical professionals, e.g. surgery or intensive care and monitoring, or care requiring specific facilities or equipment.

*Associated actors and institutions:*

The hospital, hospital staff

#### **Nursing**

*Functioning:*

Concerned with providing (bedside) care in diverse forms. Also but not specifically the care for the disabled, retarded, elderly or other forms of long-term care.

*Associated actors and institutions:*

Nurses, geriatric institutions, asylums, home care institutions.

#### **Financing**

*Functioning:*

Related to all financial aspects of health care ranging from public funding to insuring and wage compensation.

*Associated actors and institutions:*

Insurance companies, sick funds, the state and the legislative force.

#### **Public health**

*Functioning:*

Health care provision at scales larger than that of the individual or the family.

*Associated actors and institutions:*

Municipal public health services (GGD's<sup>4</sup>), municipalities, cross organisations, the state.

The regime currently appears to be characterised by the dominance of the combination of the constellations of specialists, hospitals, the related academic training and the financial constellation of insuring with which it is thoroughly interwoven.

Obviously this is not a comprehensive picture of everything related to health care. Some constellations one readily thinks of are not present, like alternative medicine, occupational health care, etc. The reasons for this are diverse but what is generally considered mainstream health care is intended to be treated as or as part of one of the above constellations. The state or the government is not present as a constellation itself but as an associated actor for several

---

<sup>4</sup> *Gemeentelijke GezondheidsDienst* in Dutch, sometimes *Gewestelijke* or *Gemeenschappelijke GezondheidsDienst*, meaning regional or communal health service. Other local variants exist as well.

## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

constellations in different roles, for instance as legislator, financier or enforcer. Patients are not represented in a constellation either, ironically enough. Societal systems are *meeting* societal needs and in that sense patients are — although shaping, defining, and using it — on the outside of the health care system.

Also absent in this list are constellations that were prominent for a certain period of time in Dutch health care but have disappeared or have been subsumed by other ones. For example, the constellation around care for the poor, which essentially was the dominant form of institutionalised health care until a couple of centuries ago, that has been subsumed almost altogether in the modern welfare state. These constellations will be either treated as they come by in the historical developments of other constellations (cross organisational care) or as aspects of other constellations (care for the poor).

In the following the health care system and its constellations are described in more detail beginning with that which is of influence on it without being part of it: its environment, the landscape.

### 1.4 The landscape

The 'envelope' of the societal system is the landscape, its embedding in the developments of society and the other societal systems that influence it. In the case of Dutch health care this entails a host of things and to discuss their influences on the functioning of the health care societal system it is convenient to describe the landscape in terms of structures, cultures and practices also. This appears straightforward for structures and cultures; think for instance of institutional structures that influence, but are not restricted to health care like legislation or financing arrangements, or think of cultural aspects like the political climate. Speaking of the practices on the landscape is somewhat paradoxical, the high level of aggregation and abstraction appear unfamiliar to the rules and routines that are at the basis of what is meant with practices here. However, there are certain practices that are much broader than health care — hygiene, for example — and therefore their influence needs to be taken into account as such. Also the aspects of the practice of medicine in a broader context than just the Dutch system is taken into account here, like international developments in (medical) science. Events, like an epidemic or war will be described in terms of how they change the structures, cultures and practices of the landscape and the societal system.

Properties and influences of the landscape will be treated roughly as follows:

#### Structures

Broad societal trends be they economical, political, demographical or otherwise, like centralisation and growing government influence or liberalisation and privatisation. The Dutch legislative system in so far as it relates to health care consequently also the political system and affairs of the Dutch nation, for instance when it concerns the welfare state.

#### Cultures

Perspectives and world views. Developments like secularisation and individualisation. The attitude of and towards medicine and health care and the influence of religion thereupon.

## 1.5. TRANSITION OVERVIEW AND QUESTIONS RAISED

### Practices

Habits and routines throughout society. The way people work and live.  
Also international developments in scientific and medical practice.

## 1.5 Transition overview and questions raised

### 1.5.1 Bird's-eye view

It was mentioned earlier already that the current health care regime appears to consist of the constellations around the specialists, the hospital and financing. This is not to say that all of current health care in the Netherlands is a part of that regime, but rather that it is typical of it. Moreover, the way health care is organised, perceived and experienced is dominated by the structures, cultures, and practices of these three constellations. Other constellations take over facets of this way of functioning when one would at first glance reckon that it would be unsuitable, like the implementation of the complaint-diagnose-treatment practice in long term caring situations.

Typical for regimes, like discussed in book II, is a strong degree of self-referential and self-reinforcing dynamics, by which is meant that the evolution of the regime under non-transitional conditions, so to say, would strengthen the current functioning. In other words, if for whatever reason the regime needs to change it will prefer to do so by seeking recourse to the structure of specialisation, in a culture that regards health care as curing illness via the by now well known triplet as practice.

Why is this such a strong regime and how did it emerge historically? To what extent was this development a historic inevitability and in any case what patterns and conditions led to its formation?

The time frame of this case study begins in the early 19<sup>th</sup> century, roughly when the modern Dutch state came into being<sup>5</sup>. Curative treatment back then was virtually non-existent, the medical profession was to a large extent a craftsmanship and medical science still in its infancy. The guild system had been abolished by the French during the occupation to not be restored when they left. What is the overall picture of the development of the Dutch health care system to what it is, or at least was at the end of the 20<sup>th</sup> century?

With the developments of the constellations, as sketched in the previous, in mind, it appears that the system went through at least three phases. Hardly coincidental, this is consistent with the phases that are often distinguished in societal transitions in transition studies. Obviously these phases do not neatly cut the line of time in three but smoothly flow over into each other.

First there was a period of pre-development lasting roughly until 1870. In this period the medical profession in the Netherlands took shape. Laws became effective that defined and protected the medical professions and arranged the authorization of the practice of medicine. Health care decoupled from general care, from incarceration and from temporary lodging. Professional associations emerged for physicians, pharmacists, etc. Medical science advanced in various ways and more treatments became possible for patients. Institutions for the

---

<sup>5</sup>The Netherlands became an independent kingdom in 1815 of which Belgium separated itself in 1830. In the year 1848 a new constitution became effective reforming the Dutch state to a parliamentary democracy, strongly limiting the power of the monarch.



## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

financing of health care emerged: sick funds, laws for care of the poor. Arrangements for public health, hygiene and such were advocated and instated by the hygienist movements and municipalities facing the burdens of industrialisation, urbanisation and epidemic outbursts.

A second phase, one of acceleration, lasts approximately until the 1930's. Several constellations saw their transition paths go through accelerated dynamics in that time, though not necessarily at the same time. After the 1868 Thorbecke-laws and the educational reforms a couple of years later, the end of the 19<sup>th</sup> century sees the emergence of medical specialties as a result of ongoing developments in medical science. Advances that also made it possible for hospitals to become safer places where more and more people left better than they entered. Nursing became a real profession through the pioneering work of Fliedler in Deaconess hospitals and the iconic Nightingale, but maybe more so through the proliferation of the cross organisations throughout the Netherlands that also assumed the dominant role in public health. In this turmoil of new developments health care financing remains an unresolved matter, with various sick funds and private insurance although in 1930 a law becomes effective insuring income during periods of sickness of workers.

Curiously enough, a phase of stabilisation is initiated by WWII, when the German occupier instates the sick fund act. This arranges mandatory coverage of various treatments, including specialised treatment in hospitals. It entails a compulsory sick fund membership for those with an income under a certain level. After WWI already, but certainly after WWII the pharmaceutical industry became a force to be reckoned with. Although some attempts at large scale reforms towards a national health service are taken, the system crystallises out around this core, growing to immense proportions in terms of public spending on health care. The AWBZ<sup>6</sup>-law in 1964 consolidates this system incorporating long term and mental health care (a medical specialty since long, but rather *sui generis*) in the regime through direct funding of their associated institutions.

### 1.5.2 Questions

Any historical overview raises more questions than it answers, here are some that arise in this case:

- The first question was already raised in the above: How could this current health care regime emerge and why does it appear to be so strong?
- Stronger put: why is there even a tendency for other constellations to align their structures, cultures, and practices with that of the regime?
- Why did the Dutch system, in spite of various attempts, never adopt a national health service, or any nationally organised system of public health for that matter?
- Related to this is the question why never a powerful constellation around prevention has formed? Considering the influential role of the two waves of the hygienist movement, for example, this appears counter-intuitive.

---

<sup>6</sup>The general law on special healthcare costs, or *Algemene Wet Bijzondere Ziektekosten*, in Dutch.

## 1.6. PATTERNS IN HEALTH CARE

- Similarly, one wonders why more integral forms of care, like offered by the cross organisations, did not make it in the long run?
- Why have the professions of the surgeon and the physician become so integrated?
- Why has nursing gained and then lost so much of its societal standing?
- Considering the historical development of the health care systems, its patterns and the conditions under which they unfolded, can one understand the current conditions of the health care system? For instance, the tension between the 'curing-imperative' and the growing health care expenditure.

In the following, the historical threads leading to this picture of an overall transition as sketched above will be unravelled in terms of patterns and conditions of transitional change. This 'grand scheme' will be used to take a closer look to some developments that might shed some light on the questions raised here.

## 1.6 Patterns in health care

### 1.6.1 The grand scheme I: The landscape

Although the influence the landscape has had will be described in terms of patterns working on the specific constellations it is nevertheless appropriate to sketch the landscape in broad terms over the course of the transition. As said this will be done in terms of structures, cultures and practices. The time span will be divided into four not entirely arbitrary eras, such that the beginning, end and 'turning points' of the phases described in the bird's-eye view fall in these blocks. The descriptions in these blocks are deliberately broad as details, specific influences, and dates are to appear in the descriptions of the constellations. The years are an indication only.

#### 1815-1848 Pre-development

##### *Structures:*

Dutch state becomes a fact and legislation becomes effective defining the medical professions. Degree of urbanisation is high (60–70%) but stable. Declining industry. Protestant Christian faith dominates institutions.

##### *Cultures:*

Aftermath of enlightenment: individual rights become recognised. Distinction made between the sick, the insane and the criminal.

*Practices:* No relevant practices identifiable.

#### 1848-1918 Taking off to acceleration

##### *Structures:*

The Netherlands become a parliamentary democracy. More specific legislation is made for health care, reforms in educational system. Urbanisation concentrates in the large cities. Dutch society becomes segregated to denomination ('pillarisation'), Catholics and socialists obtain their respective pillars.

## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

### *Cultures:*

Socialism becomes an important idea, reframing how society is stratified, fear of revolution influences politics. 'Reveille' of protestant welfare work, women's emancipation, emergence of civil society. Idea of disease as something to be cured arises. Hygiene as a concept emerges, perhaps under influence of recurring epidemic outbreaks.

### *Practices:*

Urban life changes needs and demands for health care. New medical instruments invented, hygiene and antisepsis become part of medical practice.

### **1918-1945 Touching down from acceleration**

#### *Structures:*

New era of industrial production after WWI, chemical industry emerges. Legislation insuring workers' payments during sickness becomes effective. Later the German occupier arranges health care costs insurance.

#### *Cultures:*

Perception is widespread that disease is absence of health, something to be cured.

#### *Practices:*

Medical practice becomes the practice of curing.

### **1945-1990 Stabilisation**

#### *Structures:*

Economic boom after WWII. First Labour governments. Welfare state emerges arranging many aspects of health care. Large role for government in public life. Height of denominationally segregated society in 1960's followed by ongoing de-pillarisation.

#### *Cultures:*

Secularisation and individualisation. Criticism about health care system arises: beginning of the contested health care system.

#### *Practices:*

Evidence based medicine, medicalisation in general.

### **1.6.2 The grand scheme II: Constellations and patterns**

In this section the development of each constellation is described as a short story line draped on the skeleton provided by the patterns, after which the changes in structures, cultures and practices are summarised. One constellation is described only with patterns since it was a transient, emerging halfway and being absorbed by others before the end of the timespan of this case study (the cross organisation constellation).

## 1.6. PATTERNS IN HEALTH CARE

### General practice

Although everyone has an idea of what a general practitioner is and does, defining the profession is not that simple as Bremer (2006) already remarked. In any case the 19<sup>th</sup> century knew already of doctors with their own practice, often at their homes, sometimes with an apothecary, doing house calls, all of which are very much defining characteristics. The legislation at the beginning of that century demanded that these people graduated at one of the universities<sup>7</sup> or at one of the local committees<sup>8</sup>. Later, in the law of 1865 an academic training became mandatory and the local committees were no longer relevant. A basic degree in medicine<sup>9</sup> was sufficient to become a general practitioner.

#### *1865 / Reconstellation (Stress)*

The Thorbecke-laws define and protect the title ‘physician’. Merging the professions of the surgeon and the academical doctor. These laws also limit authorisation to practice medicine, and thus general practice, to those having passed a state exam.

Once the specialists gain influence in the beginning of the 20<sup>th</sup> century the general practitioners begin feeling insecure in their position.

#### *1912-1930 / Adaptation (Stress)*

Internal struggles in the medical association lead to positioning of specialists with respect to general practitioners and steps toward a system of referral.

Within the medical society (NMG<sup>10</sup>) general practitioners still outnumber the specialists by a factor of two in the 1930’s, but in terms of status and perceived representation they are on the decline. Therefore in 1925 they founded the Committee for the Study of Medical Professional Interests<sup>11</sup> which became the Dutch Community for General Practitioners (NVH<sup>12</sup>). This committee at first was not valued greatly, but later, around 1938, 1939 almost one third of the general practitioners were members and the committee contributed a lot to the emancipation of the profession and even to the definition of what it was to be a general practitioner.

#### *1925-1939 / Empowerment (Stress)*

Though still stronger in numbers, general practitioners perceive their position weakened and unite within the association (NMG). This culminates in the founding of the Dutch Association of General Practitioners (later to become the NVH). This association was highly influential in the emancipation and definition of the profession.

---

<sup>7</sup>Back then referred to as *hoge scholen*, nowadays as *universiteiten*. *Hoge scholen* still exist in the Netherlands with a different status, somewhat lower than that of the *universiteiten*.

<sup>8</sup>A system, instated by the Dutch king, of provincial committees mandated with the licensing of non-academically trained medical practitioners.

<sup>9</sup>*Basisarts*, in Dutch.

<sup>10</sup>*Nederlandse Maatschappij tot bevordering der Geneeskunst*, in Dutch. Nowadays a royal society, the *Koninklijke Nederlandse Maatschappij tot bevordering der Geneeskunst*, the KNMG.

<sup>11</sup>In Dutch, the *Comité voor de Studie van Medische Beroepsbelangen*.

<sup>12</sup>In Dutch, the *Nederlandse Vereniging van Huisartsen*.

## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

In 1973 the law required a specialisation to general practice at one of the universities and from that point on the general practitioner was registered as such in a national register.

On the structural side of things it is important to realise that the original clientele of the general practitioner consisted mainly of the reasonably well to do and privately insured who were visited at their homes. This means that up to the end of the 19<sup>th</sup> century the publicly insured (sick funds) patients, especially in urban areas, sought treatment with specialists in their polyclinics for which either the municipalities paid through care for the poor arrangements or by the sick funds. When sick funds and their patient base grew this led to a desire for the general practitioner to secure their financial position. One of the ways in which this eventually was done was through a system of referral, which made the general practitioner a gatekeeper for specialist care, adding the sick fund patients to their clientele. The system of referral first emerged in a formal way in the medical association (NMG) in 1912 as a proposal that did not obtain approval formally (Juch, 1997) but was nevertheless adopted by the local departments. From then on it found its way through legislation via the German sick fund act in 1941 to later become anchored in the law AWBZ in 1964 (Schouwstra, 1995)

### *1930-1964 / Reconstellation (Stress)*

Changes in educational system, the German sick fund act (1941) and later the law AWBZ further defined and consolidated the position of the specialist. This is relevant for general practice as its relation to specialist treatment is finally codified. The general practitioner becomes a gatekeeper of the medical system.

The general practitioner re-enters university from a different angle, as a sort of specialist with the specialty of generalism. First professors are appointed for chairs of general practice, later an educational trajectory for general practice appears. In 1973 the law required a specialisation to general practice at one of the universities and from that point on the general practitioner was registered as such in a national register.

### *1968-1973 / Adaptation (Stress)*

General practice becomes the medical specialty of generalism. First in academia, later formalised by law and registration.

| <b>From</b>  |            | <b>To</b>   |
|--|------------|---|
| The private practice, sometimes with a pharmacy. Trained academically or in a clinical school, authorisation through regional councils. Paid via honorarium. | Structures | The private practice, sometimes with a pharmacy, sometimes with several practitioners, mandatory academical training, gatekeeper for specialist treatment and administration of prescription drugs. Paid via insurance. |

Continued on next page...

## 1.6. PATTERNS IN HEALTH CARE

| From   |           | To   |
|--|-----------|--|
| Community figure of high social esteem. Offering more than medical care and advice only to patients and families alike. Cure, care and community based care. Generalist with human interest. | Cultures  | Advising, friend-like relation with patient and often the family. Somewhere between cure, care and community based care. Generalist with human interest.   |
| House calls, visiting hours, family care, for all illnesses (in principle), not restricted to a certain category of patients. General treatment, including even surgery.                     | Practices | House calls, visiting hours, family care, for all illnesses (in principle), not restricted to a certain category of patients, directly helping for 'small' complaints, referral to specialist care and prescription of medicine. |

Table 1.1: Development of general practice constellation

### Mental health

Initially, that is up to the French revolution, mental health care as such was non-existent and restricted to locking the insane up in madhouses<sup>13</sup> or prisons if they were a burden to society. The French revolution changed the perception, and the idea gradually became accepted that an insane person was mentally ill and therefore ought to be treated as a patient (Dankers & van der Linden, 1996). In the Dutch system this led to a series of legislation that led to a system of mental health care organized around the asylum<sup>14</sup>. First there was the royal decree in 1818 that the purpose of an asylum is to cure the mentally ill (Schut, 1970). In 1871 a law becomes effective that says that the insane are not to be locked up in penitentiary institutions and the society for care for the mentally ill<sup>15</sup> is willing to take up all the imprisoned mentally ill in 1883 (Schut, 1970). Roughly speaking one sees that, although the practice remains that of institutionalisation, the insane gained recognition as patients, as mentally ill rather than e.g. possessed. This occurs under the influence of the landscape developments like enlightenment and growth of urban populations.

#### *1818-1883 / Reconstellation (Tension)*

Enlightenment and French revolution inspire laws formally distinguishing between the sick, the mentally ill, vagabonds, and criminals. Incarceration of the mentally ill in penitential institutes is prohibited.

In the second half of the nineteenth century the somatic approach became prevalent in mental institutions (Vijselaar, 1982). This was also a consequence of overpopulation of the asylums and the successes of natural sciences. This development changed the culture, as said, towards more a natural science attitude, while the structure remained that of the asylum<sup>16</sup>.

#### *1850-1900 / Adaptation (Stress)*

Asylums become more like hospitals under the 'somatic approach': being

<sup>13</sup>Back then referred to as *dolhuysen* in the Dutch language.

<sup>14</sup>In Dutch this was referred to as *gesticht*.

<sup>15</sup>The *maatschappij tot verpleging van krankzinnigen*.

<sup>16</sup>Though often renamed into 'psychiatric institution', *psychiatrische inrichting*.

## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

treated similarly as having a physical illness, in the sense of treatment in bed. Under the stress of overcrowded asylums and natural scientific successes.

The more traditional verbal relation with the mental patient re-emerged later in the beginning of the 20<sup>th</sup> century under the influence of the work of Freud and Breur, a process which is sometimes referred to as re-psychologisation (Dankers & van der Linden, 1996) This is part of the continuation of the adaptation pattern of this constellation, like the development at the turn of the twentieth century when the practices started resembling more and more that of the normal hospital (patients in beds) a constellation with which it coevolved in that period.

From the nineteen twenties onward mental health care goes through patterns of adaptation and empowerment leading to a 'medicalisation'. This is the period that new mental illnesses are diagnosed (notably schizophrenia) and medication leads to a depopulation of the asylums (Dankers & van der Linden, 1996). The joint rise of this form of clinical psychology and social psychology leads to the process of what is called deinstitutionalization. From the nineteen fifties the world health organisation starts to recognise mental disorders as diseases (actually 1949 with the sixth edition of the ICD<sup>17</sup>) and in 1952 sees the first edition of the DSM<sup>18</sup> by the American Psychological Association. These developments led to a more hospital-like approach to mental health care, with patients admitted for short periods.

### *1920-1950 / Adaptation (Tension)*

Medicalisation and de-institutionalisation. New mental illnesses recognised (like schizophrenia). Patients treated more and more like in a hospital in the sense of brief stays.

Medicalisation of mental health care gains a new dimension when the effectiveness of CPZ<sup>19</sup> opens up a new world of psycho-pharmaceuticals. Mental illness now became treatable with medicine, whereas drugs before mainly served for sedation of mental patients.

### *1952-1975 / Adaptation (Tension)*

The discovery of the psychotherapeutic effectiveness of CPZ brings pharmaceuticals into mental health care.

Simultaneously after the second world war societal well being became important (Dankers & van der Linden, 1996) which led to several (public) institutions for mental health care like the RIAGG's<sup>20</sup> and arrangements under the law AWBZ.

### *1964 / Reconstellation (Stress)*

The AWBZ-law arranges financing for, amongst others, mental health care, making it no longer a municipal matter. Mental health institutions now directly funded by the Dutch government.

---

<sup>17</sup>International Classification of Diseases.

<sup>18</sup>Diagnostic and Statistical Manual of Mental Disorders.

<sup>19</sup>Chlorpromazine, the oldest typical antipsychotic, sold as chlorpromazine hydrochloride under the name Thorazine in the US and Largactil in Europe.

<sup>20</sup>In Dutch, the *Regionale Instelling voor Ambulante Geestelijke Gezondheidszorg*, regional institution for ambulant mental health care.

## 1.6. PATTERNS IN HEALTH CARE

| From  |            | To  |
|---|------------|---|
| The ward with resident psychiatrists and nurses. Often financed as care for the poor. Also private psychotherapeutic counselling. | Structures | Although still present, currently focussed on care outside clinics through counselling and medication. Specialists (psychologists, psychotherapists and clinical psychologists) and their institutions are either affiliated in the professional association (called GGZ <sup>21</sup> ) or work in a private practice setting. |
| Strong authority of the psychiatrist and nurses. Disparate views on insanity and mental illness.                                  | Cultures   | Strong authority of the psychiatrist, scientific attitude, mental problems viewed as illnesses, focus often on helping the patient returning to normality or coping with disorder, rather than curing.  |
| Admittance and detaining of patients, institutionalisation. Somatic approach (patient treated in bed, as in hospitals).           | Practices  | Also that of complaint, diagnose and treatment, but because of the nature of the specialism more in dialogue with the patient and often on a more long term basis.  |

Table 1.2: Development of mental health constellation

### Pharmacy

The current constellation around Dutch pharmacy has two distinct but strongly tied faces. The one being that of the dispensing chemist with its culture of craftsmanship in which that of the apothecary lives on, an academically trained entrepreneur usually with his own store. In the Dutch system there is a difference between a general drug store that usually also sells products for personal hygiene etc., and stores in which also medication on prescription is being sold<sup>22</sup>. This gives a rather stable structure of commercially operating and academically trained people administering medicine, combined with a culture of a combined commercial and scientific attitude, practiced in designated selling points<sup>23</sup>. Although once its core business, the modern chemist usually does not prepare the drugs it sells, this is the realm of the other side of the constellation, the pharmaceutical industry. The pharmaceutical industry is strongly international, but the Dutch system has its own share of large industry players as well, either as strictly pharmaceutical or via the general chemical industry (Organon, AKZO Nobel, Brocades, etc.).

The preparation, distribution and administering of medicine is what could be called pharmacy. Historically this has long been connected with the apothecary. Indeed, as early as in the 13th century the holy Roman emperor Frederik II formally ordained that the preparation of medicinal substances is a task of the apothecary (Verhoog, 1998). In later times, the early years of the kingdom

<sup>21</sup>Short for *Geestelijke GezondheidsZorg*, mental health care.

<sup>22</sup>The former would be a *drogist* and the latter is an *apotheeek*. The words are akin to the English words drug store and apothecary, as one recognises easily.

<sup>23</sup>Such a selling point would be an *apotheeek*, as noted.



## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

of the Netherlands, the separation between pharmacy as a profession and the rest of the practicing of medicine was arranged by law (in 1804 and 1818). These laws ensured that the practicing of medicine and the profession of the pharmacist could only be done by authorised professionals, whose authorisation was to be given by regional committees. Under this law a doctor could not sell medicine and a pharmacist was not allowed to prescribe medicine (except for instance in some rural areas without an apothecary). In these days the education of a pharmacist could either be like under the guild system, from a master or in so called clinical schools that emerged in the early twenties of the 19<sup>th</sup> century. The new law in 1865 took the responsibility for authorising the practicing of pharmacy from the regional committees and gave it to the state. A decade or so later, once the teaching of pharmacology was restricted to universities (1876), the profession of pharmacist definitively became an academic profession, with a degree demanded for practicing it 1878. The state-committee summoned by royal decree to prepare the law that became effective in 1865 was the immediate reason for the founding of the Dutch association of pharmacists NMP<sup>24</sup>, that addressed the government with advice to promote the interests of their profession (all this: Bierman, 1988) and safeguarded the quality of the profession. As far as the profession and practicing of pharmacy is concerned the structures laid by these laws and the refinements in the law of 1963 remained fairly stable.

### *1841-1865 / Empowerment (Stress)*

Pharmacists unite in the Dutch association for pharmacy (NMP, later KNMP) when the state committee to investigate new health care legislation is summoned. Their lobby secured the position of the pharmacist as a medical profession in the 1865 laws.

As said, the production of medicine was historically the task of the pharmacist. The 19<sup>th</sup> century however saw the rising of a pharmaceutical industry. This was for a large part a side effect of the more general trend of mass production and the emergence of a chemical industry. In the Dutch system, the sudden expansion of the reach of newspapers because of a tax measure on them being loosened<sup>25</sup> in 1869, and the brand-law<sup>26</sup> becoming effective in 1880, made large scale advertising possible which created a mass market for pharmaceuticals (Huisman, 1999). These, however, were products that would qualify mostly as quackery, which was the incentive for the founding of an association against quackery<sup>27</sup> in 1880. The end of the 19<sup>th</sup> century also saw the emergence of a so-called 'ethical' pharmaceutical industry, which valued scientific methodology and proven effectiveness of the medicines produced. In the Netherlands examples of these were Brocades, which advertised with 'apothecary quality' and later Organon. Especially the latter is noteworthy since at first the community of pharmacists viewed the upcoming pharmaceutical industry as a threat, taking over their traditional role of preparing medicines. The ethical companies, however, employed pharmacists and had an active advisory and research

<sup>24</sup>The *Nederlandse Maatschappij tot bevordering der Pharmacie*. Later they acquired the royal adjective, becoming the *Koninklijke Nederlandse Maatschappij tot bevordering der Pharmacie*, KNMP.

<sup>25</sup>The disappearance of the so-called *dagbladzegel* a mark on daily periodicals.

<sup>26</sup>The *merkenwet*.

<sup>27</sup>In Dutch, *Vereniging tegen de Kwakzalverij*.

## 1.6. PATTERNS IN HEALTH CARE

role for them (Huisman, 1999; Verhoog, 1998). The First World War pushed the chemical industry to new heights and the discovery of antiseptics and antibiotics created new possibilities and markets.

### *1914-1930 / Adaptation (Tension)*

Isolation in WWI and post war boom in chemical industry make large scale industrial production of pharmaceuticals possible.

After the Second World War the pharmaceutical industry could grow again through the new field of pharmaceuticals following the discovery of the effectiveness of CPZ (Caldwell, 1970). Up till then psychopharmaceutics were mainly sedatives. The main point is that pharmaceutical industry grew tremendously and with some large jumps from halfway through the 19<sup>th</sup> century up till now. But this growth was also gradual in the sense that it was in pace with growing demands and growing possibilities after the Second World War, with pharmaceuticals entering and changing the face of mental health for good, thus connecting the mental health constellation with that of pharmacy.

### *1952-1975 / Adaptation (Tension)*

The discovery of the psychotherapeutic effectiveness of CPZ brings pharmaceuticals into mental health care.

| <b>From</b>   |            | <b>To</b>  |
|---|------------|--|
| Industrial and small-scale production of drugs, sold in pharmacies by pharmacists. Preparing and selling of drugs separated from practicing medicine. | Structures | Distribution, via pharmacies, of prescribed drugs that are industrially produced. Preparing and selling of drugs separated from practicing medicine. |
| Proud traditional craftsmanship and large scale 'ethical' industry.   | Cultures   | Scientific attitude, 'evidence-based' combined with tradition and medical entrepreneurship.  |
| Production and preparation of drugs in industry and (limited) pharmacy. Distribution via pharmacy and rural general practitioners.                    | Practices  | Production in pharmaceutical industry, distribution via prescription in pharmacies and rural general practitioners.                                  |

Table 1.3: Development of pharmacy constellation

### Specialist health care

The rise of the specialists took place at a paradoxical time for the Dutch health care system. As said, the advent of new medical technology (like the invention of the eye-mirror in 1851) and the progress of medical science in general (the use of anaesthetics and a- and antisepsis in the second half of the 19<sup>th</sup> century for example) cleared the way for a new generation of medical professionals. The proliferation of the medical profession also led to tremendous diversification and there was fear in the sector that this would lead to a devaluation of the medical community as a whole. The way the medical community attempted to face this was to unite — a popular societal dynamic at the time — in the Dutch medical

## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

association (NMG) in 1849.

### *1850-1930 / Adaptation (Tension)*

Progress of medical science saw growth of the number of specialists and their becoming a group within the medical community.

In those years a state committee was working on new legislation for Dutch health care and the NMG put in a lot of effort into seeing to it that this legislation united the medical community into what was called unity of standing<sup>28</sup>, which meant that all physicians were educated at the same or similar level, and would have passed a state-supervised exam to warrant that, and all these physicians would be authorised to practice all that was part of the medical profession. This legislation, taking into account a lot of the wishes of the NMG, became effective in 1865<sup>29</sup> which, combined with a reformation of the educational system, led to the desired unity. This latter reformation allowed people with a non-classical education to become *doctores*, tearing down an ancient boundary between the surgeons and the physicians and leading to a large inflow of new physicians (four out of five potential physicians stemming from this flow in 1895 already). The irony is that these developments also pushed the emergence of the specialties that would effectively end this unity in due time<sup>30</sup>.

### *1865-1878 / Reconstellation (Stress)*

Necessity to organise health care and higher education leads to Thorbecke and education laws. More recognition for specialists and more educational possibilities to become one (HBS<sup>31</sup>).

The late 19<sup>th</sup> century, then, saw this new generation and new type of medical professional, the specialist. Adding the word 'specialist' to the doctors name was a way to gain clientele and to mark their distinction with general practitioners. The specialist would boast a more scientific attitude and approach to medicine. Although the patients and the working environment of the specialists differed from the general practitioners the latter felt threatened in their livelihood. The specialist would treat mostly poor or publicly insured patients in hospitals or in polyclinics (which they often privately ran). This approach also allowed them to gain experience and do research. The general practitioner in contrast held a private practice with a visiting hour and preferred to do house calls on well-to-do patients that were privately insured or otherwise able to pay the doctors fee.

In many senses this summarises the different cultures of the specialist and the general practitioner. The general practitioner being the 19<sup>th</sup> century doctor that was also a friend of the family with an advising role in affairs that went well beyond the medical, a public figure in the neighbourhood with a steady, easy going, way of going about his medical business. Whereas the specialist would be the hard working 20<sup>th</sup> century physician, a trained professional with a scientific passion for his work.

---

<sup>28</sup>In Dutch this was referred to as *eenheid van stand*.

<sup>29</sup>The four Thorbecke-laws, as they are commonly referred to.

<sup>30</sup>This paragraph mostly (van Lieburg, 1999) all the rest (Juch, 1997).

<sup>31</sup>*Hogere BurgerSchool*, comparable in level of education to the *gymnasium*, back then and nowadays the highest level in secondary education. The HBS only did not provide education in the classical languages.

## 1.6. PATTERNS IN HEALTH CARE

### *1890-1920 / Empowerment (Pressure)*

Specialists become a group to be reckoned with in the medical community through their work in their polyclinics and their subsequent penetration of the hospitals.

The developments that subsequently led to the dominant position of the specialists and the structuring influence they have had on the Dutch health care system appears to contain three main threads:

The first thread is emancipation through organisation and registration. At the turn of the 19<sup>th</sup> century specialists start organising themselves around their specialties, mostly out of scientific interests. Later in 1910 they unite for more non-scientific reasons — an imminent, but apparently never to arrive national health care act — in the general Dutch association of medical specialists<sup>32</sup> and later also within the NMG in different forms from 1915 on. Some internal wars for territory were fought between the general practitioners and the specialists. Within the association (NMG) the general practitioners still by far outnumber the specialists and they appeared to have their own share of privately insured patients, which they preferably treated in their homes and if necessary referred to a professor for clinical treatment in an academic hospital. The specialists, operating from and in their polyclinics on the publicly insured<sup>33</sup> or poor<sup>34</sup>, however, grew in number as did their group of patients. Since more patients became insured by the sick funds, as they can be called, these funds, became better organised and grew and consequently the general practitioners began to see the specialists as their competitors. No surprise that in the early 20<sup>th</sup> century the general practitioners began to advocate a system of referral where they would act as gatekeepers for specialist treatment. Since this was contrary to the dogma of ‘free choice of physician’ which the association (NMG) held high, this was blocked up till as late as 1912 (Juch, 1997). From that point on, however, local departments of the association took it upon themselves to initiate a system using referral notes for specialist care, something that was met with mixed reactions by the specialists but marched on to become part of the model contract with sick funds in 1927. In the 1920’s the organisation of specialists within the association started to investigate the necessary conditions for a physician to call him- or herself a specialist. This was to lead to the committee for the registration of specialists<sup>35</sup>) in 1931. Although this was a matter of interest within the profession and of private-law, the health insurers quickly adopted it and started to only pay to registered specialists (Schouwstra, 1995).

### *1912-1930 / Adaptation (Stress)*

Internal struggles in the medical association lead to positioning of specialists with respect to general practitioners and steps to a system of referral.

This can be seen as the start of the second thread: the recognition of the medical specialties by health insurers and legislation. Since in the Dutch situation the universities and their education *followed*, rather than initiated the specialisations that took place in the practical medical world, the way the health

<sup>32</sup>The, ANVVM, the *Algemene Nederlandse Vereniging voor Medisch Specialisten*.

<sup>33</sup>The so-called *ziekenfondspatiënten*.

<sup>34</sup>Via an arrangement called *armenzorg*, care-for-the-poor.

<sup>35</sup>SRG, the *Specialisten Registratie Commissie*.

## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

insurers utilised the registration completed the *de facto* recognition of the specialists as a medical constellation. The *de jure* recognition was on its way too, however. In an early draft of the elusive national health insurance act in 1920 a distinction was already made between specialist treatment and general practice and the National Insurance Bank<sup>36</sup> officially distinguished between specialists and general practitioners in 1922. A national health care act<sup>37</sup> was eventually forced upon the Dutch health care system by the German occupier in 1941 which finally formally distinguished the general practitioner and the specialist, including a gatekeeper role for the former with a system of referral. This system still prevails although the law has been replaced and altered (e.g. with the introduction of the AWBZ in 1964) a couple of times.

### 1930-1964 / Reconstellation (Stress)

Changes in educational system, the German sick fund act (1941) and later the law AWBZ further defined and consolidated the position of the specialist.

The third thread was coevolution with the hospital-constellation. The 1920's saw the decline of the privately ran polyclinics. The costly new equipment (x-ray, ECG, etc.) combined with low incomes often forced specialists to seek refuge in hospitals. Hospitals in those days were becoming the modern institutions we are all more or less familiar with so that these constellations would find each other was to be expected. At first hospitals opened polyclinics and specialists could work there and make use of the hospital facilities. This worked to the higher esteem of the specialists and the hospitals. In the hospital the specialist eventually found a natural habitat and from there on the division in specialities became a leading principle in the rest of Dutch health care.

| From   |            | To  |
|--|------------|---|
| Structures around the specialties and the required academic training and skills.   | Structures | The medical specialties themselves, embodied in academic training and the institutions where the specialties are being exercised (the hospitals). |
| Hard working and passion for the trade, natural scientific attitude, focus on curing, strong authority for the physician or surgeon. | Cultures   | Hard working and passion for the trade, natural scientific attitude, focus on curing, strong authority for the physician or surgeon.              |
| Treatment in polyclinics. Mostly directly approached by patients.  | Practices  | Treatment in hospitals. Usually via referral by a general practitioner.   |

Table 1.4: Development of specialist health care constellation

### Hospital based care

The hospital is in many senses representative for the medical system as a whole (Houwaart, 1996). In it, treatment of patients at the highest level, scientific

<sup>36</sup>The *Rijks VerzekeringsBank* (RVB), nowadays the *Sociale Verzekeringsbank* (SVB).

<sup>37</sup>The *Ziekenfondsbesluit*.

## 1.6. PATTERNS IN HEALTH CARE

knowledge production and the education of people in the medical professions come together, interact and stimulate each other. In a sense the hospital evolved from the guesthouses<sup>38</sup>, the history of which goes back to early medieval times. At first the main function of these establishments was one of care or even simply offering a place of temporary lodging, hospitality as one could say. It is, however, the 19<sup>th</sup> century which saw the emergence in the Netherlands of the modern hospital. A first step towards this was separating the function of temporary lodging from that of providing health care.

### *1818-1854 / Reconstellation (Stress)*

Legislation comes into effect that makes hospital a place for the sick (no longer as a place for temporary lodging for instance) and municipalities become responsible for the health care of their poor.

After a period in which the care in guesthouses and hospitals was at its very worst, roughly from 1650 to 1850 (van der Heyden, 1994), several developments were to lead to the emergence of the modern hospital. The low quality of the care was itself a driver in these developments as it led to the founding of for instance deaconess hospitals and hospitals as private enterprises<sup>39</sup>) that competed with the care offered at the municipality financed and organised guesthouses.

### *1848-1920 / Adaptation (Tension)*

A new spirit (*r veil*) and the new constitution banning repression of the Catholic community leads to the founding of many new hospitals (Deaconess, Catholic and private).

Roughly around the same time in the early 19<sup>th</sup> century the doctors were starting to become dissatisfied with the medical education given. They complained of a lack of a practical, empirical component in their education for which some students journeyed to the famous clinic in Paris (Mooij, 1999). These doctors took to use the guesthouses as clinics for their clinical education. More or less parallel to this was the situation concerning the clinical schools that educated the non-academic medical professionals (surgeons, pharmacists, etc.). When in the twenties of the 19<sup>th</sup> century the clinical schools emerged (per royal decree 1823), they would be doing the practical part of their education in the guesthouses.

Other developments that made the modern hospital possible were of a medical scientific nature. From 1847 on anaesthetics made complicated operations possible and the 1860's to the 1870's saw the introduction of a- and antiseptics that allowed patients to survive these operations.

### *1860-1900 / Adaptation (Tension/Stress)*

Anaesthetics, hygiene, a- and antiseptics transform the hospital into a place of non-lethal surgery.

Since the circumstances in the guesthouses were still very harsh, the second half of the 19<sup>th</sup> century saw the founding of new hospitals with a confessional or non-religious signature. This was both made possible by and made the development possible of a growing, both in number and status, of the nursing

<sup>38</sup>In Dutch referred to with the similar *gasthuizen*.

<sup>39</sup>*Diakonessenhuizen* and *particuliere ziekenhuizen*.

*BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"*

profession — a development that also was a phenomenon of women’s emancipation of those days. Treatment in these new hospitals was, however, costly and not accessible for the lower classes. In 1854 a new law for the care of the poor came into effect<sup>40</sup> which made the municipalities responsible for the health care of the underclass (van der Heyden, 1994). Add to this several outbursts of cholera and the necessity for modernisation of the guesthouses (under municipal responsibility) led to the collection of modern hospitals that is characteristic of the current medical system.

The more complicated equipment that became a necessity for operations and other hospital treatments made it the new environment of choice for doctors, as opposed to the house of the patient which it used to be. More specifically it became the environment of choice for medical specialists. Hospitals needed their skills and knowledge and the specialists the facilities only hospitals could afford.

*1920-1941 / Adaptation (Pressure)*

The specialists become more dominant in the hospital and the treatments available there. Specialists bring prestige and knowledge in the hospitals and they profit from the facilities.

Still hospital care was very expensive and often not covered by sick funds or health insurance. The German sick fund act arranged this by force, making hospital care generally accessible, paving the way for the very central role the hospital has in the current medical system, defining the public image of health care to a large extent.

*1941 / Reconstellation (Stress)*

The German sick fund act makes hospitalisation more accessible for sick fund patients. It also decouples to a large extent the insurance fees from the covered costs.

| <b>From</b>   |            | <b>To</b>   |
|---|------------|---|
| Locus where the poor sick are found. Internally structured in wards where patients lie, or are being operated upon.   | Structures | Locus where specialists, professional nursing and equipment, often hightech or otherwise costly are found. Internally structured around specialties (body parts, techniques or illness).                                |
| Regarded end station of medical system. Badly maintained and typically unhygienic, associated with poverty and death. | Cultures   | Cleanliness, professionalism and efficiency all in service of curing the patient from his illness. Leading role for professional judgement of physician.  |
| Patients residing in beds awaiting the, often only visiting, doctor’s round or surgery.                               | Practices  | Admission and treatment in a hospital following a diagnose indicating such as necessary. Usually this involves a hospital stay and being assigned a bed at which the patient receives his or her care until discharged. |

Table 1.5: Development of hospital based care constellation

<sup>40</sup>The *Armenwet*.  
BOOK III, page 30

## 1.6. PATTERNS IN HEALTH CARE

### Nursing<sup>41</sup>

In the sense of caring for the sick, nursing is among the oldest facets of health care. And with the fact in mind that medical knowledge was very limited up until quite recently as were the possibilities for treatment and actual curing of disease, care was long the only relevant form of health care. In the 19<sup>th</sup> century some developments in the Netherlands paved the way for a new angle to the nursing profession. Changes in legislation, namely the 1818 and 1854 laws on the medical profession and care for the poor, were important as they made municipalities responsible for the health care for the poor. Moreover these laws separated providing food and shelter from health care for the poor, which allowed the reinvention of the hospitals.

#### *1818-1854 / Reconstellation (Stress)*

Legislation comes into effect that makes hospital a place for the sick, redefining the nursing profession as caring for the sick.

Also catholic hospitals began to appear since the repression of the catholic community ended with the new 1848 constitution.

The religious *revéil* that revived the interest of the protestant Christian community in welfare and charity work led to the emergence of Deaconess hospitals. Though named after positions in early Christianity, they in fact modelled themselves in the image of catholic orders. Midway through the 19<sup>th</sup> century, when nursing was still limited to basic tasks (changing bed linen and such), the Deaconess hospitals began educating their nurses. Under the influence of the german Fliedler this education was innovative and of high quality including hygiene. The icon of nursing Florence Nightingale, for example, was trained in Fliedler's school. Nightingale's school of training nurses was highly influential and emancipated the nurse profession and conversely the profession of nurse became an important factor in women's emancipation.

#### *1850-1900 / Empowerment (Pressure)*

Training in the Deaconess order in the Fliedler tradition and the influence of Nightingale make nursing a proper educated profession held in high societal esteem.

The early 20<sup>th</sup> century also saw the rise of the medical specialists (see the specialists constellation for details) and the transformation of the hospital into their natural habitat. The nurses, who once dominated the hospitals, with a lot of authority for the head nurse for instance, were reduced in status to the assistants to the physicians. These developments led to the nursing profession becoming a symbol for women's emancipation to rebel against in the 1960's and 1970's.

#### *1920-1941 / Adaptation (Tension)*

The specialists become more dominant in the hospital and the treatments it offers. This redefines the nurse as the physicians assistant where she once had a quite dominant role.

---

<sup>41</sup>See for this constellation the sources (Binnenkade, 1973; Dane, 1980; Goudswaard, 1994; Knoop & Schuringa, 1998; Spijker, 1979; Zwols, 1985)



BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

The Dutch system also has a parallel development of the nursing constellation in the so called cross organisations<sup>42</sup>. In 1875 the health inspector for the province of North-Holland founded the first cross organisation: the white cross. He did so out of his dissatisfaction of how the local authorities neglected their responsibilities for public health and prevention of infectious disease and epidemics (several outbursts of cholera, scarlet fever and smallpox) and envisioned a sort of red cross in peacetimes, hence the cross<sup>43</sup>. The cross organisations did much in the field of public hygiene and preventive health care and, especially relevant, they trained nurses. For this they became literally emblematic, a professionally trained nurse received a little white cross to wear on her dress. The cross organisations were highly successful, growing and merging in the first decades of the 20<sup>th</sup> centuries to the point were almost all of the Netherlands was served by them. Essentially they were a kind of union of which one became a member in order to make use of their services. The cross organisations were held in high esteem by the government which gave them structural access to funding. This was in the end also a weakness, since funding came with government influence and much of the cross organisation's original tasks were subsumed in public health initiatives favoured by the progressive/social-democrat government in the 1970's, followed by later budget cuts under liberal conservative governments, reducing cross organisations to providing home-care where they once were in fact a constellation themselves.

1875-1970 / Empowerment (Pressure)

Rise of the cross organisations setting the standards for nursing and its education. Cross organisations take a central and rather autonomous role in public hygiene and preventive health care. From 1940 to 1970 membership numbers increase from 30 to 80 percent of citizens.

Where the nursing profession was at first a symbol for women's emancipation the societal appreciation declined in the second half of the twentieth century. Nursing became a profession associated with hard work and low wages, low in the hierarchy of the medical system. Educational reforms later countered this development somewhat.

1950-1990 / Adaptation (Tension)

Nursing profession's societal appreciation changing, losing its emancipatory appeal, stabilising after higher educational reforms giving it a more solid basis.

| From  | To   |
|---|--|
| Typically within institutions that leave little autonomy for the nurse, be they hospitals, a cross organisation, or a church order institution. | Structures Currently mostly in assisting roles for physicians (in hospitals) or within a dedicated care institution (e.g. nursing home, mental institution). Also home care. |

Continued on next page...

<sup>42</sup> *Kruisverenigingen*, as they are called in Dutch.

<sup>43</sup> The inspector, one Penn, had the bright idea of inverting the logo of the red cross to use as an emblem. He was apparently unaware that Dunant had the same idea inverting the Swiss flag into his red cross logo. The Swiss government later urged the white cross to stop using their national flag.

## 1.6. PATTERNS IN HEALTH CARE

| From  |           | To   |
|---|-----------|--|
| Self-chosen non-prominence of the nurse, stemming from an idea of vocation or higher goal (not necessarily in a religious sense) in serving and caring. Strong association with women's emancipation, though, and much authority for head nurses. | Cultures  | Still much a culture of assisting other medical professionals. Associated with skilled professionalism and the medical hands-on work.                            |
| Caring in its broadest sense, though mostly associated with bedside care. All tasks the physician wants done (the nurse as the 'doctor's hand').  | Practices | Various tasks the patient is unable to perform, maintaining hygiene, assistant to the physician up to and including small medical tasks (bandages, shots, etc.). |

Table 1.6: Development of nursing constellation

### Cross organisations (transient constellation)

Privately organised health care organisations. See under nursing, public health and financing.

#### *1875-1970 / Empowerment (Pressure)*

Rise of the cross organisations setting the standards for nursing and its education. Cross organisations take a central and rather autonomous role in public hygiene and preventive health care. From 1940 to 1970 membership numbers increase from 30 to 80 percent of citizens.

#### *1956-1973 / Adaptation (Stress)*

The 1956 'healthlaw'<sup>44</sup> with the founding of a central public health council (mixed public-private) initiates growing influence of government. More financial support from the government is given at the expense of autonomy.

#### *1974-1980 / Reconstellation (Tension)*

The 1974 white paper<sup>45</sup> is the beginning of the welfare state subsuming a lot of the public health aspects of the cross organisational work. In 1980 the law AWBZ becomes effective for cross organisations arranging financing via sick-fund system.

#### *1982-1990 / Adaptation (Tension)*

Right-winged liberal governments cut in budgets. Reducing the role of the cross organisations to home-care organisations.

### Public health<sup>46</sup>

Initially, in the 19<sup>th</sup> century, public health grew steadily. The hygienist movement of physicians concerned with the public cause saw their efforts incorporated in diverse forms in the regulatory system and the second, more revolutionary minded, generation of hygienists held an extensive lobby for public health measures starting in 1895 with their 'Congresses for Public Health Arrangements'<sup>47</sup>.

<sup>44</sup> *Gezondheidswet*.

<sup>45</sup> The *Structuurnota Gezondheidszorg*. A white paper is comparable to a Dutch *structuurnota*.

<sup>46</sup> See for this constellation the sources (Kappelhof, 1990; Kerkhoff, 1994; Querido, 1965).

<sup>47</sup> *Congressen voor openbare gezondheidsregeling*.

## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

### 1840-1920 / *Empowerment (Pressure)*

A strong lobby of two waves of the hygienist movement induced a steady growth of the constellation. Many public health and hygienic arrangements being adopted.

The larger cities faced all the problems of industrialisation in the period from 1870-1930, crowding, poverty, etc. which effectively forced them to initiate public health services like efforts for pest and disease control, emergency aid, school physicians. This initiative was followed later by smaller cities and towns.

### 1870-1930 / *Reconstellation (Tension/Stress)*

Large cities initiate public health services like efforts for pest and disease control, emergency aid, school physicians to counter effects of industrialisation (crowding, poverty). This initiative followed by smaller cities and towns.

In the beginning of the 20<sup>th</sup> century the desire for cost control leads to municipalities demanding more control over the activities of the civil organisation they preferred to subsidise instead of delivering the health care themselves. The cross organisations have an important role in delivering this care.

### 1875-1970 / *Empowerment (Pressure)*

Rise of the cross organisations setting the standards for nursing and its education. Cross organisations take a central and rather autonomous role in public hygiene and preventive health care. From 1940 to 1970 membership numbers increase from 30 to 80 percent of citizens.

Several attempts to create a national health service failed. Van Raak (2010) points out four. The first, as early as 1849 already, stranded because the responsible minister (Thorbecke) was a fierce opponent for liberal reasons. This was applauded by most of the medical community who preferred a free practice of medicine instead of becoming government employees. A second attempt in 1920, aiming to instate mandatory public health services like the ones already present in the large cities, met opposition in parliament and the medical community. The third attempt was directly inspired by the British national health service, with which the Dutch government became acquainted while exiled during the Second World War. However, Dutch society in those decades directly following the war was denominationally segregated<sup>48</sup> to such an extent that the law that was eventually passed in 1965 only reconfirmed the position of the private parties and further marginalised the public health constellation.

### 1848-1982 / *Adaptation (Tension)*

Several attempts to create some form of national health service all fail. Public health remains a government subsidised privately organised matter. Government takes a lot of control over the cross organisations roughly from 1920-1940.

---

<sup>48</sup>This segregation is referred to as *verzuiling* in Dutch. This translates to 'pillarisation'. The pillars were: Protestant, Catholic, socialist, and sometimes a liberal (neutral) is distinguished.

## 1.6. PATTERNS IN HEALTH CARE

The last attempt was in 1974 when the ruling coalition was predominantly progressive and social democratic. This attempt failed since the following governments were far more right winged and facing economical crises, precisely when the law laying the foundations for such a system was passed.

### *1973-1982 / Reconstellation (Tension)*

The welfare state subsumes a lot of the public health aspects of the cross organisational work.

In the following decades, however, public health did broaden its functional scope with venereal disease education, disaster management and preventive disease screening. In 1990 a law was passed making it compulsory for a municipality (or municipalities in cooperation) to run a public health service, called a municipal health service (GGD).

### *1982-1990 / Empowerment (Stress)*

Public health broadens its scope to for instance venereal disease education and preventive disease screening. Running of a public health service becomes compulsory for municipalities or co operations thereof.

| From   |            | To   |
|--|------------|--|
| Regionally and privately organised in the form of cross organisations with a system resembling workers' unions, or locally and publicly by cities. | Structures | Locally organised at the level of the municipality (GGD). Medical staff directly employed as civil servants.   |
| Integral approach to health care, especially the social and societal aspects of health and disease.  | Cultures   | Focus on the context of health care and disease. Prevention, education, sanitation and hygiene highly valued.  |
| Prevention of infectious disease, public hygiene, home care, neighbourhood care, midwifery, education of the public and of nurses.                 | Practices  | Vaccination, infectious disease control, some care for the poor, advising governments. Contrast with complaint-diagnose-treatment approach. Also some direct intervention. |

Table 1.7: Development of public health constellation

## Financing

Apart from charity and care for the poor, health care financing and 'insurance' was in the Dutch context a matter of the guild system. With the collapse of that after the French occupation there was a gap for insurance in the second and third decade of the 19<sup>th</sup> century. Although slowly in the beginning the course of that century saw a plethora of sick funds emerging, insuring for medical costs or wages not earned or both.

### *1815-1910 / Adaptation (Tension)*

Steady growth of sick funds after slow start, filling the gap left by the collapse of the guild funds. Growth also driven by a larger and growing

## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

population of the working class.

To avoid the medical profession becoming subjects to the whims of an insurance market, the professional association (NMG) began their own sick funds in 1913 and already required their members to accept payment only from their funds, or from others that met their criteria, from 1912 onward (van Lieburg, 1999).

### *1910-1915 / Adaptation (Stress)*

The medical association (NMG) sets up guidelines for the sick funds that physicians should accept. Later the NMG starts its own funds and boycotts funds that do not meet their requirements.

The first four decades of the 20<sup>th</sup> century saw a continuation of the growth of the sick funds because of industrialisation, population growth and more widely available health care. Something peculiar however that specialised hospital care was almost never insured via normal sick funds, for this one had to rely on other commercial insurers or personal capital. This was a driver for the emancipation of the specialists who exercised their profession in polyclinics on a clientele of poor people paid for via the municipality's care for the poor and sick funds. In the 1920's and 30's some sick funds for the middle class began to cover specialist hospital care as well which went hand in hand with the gradual 'takeover' of the hospital by the specialists. In any case Dutch politics failed to provide proper sick fund legislation and bill after bill stranded in endless debates and fruitless attempts. Loss of wages was however taken care of separately in a 1930 law<sup>49</sup>.

### *1890-1920 / Empowerment (Pressure)*

The poly-clinical treatments by specialists increases highly the relevance of the sick funds.

WWII arranged this matter by force, the German occupier forced the sick fund act<sup>50</sup> on the Dutch system. This act entailed a mandatory membership of a sick fund up to a certain income. Although this act in a certain sense only codified current Dutch practice, the change was fundamental. The coverage was extended to include hospitalisation, dental care and such and patients could no longer be refused by sick funds. This changed the playing field for the sick funds, many smaller ones disappeared and since coverage of treatments and acceptance of patients was arranged by law the previously existing insurance market was essentially killed. The German sick fund act was not revoked after the war.

### *1941 / Reconstitution (Stress)*

German sick fund act ends disorderly system, instates mandatory coverage of, amongst other things hospitalisation, introduces mandatory acceptance of patients and for patients up to a certain income mandatory membership.

After WWII the health care sector grew tremendously and with it the public spending to maintain its financing system. The decades after WWII saw the emergence of public insurances<sup>51</sup>, of which the AWBZ is relevant for health care.

---

<sup>49</sup>The *Ziektewet*.

<sup>50</sup>Again the *Ziekenfondsbesluit*.

<sup>51</sup>*Volksverzekeringen*, as they are called in Dutch.

## 1.6. PATTERNS IN HEALTH CARE

Such an insurance is a sort of sick fund paid for in the form of an income tax. It made it possible for mental and long-term health care to become financed within a national frame where it first was a municipal matter, an arrangement like the care for the poor. This AWBZ-approach divided health care financing into three compartments: 1. Directly funded long-term and mental health care. 2. Sick funds and health care insurances with mandatory acceptance and membership with coverage for essential health care. 3. Voluntary insurance for non-essential health care.

### *1964 / Reconstellation (Tension/Stress)*

The law AWBZ becomes effective arranging much of health care financing in three compartments: direct coverage (mental and long-term health care), sick funds for essential treatments and non-compulsory insurance for non-essential health care.

Current reforms in the health care financing system try to rearrange the latter two to allow more market dynamics to reduce the public costs of health care and provide more choice for the patients.

### *2001-2006 / Reconstellation (Tension)*

Reorganisations are attempted and become effective in 2006 introducing (limited and regulated) market dynamics in the system of health care financing.

| <b>From</b>  |            | <b>To</b>   |
|--|------------|---|
| Either based on private capital, via a sick fund, (religious) charity or publicly arranged care for the poor. No obligation to be insured. Essentially a health care market, insurances and sick funds available in free competition. Free choice of doctor/physician. | Structures | Publicly enforced mandatory insurance with private corporations, focus on paying for curative treatment. Direct funding of some third party institutions (mainly long-term and mental health care). |
| A de facto free market attitude. Both health care and insurance available in free competition. Some organised solidarity within the medical community and workers' sick funds.   | Cultures   | Hybrid: between arranging for a basic right to health care and a commercial insurance market; between a public, societal, affair and individual responsibility.                                     |
| Doctors had honorariums, pharmacists sold drugs directly, doctors in academic hospitals were paid as such and specialists in polyclinics were paid mostly via sick funds.  | Practices  | Essential treatments paid for by insurance companies, with whom one is mandatorily insured. Some direct funding of institutions.  |

Table 1.8: Development of financing constellation

## **1.7 A closer look: addressing the questions**

The grand scheme has mostly been a disentanglement of the historical development into the patterns of the several constellations. In order to address the questions risen in the previous, a closer look needs to be taken at the interactions between the constellations and how their developments influenced each other. Reconstructing patches of the web now that the threads have been identified, in other words.

It might be appropriate first to reiterate the questions here:

- How could the current health care regime emerge?
- Why the tendency for other constellations to conform their structures, cultures and practices with this regime?
- Why was a national health service never adopted, nor any other nationally organised system of public health?
- Why hasn't a powerful constellation around prevention formed?
- Why did integral forms of care that were present not survive?
- Why have the professions of the surgeon and the physician become so integrated?
- Why has nursing lost so much of its societal standing?
- Can we understand the current tension, stress and pressure on the health care system, if any such are present?

The answer to the first question is in a certain sense the entire case study itself and it should emerge in the reconstruction that follows. The last question will be addressed in the conclusion and discussion section. The other questions will be answered in the reconstruction, not necessarily in this order, but rather by following the threads, the chains of patterns and letting them follow naturally.

Early on in this case study the remark was made that the current regime, the essence of how the health care system functions, is centred around medical specialties. This shows in the way several aspects of the system are structured, ranging from departments in hospitals to the physician's academic training to the way insurance companies cover treatments. It is also apparent in the functioning of constellations that have more 'distance' from the regime but still adhere to it. For instance the nursing or more broadly speaking 'care' constellation assuming a subordinate role with respect to the treatment of a patient which is directed by specialists. This is noteworthy since it could very well have been the other way around, with a patient under supervision of a care professional, like a nurse, with specialist help being called in if deemed necessary. Another example is the tendency in mental health care to treat mental issues as if they were physical illnesses, that is, in short, medication centred treatment. Also symptomatic of the dominance of the regime is the role assumed by the per definition generalist constellation of general practice, as the gatekeeper of specialist care and treatment.

### 1.7. A CLOSER LOOK: ADDRESSING THE QUESTIONS

If this regime is embodied by the constellations of specialists, hospitals and financing then it is only natural to see what patterns led to their current functioning and relations. First of all it is important to notice that both hospitals and financing of health care already existed in some form for a long time, before the medical specialties were to appear in the 19<sup>th</sup> century. It then suggests itself that the conditions that drove the development of specialties also drove in direct or indirect manner the course of development of the hospitals and financing.

This condition was obviously that of tension, in the sense that health care lagged behind the progresses in medical science. This entailed new scientific knowledge yet to be implemented as well as new instruments and equipment that became available. This drove distinct adaptation patterns in both the constellations of hospitals and of specialists.

For the hospitals, back then in the second half of the 19<sup>th</sup> century still the domain of medical professors, surgeons and nurses, it meant the introduction of hygiene, of a- and antiseptic working and anaesthetics. The latter was quite an immediate success, since it allowed longer operations without the pain distracting the surgeon. The introduction of hygiene and a- and antiseptic working was embraced more reluctantly and with large scepticism.

For the specialists the landscape developments drove a much more fundamental adaptation pattern. Although the notion of a specialist already existed in the early 19<sup>th</sup> century, it became more and more necessary to specialise if one was to remain at the front of developments. No surprise that some of the oldest medical specialties followed the introduction of a medical instrument, notably ophthalmology with the ophthalmoscope.

Although the same conditions drove adaptation patterns for the specialists and hospitals they still remained very separate constellations indeed, although in hindsight one can say<sup>52</sup> that the stage was being set for their later confluence. In any case the financing constellation and the specialists were about to go to an empowerment pattern driven by the pressure they themselves generated:

The rise of the specialists as a constellation was already to a certain extent self-reinforcing as they found their own niche of practicing in polyclinics from, roughly from 1890 to 1920. On the one hand the polyclinic work gave the specialists a large pool of patients to help and, for the development of the constellation perhaps more important, to learn from. On the other hand polyclinics were a place to receive affordable treatment for those who had to rely on sick funds, that more often than not did not cover hospitalisation.

The financing constellation initially went through an adaptation pattern, growing steadily, filling the void left by the guild funds and profiting from a growing working class. Then, when polyclinical treatment became an important factor in medical care, the relevance and status of the sick funds grew, which worked to their empowerment greatly while they thrived and a host of them emerged.

Here the patterns in the development of financing and the specialists have begun to mutually influence each other, one could almost speak of coevolution<sup>53</sup>. Subsequently the financing constellation goes through an adaptation pattern under the stress of the situation with the plethora of sick funds around. The medical association (NMG) takes the sick funds very seriously in the period of

<sup>52</sup>One can always say a lot of things in hindsight.

<sup>53</sup>Though this is said with hesitation since the theoretical framework is not explicitly evolutionary.



## *BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"*

1910-1915, by setting up guidelines for good sick funds. Later on the association founds its own funds and starts boycotting the funds that do not meet their requirements.

This stress in turn leads to adaptation for the specialists. The medical association goes through internal struggles around the position of the specialists. The association is dominated in numbers by general practitioners that see a large group of sick fund patients staying out of their reach. This leads to the first steps in a system of referral where the general practitioner becomes a gatekeeper for specialist care. Here the skeleton emerges for a system around specialist treatment assessed first by a general practitioner and financially covered per treatment.

Around the 1920's the constellations of the hospital and the specialists went through patterns of adaptation and empowerment respectively, which were to bring them together much in the form still present these days. Although polyclinical treatment was highly valued, the clinics themselves were on the decline because of rising costs of, amongst others, costly medical equipment. The hospitals however recognised the increasing importance of the specialists and were in a position to purchase large and costly medical equipment and their adaptation was that they brought in the specialists, so to speak, opening polyclinics in their hospitals, empowering the specialists since these now had access to hospital facilities.

The **integration of the professions of the surgeon and the physician** can now also be understood better. The medical specialists had a strong culture of practicing their art, a hands-on attitude to medicine cultivated in their polyclinics and simultaneously the legislation and educational reforms, strived for a so-called unity of the medical profession directing surgery towards becoming a medical speciality itself. In the hospital the twain met.

At this point, somewhere near the end of the interbellum, the health care regime as it functioned at least until the end of the time span of this case study was to a large extent in working order. A series of reconstellation patterns institutionalised this. First there were reforms in the academic education in medicine in the 1930s, whereby medical specialties were recognised and physicians subsequently officially registered as such. Second there was the German sick fund act in the Second World War. Where Dutch politics had time and again failed to arrange overarching legislation concerning health care financing, the German occupier forced it upon the system in 1941. The sick fund act arranged mandatory insurance for all people below a certain income and mandatory acceptance of patients by sick funds. Also coverage of treatment was broadened almost to all treatment including specialist treatment and hospitalisation. The third reconstellation concerned the law AWBZ of 1964, which structured health care in three compartments, direct funding for long-term and mental health institutions, the sick funds for essential health care and non-compulsory insurance for non-essential health care.

The early confluence of the financing system with the system of specialists which led to a standard where coverage was per treatment, had a conceptual and structural simplicity to it which is probably a factor in its robustness and success. This alone however is not enough to explain **why other constellations appear to conform** to this way of functioning.

For the mental health constellation the paradox appears not that profound. Although psychiatry is definitely a world apart it has been a medical speciality

### 1.7. A CLOSER LOOK: ADDRESSING THE QUESTIONS

almost from the beginning, in fact one of the oldest medical specialties. And although a lot of mental health institutions are directly and publicly funded via one of the aforementioned AWBZ-compartments, medication — a substantial part of mental illness treatments — is prescribed and paid following the normal insurance system.

Speaking of which, the pharmacy constellation conformed to the regime functioning in a very early stage already. That is, they united in 1841 in an association, empowering themselves, and their successful lobby secured their position with respect to the other medical professions, as an aiding medical profession<sup>54</sup>, in the 1865 Thorbecke-laws. Later on the system of referral with the general practitioner prescribing medicine or referring to specialist care, and again the structures of health care financing, consolidated the position and functioning of the pharmacy constellation with respect to the regime.

One is tempted to speculate whether the strong structuring power of the regime functioning is the reason **why there appears so little room for more integral forms of health care**. Integral care would per definition structurally go against the grain of the orientation around medical specialties and the complaint-diagnose-treatment paradigm. Moreover on the cultural side of things it would be an entirely different attitude towards health care than the curing imperative, the view of health as absence of illness.

Could this then also be a reason for **the decline of the social status of the nursing profession**? Because the culture of the nursing constellation is about the general care for patients as opposed to delivering specific treatment for specific ailments? Perhaps another cultural aspect of nursing had more to do with it, that of the serving attitude which made the adaptation to the rise of the specialists in hospitals probably seem natural. Which leads one to think that a structural aspect was at work here too: as the hospitals took on the structures of the medical specialties, the nurses in their assisting role, naturally became distributed over the departments. This continuous adaptation under tension made the nursing profession lose much of its emancipatory appeal which was counteracted by reforms in higher education giving the nurse a more solid educational basis, again based on specialties.

The above was about how the current regime emerged as the result of conditions for, driving patterns of, transitional change. How the functioning of this regime is a dominant, structuring force to which other constellations conform, was also addressed. One of the questions raised, however, is more or less about the opposite: **why did a public health system, or a national health service not emerge** in the Dutch health care system?

Answering questions about why something did *not* happen is precarious. Explaining non-occurrence would amount to singling out a mechanism that was sure not to lead to the opposite of the non-occurrence to be explained. Not only is there in principle an unlimited amount of such mechanisms, in theories of complex social phenomena like transitions it is already arguable if occurrence itself can be attributed to some causal mechanism. Still, one wants to know.

In the case at hand it might give some insight to know that attempts to form a system of public health, even on a national level were made. No less than four attempts to a sort of national health service were made and on the regional and municipal level such systems have existed. Then, why was this not

---

<sup>54</sup>A profession to aid other medical professions, which sort of says it all.

## *BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"*

institutionalised and absorbed into a slightly different regime functioning than the one that did emerge?

Two hypotheses suggest themselves to explain this. The first is that simply no conditions were present that would drive patterns transitionally changing the health care system from a public-private hybrid to a nationally organised, publicly financed system. The other one is that patterns that would transitionally change the health care system in such manner, would thereby also induce conditions that would drive patterns counteracting such developments. Both hypotheses notwithstanding that such patterns were attempted and advocated and appeared a solution to problems at hand.

Considering the attempts at a national health service to begin with, one sees that in all four cases the initiated reconstellation pattern — government proposals for legislation — was to create either tension with some landscape development or property or stress within the health care system itself<sup>55</sup>. The first and second attempts (1849 and 1920) for example ran into a political climate — liberal-confessional — hostile to such initiatives, in other words the patterns would induce tensions with the political system. They would also induce stress within the health care system since the medical community opposed the idea of becoming state employees and feared for their highly valued free practice of medicine. The third attempt was after the Second World War. During the war, the Dutch government that was in exile in the UK got acquainted with the British national health service and was inspired to propose such a system for Dutch health care. The Dutch society, reflected in its religious-political factions, was at the summit of its pillarisation (denominational segregation). Such a system revision would therefore entail much tension with the social-political landscape. In the end (1965) the law that was finally adopted was so heavily edited that it confirmed rather than changed the then current situation. In 1974 however the right conditions did appear to be present in the form of tensions. Curative health care was fragmented, inefficient and expensive for the government, moreover the political climate had never in the Dutch history been so left-wing. These tensions indeed drove a reconstellation pattern that was about to introduce a national health service-like system. The law was actually passed in 1982 but was unfortunate enough to meet changing conditions about the time it became effective, a new tension — recession — emerged and the political climate swung to the right. The law was retracted.

Ironically enough, private initiatives for public health, specifically the cross organisations, also suffered from the dynamics sketched above. At various points the conditions for the cross organisations as the designated entities for public health services appeared more than favourable. Indeed in 1970 as much as 80% of the Dutch were member. The cross organisations were already pillar-proof, in that there was one for every denomination and they were privately organised in a worker's union kind of way. The benign political climate — apparently also a form of tension — and the stress of risen costs, invited increased government funding in an adaptation pattern for which the cross organisations paid with their autonomy. In the second half of the 1970's more and more of the cross organisational functioning is subsumed in the growing welfare state and eventually they are financed completely through the sick funds via the law AWBZ in 1980, effectively reducing them to home care organisations. The irony is doubled since

---

<sup>55</sup>See also the description of the financing constellation for more details.

## 1.8. A CLOSER LOOK ON THE BIG PICTURE

in recent years the cry is heard for a more integral, patient centred, health care, which was much like what the cross organisations offered in addition to their (other) public health services.

There remains the question **why a proper constellation around prevention never really emerged**. Considering the influence the hygienist movement has had in the 19<sup>th</sup> century, which could be a pressure driving the empowerment of such a constellation; or considering the tension that urbanisation and industrialisation posed that was a driver for the constellations of specialists and financing for instance. At the one hand the answer lies in the public health developments discussed above, with which a preventive health care constellation would have a lot of functional overlap. Indeed one of the latest patterns affecting the public health constellation was one of empowerment in the sense that each municipality was to maintain a GGD (municipal public health service). These GGD's among other things have a role in preventing infectious and venereal diseases, dealing with epidemics and health education. On the other hand much of the societal need for prevention of disease has been subsumed by other systems, like waste disposal, pest control, and sewage systems which are dealt with outside the health care context.

## 1.8 A closer look on the big picture

The questions answered in the previous section were raised in a transition overview. In that section three phases were distinguished for the overall transition. It is interesting to see whether, with the detailed, pattern-based, picture of the grand scheme, these phases still can be distinguished and whether they still make sense.

Since the overall transition has led to a health care regime consisting of the closely intertwined constellations around the hospital, health care financing, and specialist care, it appears logical to investigate whether patterns shaping the developments of these constellations can be identified that somehow demarcate the phases of the overall transition. This appears to be the case.

After a period of pre-development, all three constellations mentioned above go through a relevant transition pattern around the arrival of the 20<sup>th</sup> century. The specialists empower themselves through their polyclinical work, the hospital becomes a hygienic place ready for a next level of medical practice, and the sick funds become a relevant factor in the financing constellation. This latter pattern also strongly couples the development of the constellations of specialists and financing.

This period is followed by a period in which these protagonist constellations go through several patterns that elaborate these developments, for example the hospital constellation becomes coupled with the specialist one. This period could very well be called the acceleration phase, since in this period the rest of the societal system reshapes as well, conforming to what is to become the health care regime. For example, within the medical association the roles of general practitioners versus specialist crystallise out, the nurse becomes the 'doctor's hand'. In the meanwhile the technical-scientific possibilities in medicine take flight.

A phase of stabilisation is entered from the 1930's on. All three constellations go through reconstellation patterns in this period, consolidating the pre-

## BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

vious developments. Changes in the educational system academically embed the specialties wholly and especially the 1941 German sick fund act gives the new regime an important institutional basis. The stabilisation eventually culminates with the law AWBZ becoming effective. The demarcating patterns are tabularised in table 1.9.

### 1.9 Conclusion and discussion: ironies and paradoxes

In the process of unravelling and reconstructing questions about transitions in Dutch health care were raised, framed and addressed which perhaps might not have been possible if this framework was not used. Several questions were raised already in the beginning of the case study that were subsequently reformulated and addressed in the previous section. It would be superfluous to repeat the treatment of the questions here, instead it might be nice to highlight that using this theory either pointed out some ironies and paradoxes in the current Dutch health care system or shed new light on them. For example:

Health care is about curing illness. This appears so obvious that it might not even raise an eyebrow at first. It is symptomatic — *excusez le mot* — that one of the most important factors in the success of modern health care, namely the scientific approach to curing illnesses, has in a certain sense led it away from care for the patient. It appears that for much of the practice of health care the patient is more the carrier of his or her illness than a person with a problem. This is also reflected in that a constellation of patients could not be identified in any useful way. Whether this is entirely a bad thing is however disputable as well since it is also at the core of the effectiveness of modern medicine.

This touches on the related irony that the health care system is becoming contested where it is at its best. More successful than ever in a sense by focussing on curing and a scientific approach to that, the well being of the patient as a whole has somewhat left the scope of the health care system. It appears that this stress will become more important in the coming time since at the one hand we can now do things of which it is questionable if we want them, referring to ethical debates concerning genetics, euthanasia and extreme longevity, whereas at the other hand we appear not to be able to do some things we want, such as caring for an growing ageing population or provide for more abstract needs for well being like happiness, the lack of which is becoming regarded more and more as a medical condition.

In view of the recurring cries for a more integral approach to health care it is interesting to find out that the institutions for integral care that were present in the Dutch context, like the cross organisations, disappeared. These organisations by their own merits and efforts became a relevant constellation with many facets of functioning related to public health and hygiene, education, care and much more. The various failures of Dutch politics in forming a national public health service were an important factor in their waning, since much of their functioning became absorbed in the facilities of the well-fare state. As well as did the tendency to favour denominationally segregated institutions as a solution to public health, leading to more government influence in exchange for funds. Budget cuts later on did the rest.

1.9. CONCLUSION AND DISCUSSION: IRONIES AND PARADOXES

| Take-off  | Touch-down  |
|---|---|
| <p><i>1890-1920 / Empowerment (Pressure)</i><br/>Specialists become a group to be reckoned with in the medical community through their work in their polyclinics and their subsequent penetration of the hospitals.</p> <p><i>1860-1900 / Adaptation (Tension/Stress)</i><br/>Anaesthetics, hygiene, a- and antiseptis transform the hospital into a place of non-lethal surgery.</p> <p><i>1890-1920 / Empowerment (Pressure)</i><br/>The poly-clinical treatments by specialists increases highly the relevance of the sick funds.</p> <p style="text-align: center;">PRE-DEVELOPMENT</p> | <p><i>1930-1964 / Reconstellation (Stress)</i><br/>Changes in educational system, the German sick fund act (1941) and later the law AWBZ further defined and consolidated the position of the specialist.</p> <p><i>1941 / Reconstellation (Stress)</i><br/>The German sick fund act makes hospitalisation more accessible for sick fund patients. It also decouples to a large extent the insurance fees from the covered costs.</p> <p><i>1941 / Reconstellation (Stress)</i><br/>German sick fund act ends disorderly system, instates mandatory coverage of, amongst other things hospitalisation, introduces mandatory acceptance of patients and for patients up to a certain income mandatory membership.</p> <p><i>1964 / Reconstellation (Tension/Stress)</i><br/>The law AWBZ becomes effective arranging much of health care financing in three compartments: direct coverage (mental and long-term health care), sick funds for essential treatments and non-compulsory insurance for non-essential health care.</p> <p style="text-align: center;">ACCELERATION</p> <p style="text-align: center;">STABILISATION</p> |

Table 1.9: Demarcation of phases in the overall transition

## *BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"*

Attempts to unify the medical profession increased its tendency to specialisation. It appears in hindsight that specialisation and the road to a complaint-diagnose-treatment approach to health care was inevitable. One illustration of this in this chapter was the integration of the profession of the surgeon and the physician. The demand that a surgeon be a general medical professional, led to its incorporation in the educational system as its own medical specialty. Something similar happened to general practice. Thus ironically, by the time the much desired unity was accomplished diversity was the norm.

Current reforms are reinforcements of current forms. It is curious to see that the perceived problems of the health care system such as high public expenditure, the low effectiveness of the bureaucratic machinery involved, etc. are attempted to be solved by reinforcing the very structures that are at the root of these phenomena, namely focussing on health care in terms of treatments and financial coverage of care per treatment.

### **1.9.1 Points of consideration and things left to be done**

Although this case study was an integral one it is certainly not complete nor exhaustive. More constellations could have been identified, which was already signalled when they were introduced. Think for instance of alternative practice, dentistry or occupational health. Depending on what one wants to know, what questions one wants to ask, more or less could be taken into account. In this chapter a first attempt was made to a transition study of Dutch health care using this theoretical frame. If it raised new questions that invite more research, then it has been successful.

**1.10 Appendix:**  
— *The pattern matrix*



BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

| Constellation / Pattern — |  |   |   |
|---------------------------|--|---|---|
| General practice          | —  | <p><b>1865</b><br/>— <i>Reconstellation (Stress)</i><br/>The Thorbecke-laws define and protect the title 'physician'. Merging the professions of the surgeon and the academical doctor. These laws also limit authorisation to practice medicine, and thus general practice, to those having passed a state exam.</p> | <p><b>1912-1930</b><br/>— <i>Adaptation (Stress)</i><br/>Internal struggles in the medical association lead to positioning of specialists with respect to general practitioners and steps toward a system of referral.</p>                      |
| Mental health             | <p><b>1818-1883</b><br/>— <i>Reconstellation (Tension)</i><br/>Enlightenment and French revolution inspire laws formally distinguishing between the sick, the mentally ill, vagabonds, and criminals. Incarceration of the mentally ill in penitential institutes is prohibited.</p> | <p><b>1850-1900</b><br/>— <i>Adaptation (Stress)</i><br/>Asylums become more like hospitals under the 'somatic approach': being treated similarly as having a physical illness, in the sense of treatment in bed. Under the stress of overcrowded asylums and natural scientific successes.</p>                       | <p><b>1920-1950</b><br/>— <i>Adaptation (Tension)</i><br/>Medicalisation and deinstitutionalisation. New mental illnesses recognised (like schizophrenia). Patients treated more and more like in a hospital in the sense of brief stays.</p>   |
| Nursing                   | <p><b>1818-1854</b><br/>— <i>Reconstellation (Stress)</i><br/>Legislation comes into effect that makes hospital a place for the sick, redefining the nursing profession as caring for the sick.</p>  | <p><b>1850-1900</b><br/>— <i>Empowerment (Pressure)</i><br/>Training in the Deaconess order in the Fliedler tradition and the influence of Nightingale make nursing a proper educated profession held in high societal esteem.</p>  | <p><b>1920-1941</b><br/>— <i>Adaptation (Tension)</i><br/>The specialists become more dominant in the hospital and the treatments it offers. This redefines the nurse as the physicians assistant where she once had a quite dominant role.</p> |

Table 1.10: General practice, mental health, and nursing. Part I

1.10. APPENDIX: THE PATTERN MATRIX

| Constellation / Pattern —   |  |   |
|---|--|---|
| <p><b>1925-1939</b><br/>— <i>Empowerment (Stress)</i><br/>Though still stronger in numbers, general practitioners perceive their position weakened and unite within the association (NMG). This culminates in the founding of the Dutch Association of General Practitioners (later to become the NVH). This association was highly influential in the emancipation and definition of the profession.</p> | <p><b>1930-1964</b><br/>— <i>Reconstellation (Stress)</i><br/>Changes in educational system, the German sick fund act (1941) and later the law AWBZ further defined and consolidated the position of the specialist. This is relevant for general practice as its relation to specialist treatment is finally codified. The general practitioner becomes a gatekeeper of the medical system.</p> | <p><b>1968-1973</b><br/>— <i>Adaptation (Stress)</i><br/>General practice becomes the medical specialty of generalism. First in academia, later formalised by law and registration.</p> |
| <p><b>1952-1975</b><br/>— <i>Adaptation (Tension)</i><br/>The discovery of the psychotherapeutic effectiveness of CPZ brings pharmaceuticals into mental health care.</p>   | <p><b>1964</b><br/>— <i>Reconstellation (Stress)</i><br/>The AWBZ-law arranges financing for, amongst others, mental health care, making it no longer a municipal matter. Mental health institutions now directly funded by the Dutch government.</p>  | <p>—</p>  |
| <p><b>1875-1970</b><br/>— <i>Empowerment (Pressure)</i><br/>Rise of the cross organisations setting the standards for nursing and its education. Cross organisations take a central and rather autonomous role in public hygiene and preventive health care. From 1940 to 1970 membership numbers increase from 30 to 80 percent of citizens.</p>   | <p><b>1950-1990</b><br/>— <i>Adaptation (Tension)</i><br/>Nursing profession's societal appreciation changing, losing its emancipatory appeal, stabilising after higher educational reforms giving it a more solid basis.</p>  | <p>—</p>  |

Table 1.11: General practice, mental health, and nursing. Part II

BUILDING 1. “TRANSITIONS IN DUTCH HEALTH CARE”

| Constellation / Pattern — |   |   |
|---------------------------|---|---|
| Specialist health care    | <p><b>1850-1930</b><br/>— <i>Adaptation (Tension)</i><br/>Progress of medical science saw growth of the number of specialists and their becoming a group within the medical community.</p>  | <p><b>1865-1878</b><br/>— <i>Reconstellation (Stress)</i><br/>Necessity to organise health care and higher education leads to Thorbecke and education laws. More recognition for specialists and more educational possibilities to become one (HBS<sup>56</sup>).</p> |
| Hospital based care       | <p><b>1818-1854</b><br/>— <i>Reconstellation (Stress)</i><br/>Legislation comes into effect that makes hospital a place for the sick (no longer as a place for temporary lodging for instance) and municipalities become responsible for the health care of their poor.</p> | <p><b>1848-1920</b><br/>— <i>Adaptation (Tension)</i><br/>A new spirit (<i>r eveil</i>) and the new constitution banning repression of the Catholic community leads to the founding of many new hospitals (Deaconess, Catholic and private).</p>                      |
| Financing                 | <p><b>1815-1910</b><br/>— <i>Adaptation (Tension)</i><br/>Steady growth of sick funds after slow start, filling the gap left by the collapse of the guild funds. Growth also driven by a larger and growing population of the working class.</p>                            | <p><b>1890-1920</b><br/>— <i>Empowerment (Pressure)</i><br/>The poly-clinical treatments by specialists increases highly the relevance of the sick funds.</p>   |

Table 1.12: Specialist health care, hospital based care, and financing. Part I

<sup>56</sup> *Hogere BurgerSchool*, comparable in level of education to the *gymnasium*, back then and nowadays the highest level in secondary education. The HBS only did not provide education in the classical languages.

1.10. APPENDIX: THE PATTERN MATRIX

| Constellation / Pattern —   |  |   |
|---|--|---|
| <p><b>1890-1920</b><br/>— <i>Empowerment (Pressure)</i><br/>Specialists become a group to be reckoned with in the medical community through their work in their polyclinics and their subsequent penetration of the hospitals.</p>  | <p><b>1912-1930</b><br/>— <i>Adaptation (Stress)</i><br/>Internal struggles in the medical association lead to positioning of specialists with respect to general practitioners and steps to a system of referral.</p>   | <p><b>1930-1964</b><br/>— <i>Reconstellation (Stress)</i><br/>Changes in educational system, the German sick fund act (1941) and later the law AWBZ further defined and consolidated the position of the specialist.</p>  |
| <p><b>1920-1941</b><br/>— <i>Adaptation (Pressure)</i><br/>The specialists become more dominant in the hospital and the treatments available there. Specialists bring prestige and knowledge in the hospitals and they profit from the facilities.</p>                                    | <p><b>1941</b><br/>— <i>Reconstellation (Stress)</i><br/>The German sick fund act makes hospitalisation more accessible for sick fund patients. It also decouples to a large extent the insurance fees from the covered costs.</p>   | —   |
| <p><b>1941</b><br/>— <i>Reconstellation (Stress)</i><br/>German sick fund act ends disorderly system, instates mandatory coverage of, amongst other things hospitalisation, introduces mandatory acceptance of patients and for patients up to a certain income mandatory membership.</p> | <p><b>1964</b><br/>— <i>Reconstellation (Tension / Stress)</i><br/>The law AWBZ becomes effective arranging much of health care financing in three compartments: direct coverage (mental and long-term health care), sick funds for essential treatments and non-compulsory insurance for non-essential health care.</p> | <p><b>2001-2006</b><br/>— <i>Reconstellation (Tension)</i><br/>Reorganisations are attempted and become effective in 2006 introducing (limited and regulated) market dynamics in the system of health care financing.</p> |

Table 1.13: Specialist health care, hospital based care, and financing. Part II

BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"

|                     | Constellation / Pattern —  |  |  |
|---------------------|--|--|--|
| Pharmacy            | <b>1841-1865</b><br>— <i>Empowerment (Stress)</i><br>Pharmacists unite in the Dutch association for pharmacy (NMP, later KNMP) when the state committee to investigate new health care legislation is summoned. Their lobby secured the position of the pharmacist as a medical profession in the 1865 laws. | <b>1914-1930</b><br>— <i>Adaptation (Tension)</i><br>Isolation in ww1 and post war boom in chemical industry make large scale industrial production of pharmaceuticals possible.   | <b>1952-1975</b><br>— <i>Adaptation (Tension)</i><br>The discovery of the psychotherapeutic effectiveness of CPZ brings pharmaceuticals into mental health care.   |
| Cross organisations | —  | <b>1875-1970</b><br>— <i>Empowerment (Pressure)</i><br>Rise of the cross organisations setting the standards for nursing and its education. Cross organisations take a central and rather autonomous role in public hygiene and preventive health care. From 1940 to 1970 membership numbers increase from 30 to 80 percent of citizens. | <b>1956-1973</b><br>— <i>Adaptation (Stress)</i><br>The 1956 'healthlaw' <sup>57</sup> with the founding of a central public health council (mixed public-private) initiates growing influence of government. More financial support from the government is given at the expense of autonomy.  |
| Public health       | <b>1840-1920</b><br>— <i>Empowerment (Pressure)</i><br>A strong lobby of two waves of the hygienist movement induced a steady growth of the constellation. Many public health and hygienic arrangements being adopted.   | <b>1870-1930</b><br>— <i>Reconstellation (Tension / Stress)</i><br>Large cities initiate public health services like efforts for pest and disease control, emergency aid, school physicians to counter effects of industrialisation (crowding, poverty). This initiative followed by smaller cities and towns.                           | <b>1875-1970</b><br>— <i>Empowerment (Pressure)</i><br>Rise of the cross organisations setting the standards for nursing and its education. Cross organisations take a central and rather autonomous role in public hygiene and preventive health care. From 1940 to 1970 membership numbers increase from 30 to 80 percent of citizens. |

Table 1.14: Pharmacy, the cross organisations, and public health. Part I

<sup>57</sup> *Gezondheidswet.*

1.10. APPENDIX: THE PATTERN MATRIX

| Constellation / Pattern —  |   |  |
|--|---|--|
| <p><b>1974-1980</b><br/>— <i>Reconstellation (Tension)</i><br/>The 1974 white paper<sup>58</sup> is the beginning of the welfare state subsuming a lot of the public health aspects of the cross organisational work. In 1980 the law AWBZ becomes effective for cross organisations arranging financing via sick-fund system.</p> | <p><b>1982-1990</b><br/>— <i>Adaptation (Tension)</i><br/>Right-winged liberal governments cut in budgets. Reducing the role of the cross organisations to home-care organisations.</p> | <p>—</p>   |
| <p><b>1848-1982</b><br/>— <i>Adaptation (Tension)</i><br/>Several attempts to create some form of national health service all fail. Public health remains a government subsidised privately organised matter. Government takes a lot of control over the cross organisations roughly from 1920-1940.</p>                           | <p><b>1973-1982</b><br/>— <i>Reconstellation (Tension)</i><br/>The welfare state subsumes a lot of the public health aspects of the cross organisational work.</p>                      | <p><b>1982-1990</b><br/>— <i>Empowerment (Stress)</i><br/>Public health broadens its scope to for instance venereal disease education and preventive disease screening. Running of a public health service becomes compulsory for municipalities or co operations thereof.</p> |

Table 1.15: Pharmacy, the cross organisations, and public health. Part II

<sup>58</sup>The *Structuurnota Gezondheidszorg*. A white paper is comparable to a Dutch *structuurnota*.

*BUILDING 1. "TRANSITIONS IN DUTCH HEALTH CARE"*

*BOOK III, page 54*

# References

- BIERMAN, A. I. 1988. *Van artsenijmengkunde naar artsenijbereidkunde : ontwikkelingen van de Nederlandse farmacie in de negentiende eeuw*. Rodopi.
- BINNENKADE, C. 1973. *Honderd Jaar Opleiding tot Verpleegkundige in Nederland*. Ph.D. thesis, Faculty of Medicine of the Free University.
- BREMER, G. J. 2006. *Huisarts zijn in het interbellum*. Erasmus Publishing.
- CALDWELL, A. E. 1970. *Origins of psychopharmacology : from CPZ to LSD*. Charles C. Thomas.
- DANE, C. 1980. *Geschiedenis van de Ziekenverpleging*. De Tijdstroom.
- DANKERS, J. & J. VAN DER LINDEN 1996. *Van regenten en patiënten. De geschiedenis van de Willem Arntsz Stichting: Huis en Hoeve, Van der Hoevenkliniek en Dennendal*. Amsterdam.
- GOUDSWAARD, N. 1994. *Inleiding tot de geschiedenis van de verpleegkunst*. Erasmus Publishing.
- VAN DER HEYDEN, J. T. M. 1994. *Het ziekenhuis door de eeuwen : Over geld, macht en mensen*. Erasmus Publishing.
- HOUWAART, E. S. 1991. *De hygiënisten: artsen, staat & volksgezondheid*. Historische Uitgeverij Groningen.
- 1996. *Het Ziekenhuis: De ontwikkeling van het ziekenhuis in de moderne tijd*, chapter Casus 3, pages 237–355. Open Universiteit.
- HUISMAN, F. 1999. Van bedreiging tot bondgenoot. de transformatie van de farmaceutische industrie in nederland, 1880-1940. *Tijdschrift voor sociale geschiedenis*, 25(4): 443.
- JUCH, A. 1997. *De medisch specialisten in de nederlandse gezondheidszorg: Hun manifestatie en consolidatie 1890-1941*. Erasmus Publishing.
- KAPPELHOF, A. 1990. *Tussen Aansporen en Opsporen: Geschiedenis van 25 jaar Staatstoezicht op de Volksgezondheid*. SDU Uitgeverij.
- KERKHOFF, A. 1994. *Honderd Jaar Gemeentelijke Geneeskundige en Gezondheidsdiensten*. Coutinho.



## REFERENCES

- KNOOP, A. & K. SCHUIRINGA 1998. *Door allen voor allen, een heerlijk streven! Een kleine geschiedenis van het kruiswerk in Nederland*. Nederlands Openluchtmuseum.
- VAN LIEBURG, M. 1999. *In het belang van wetenschap en kunst: Een beknopte geschiedenis van de koninklijke nederlandse maatschappij tot bevordering der geneeskunst*. Erasmus Publishing.
- MOOIJ, A. 1999. *De polsslag van de stad : 350 jaar academische geneeskunde in Amsterdam*. De Arbeiderspers.
- QUERIDO, A. 1965. *Een Eeuw Staatstoezicht op de Volksgezondheid*. Staatsuitgeverij.
- VAN RAAK, R. 2010, forthcoming. Re-defining health systems from the perspective of transition theory. In *Transitions in health systems*, edited by J. Broers & J. Bunders.
- SCHOUWSTRA, C. 1995. *Over de invloed van wet-en regelgeving op de ontwikkeling van de huisartsgeneeskunde*, chapter 6, pages 67–78. Erasmus Publishing.
- SCHUT, J. 1970. *Van dolhuys tot psychiatriach centrum : Ontwikkeling en functie*. De Toorts, Haarlem.
- SPIJKER, T. 1979. *Mooi en Beschaafd Verplegen*, volume 26 of *Verpleegkunde Studies*. De Tijdstroom.
- VERHOOG, J. M. 1998. *75 jaar Organon : 1923-1998*. Organon.
- VIJSELAAR, J. 1982. *Krankzinnigen Gesticht: Psychiatrische Inrichtingen in Nederland, 1880-1910*. Fibula-Van Dishoeck.
- ZWOLS, G. 1985. *Het ontstaan en de ontwikkeling van het beroep verpleegkundige*, volume 55 of *Verpleegkundige Studies*. De Tijdstroom.

## Building 2

# Computational model: “Building a Model on Pillars”<sup>1</sup>

In this chapter a simulation model is presented which is an a direct implementation of the theory developed in this thesis. A simulation entails letting a virtual societal system develop over time according to the conditions and patterns of transitional change. The virtual societal systems used in this chapter are not representations of existing or historical ones, rather they are general examples used to explore the dynamics of the model. This is because the principal aim of this modelling exercise is to ascertain if the theory is consistent and comprehensive enough to be a basis for modelling approaches. For which aim the model was constructed as a *Java*-programme and verified to show plausible transition dynamics.

---

<sup>1</sup>This chapter is based on the authors contribution to the 6<sup>th</sup> European Social Simulation Association Conference (ESSA 2009), held September 14–18 in Guildford, UK at the University of Surrey.

## 2.1 Introduction

This chapter builds forth from the formalisation in chapter 4 of book II. It presents a simulation model implementing the theory as directly and fully as reasonably and practically possible. Most often a simulation model is used to make predictions, do forecasting or to reproduce empirical data. This is not the case for this simulation model. The principal aim of this exercise is to ascertain whether the theory is consistent and comprehensive enough in itself to be a basis for simulation models of societal transitions. A straightforward proof of this assertion would consist of two components. First, the running model itself. Second, that it produces plausible transition dynamics.

The simulation model presented in this chapter, therefore, is not run with initial conditions representing any real-life transition case. Rather the model is verified against various possible initial conditions, that is various possible societal systems. Validation against empirical data was considered outside the scope of this exercise.

Because the model employs the concepts of the theory, verification, in the sense of the model producing plausible transition dynamics, entails simply to check if transitions occur when conditions are present. This would mean that the functioning of the modelled societal system has changed radically, that is substantially and abruptly. In terms of the concepts of the theory this implies that either that a new regime has come up or that the regime itself transformed.

## 2.2 Outline of the model

Since the motivation for this modelling exercise was a direct implementation of the theory, the structure of the model is also along these lines. In short, the theory has concepts to describe the composition and state of a societal system. From this the conditions for transitional change, tension, stress and pressure can be found. Transitional change comes in the form of the patterns empowerment, reconstellation and adaptation. These patterns are driven by the conditions. Thus, the model allows to build a virtual societal system, has methods to calculate the conditions and methods that alter the composition and state of the system when conditions are present.

### 2.2.1 Representation of societal systems

The theory was implemented in the model as a *Java*-programme. The programme provides classes for central concepts like constellations, landscape, support and societal innovations. By creating instances of these classes one builds up a virtual societal system. One special class, called composition, provides the methods that change the state of the societal system, such as changing, adding or removing constellations or innovations. An instance of the class composition keeps track of the composition of the virtual societal system. With this object the bookkeeping is done of the changes resulting from the aforementioned methods.

Following the theory like formalised in chapter 4 of book II the functioning of constellations or its counterpart in related concepts as the landscape or societal support is represented by a vector. Thus, each entity in the modelled societal

## 2.2. OUTLINE OF THE MODEL

system carries a functioning vector. In the case of this model the functioning vectors have *real* and positive facets. The number of facets is limited to two for simplicity, although the model allows arbitrary dimensions. The facets are kept positive to avoid cumbersome interpretation of what would be ‘negative’ functioning. The power of a constellation is again taken as the length of the functioning vector. The angle between two functioning vectors is a measure of how different they are. Since the facets are always positive no angle is greater than  $\frac{\pi}{2}$ .

### 2.2.2 Conditions for transitional change

When the functioning of the constellation, innovations, support and the landscape is known in terms of their vectors, the conditions for transitional change, the tension, stress and pressure can be straightforwardly calculated using the following equations<sup>2</sup>:

Tension:

$$T \equiv \frac{\mathbf{S} \cdot \mathbf{L}}{\Pi}. \quad (2.1)$$

Stress:

$$S \equiv \frac{(\bar{\mathbf{S}} - \mathbf{S})^2}{\Pi}. \quad (2.2)$$

Pressure:

$$P \equiv \frac{\mathbf{R} \cdot (\mathbf{S} - \mathbf{R})}{\pi_R}. \quad (2.3)$$

Note that for tension and stress the values for the *system* are used, whereas for pressure the pressure on the *regime* is taken. The regime is the, at a given time, most powerful constellation in the societal system.

### 2.2.3 Patterns of transitional change

The above conditions drive the transitional change in the societal system. Transitional change comes in the form of the patterns: reconstellation, adaptation and empowerment are simulated in the model<sup>3</sup> using a static class. This class contains methods that work on the societal system as represented by an instance of the composition class, which also does the bookkeeping as the system evolves, keeping track of changes.

#### Reconstellation

The pattern of reconstellation, whereby a constellation of alternative functioning emerges is implemented by methods that either make an existing constellation more powerful or transforming an societal innovation into a constellation. In both cases this represents the idea that a certain existing societal functioning is implemented — probably through some force outside the system — rather than that it develops from within the system. Making a constellation more powerful corresponds for instance to the case where a government implements some niche

---

<sup>2</sup>See chapter 4 of book II, for notation and background concerning these equations.

<sup>3</sup>In a sense the patterns are both the mechanisms producing the transitional change as their outcomes. This is what one gets when a theory is a basis for modelling as well as narrative use.

## *BUILDING 2. "BUILDING A MODEL ON PILLARS"*

technology in a large infrastructure or makes some small scale societal practice the norm. Reconstellating a societal innovation is similar except that it would correspond to the case where there was not yet a constellation as such, with that functioning, present in the system.

The methods add an amount of 'driving conditions' of power to the constellation. The idea behind this is that the conditions — measured in terms of power even — provide a sort of potential for power that can flow into the system via the workings of the patterns. In the same way power is added to innovations after which they are made constellations. Thereby their potential power becomes actual power and their functioning becomes part of the functioning of the societal system.

When the conditions are met the most deviant constellation or innovation is chosen to be reconstellated. Where most deviant is understood as differing the most in terms of functioning with the system as a whole. This is tested for by finding the constellation having the largest angle with the system. If both a constellation and an innovation are available the most (potentially) powerful one is chosen.

### **Empowerment**

The empowerment pattern also involves the emergence of a constellation with alternative functioning. In contrast with reconstellation this is the result of developments of constellations within the system and with societal innovations, thus becoming more powerful while changing. In the model this is cast in the form of a methods that take a constellation and merge it with another one or with an innovation and subsequently make it more powerful, again by adding the available potential power of the driving conditions. The merging, which is simple vector addition of the functioning, produces a new constellation with an altered functioning, which is then as said, empowered.

When the conditions are right for empowerment the most deviant constellation is chosen and merged with a constellation that is close to it, that is, has a small angle with it. Or, it is merged with the most deviant (with respect to the entire system) innovation, if such an innovation is present and potentially more powerful than the constellation that would otherwise be chosen.

### **Adaptation**

The adaptation patterns is at the other side of the balance of transitional change. It describes how the dominant constellation innovates itself to remain just that, while simultaneously responding to the conditions for transitional change. This is done by taking over functioning from or absorbing other constellations, which is modelled by methods that let the most powerful constellation, the regime, merge with the constellation or innovation that is least deviant from it. This to reflect that a regime if it is willing to innovate at all, it will do so in the least invasive way<sup>4</sup>. If a constellation and an innovation are present that are suitable to adapt the regime with the most (potentially) powerful one.

Note that the adaptation pattern reacts to conditions but does not add power to any constellation. This also implies that in principle the regime could

---

<sup>4</sup>A common notion in innovation literature as well as in transition studies, see for example (Nelson & Winter, 1977; Dosi, 1982) or (Rip & Kemp, 1998; Kemp, 1994; ?).

## 2.2. OUTLINE OF THE MODEL

become less powerful after adaptation if a constellation or innovation with which it merges would have corresponding facets of opposite sign. This is however not considered in this model.

### 2.2.4 How conditions drive patterns

Through the workings of the patterns as described above the system state is altered. Constellations change their functioning, some disappear as they become absorbed by or merge with other ones and new constellations appear when innovations are reconstellated. This results in changes in the systems composition and its total functioning. Consequently the conditions for transitional change are altered and this keeps the dynamics going.

When a simulation runs the model checks what the conditions are, it finds the values of the tension, stress and pressure and subsequently drives the patterns. This is done according to the so-called coupling hypotheses of the theory. These couplings are deliberately kept as separated as possible from the rest of the body of the theory since they are hypotheses that need to be tested against empirical data. Therefore it is convenient if they can be adjusted and studied while keeping the rest of the theory constant so to speak. In this way the hypotheses can be the parameters of their very theory. This is also the reason why they are discussed only just now. In any case the theory currently proposes the following scheme to couple the conditions to the patterns:

| Condition |   | Pattern         |
|-----------|---|-----------------|
| Tension   | → | Reconstellation |
|           | ↘ |                 |
|           | ↗ |                 |
| Stress    | → | Adaptation      |
|           | ↘ |                 |
|           | ↗ |                 |
| Pressure  | → | Empowerment     |

Table 2.1: Hypotheses on which conditions drive which pattern

### 2.2.5 Additional mechanisms and parameters

The patterns almost always involve changing constellations by merging them, either with innovations or with other constellations. This implies that almost each cycle a constellation or innovation is merged with another one and thus ‘disappears’ from the system. Since this produces rather boring dynamics, after a number of cycles one constellation is left and that is it, some mechanism must be implemented to compensate for this. Which was that with each cycle the model checks if the number of constellations has dropped below the initial number and if so introduces a new one. Similarly the number of innovations available is kept more or less constant and approximately equal to the number of constellations.

## BUILDING 2. “BUILDING A MODEL ON PILLARS”

This raises the question of what properties such a newly introduced constellation or innovation ought to have. Should it be one with ‘random’ functioning, or should its functioning depend somehow on the current composition of the system. In the latter case one could think of a constellation or innovation that bridges a gap between the societal need and the current functioning of the system for instance. However, the empirical observation was that introducing new constellations or innovations via such a scheme did not qualitatively alter the dynamics of the system. And to minimise theoretical bias in this part of the model — it being a sort of correction mechanism after all — and to keep things as simple as possible, new constellations and innovations were introduced with random functioning. Random meaning a functioning where the facets were drawn from a random generator and the (potential) power scaled to somewhat below the average power of the constellations or innovations present.

At various places it would have been possible to introduce scaling parameters to allow some fine tuning. For instance in the driving of patterns by the conditions. The power received through empowerment could be scaled by some ‘empowerment factor’. To keep the model as conceptually clean as possible and since qualitative verification of the model dynamics was the principal aim, no such scaling or fine tuning was done. The only degrees of freedom left are the initial conditions, set by supplying an initial societal system, with a landscape and support vector.

### 2.2.6 Simulation runs

Running a simulation would entail the following sequence:

1. An initial societal system is provided by the user. This is done with an initialisation file supplying the landscape, support, constellation and innovation vectors.
2. The programme calculates the tension, stress and pressure from the composition of the virtual societal system.
3. The programme invokes the patterns, changing the state of the societal system. Thus new constellations or innovations enter the system, present ones are changed or disappear.

Consequently the conditions for transitional change have changed and the programme continues to calculate the new tension, stress and pressure and invokes the patterns accordingly. The cycle 2–3 is repeated as many times as desired.

## 2.3 Model behaviour

To explore the behaviour of the model its dynamics has to be compared as function of the available degrees of freedom. As referred to earlier the model hardly any parameters to fine-tune and the main input variables in fact simply are the initial conditions. At the beginning of a simulation an initial societal system is fed into the model. This sets the following variables:

*Tension* (via the landscape vector), *stress* (via the support vector) and the *composition of the system* (via the constellation vectors). This last *en passant*

### 2.3. MODEL BEHAVIOUR

also defines the functioning of the entire system, the distribution of power over the functioning and the initial *pressure*.

At this point the questions that needed to be addressed with the model were if it produced ‘plausible’ transition dynamics and how this dynamics depended on the various theoretical and modelling parameters. The strategy employed to understand the behaviour of the model with respect to these questions was the following:

Prepare some initial distributions of constellations, that is some ‘testing’ societal systems and let them evolve in a simulation under various conditions: no tension and no stress, high tension no stress, tension and stress and relevant permutations thereof. The initial distributions explored were the following:

- *Flat*  
A number (typically 5 or 6) of constellations of equal power distributed equally over the system. See figure 2.1, for a representation of this distribution in terms of the angles the functioning makes with some reference angle, and the power of the constellations as percentage of the total<sup>5</sup>.
- *Skew*  
Similar to flat but with all constellations in a restricted part of the functioning space. See figure 2.2.
- *Regime-niches*  
One constellation of significantly larger power than the rest, that are distributed ‘flat’. This corresponds to how societal systems are often made caricatures in transition studies. See figure 2.3.
- *Regime-niches skew*  
As regime-niches but ‘skewly’ distributed. See figure 2.4.
- *Random*  
Several randomly generated societal systems were used as well.

These distributions were then run in the model through some thousand cycles, though the dynamics settled much earlier. Running a simulation with the same initial distribution yielded different results, though the overall dynamics remained similar (more on this later). Clearly the model is, and models, a complex system and exhibits the appropriate limited-predictable dynamics without being flat-out chaotic.

Typically after a couple of dozen cycles reaches a more or less stable distribution, or maintains a stable oscillation between a small number of distributions. See figures 2.5 and 2.6 for some typical runs that show the angle of the functioning of the entire system versus ‘time’. Note that these pictures have been smoothed with a B-spline interpolation of 5<sup>th</sup> order, to keep the sometimes wild fluctuations that accompany a pattern working from obscuring the general trend one is interested in.

---

<sup>5</sup>In the figures of the distributions the system as a whole is present as well with a power percentage. This percentage is not 100% as one might intuitively think, this is because of vector addition the power of the sum is not the same as the sum of the powers. The percentage the total system has is a measure of how spread-out the constellations are. If the angles between the constellations are all small, the power percentage of the entire system will be close to 100% if the spread is maximal it would be close to 71% ( $\frac{\sqrt{2}}{2}$  times 100%).



## BUILDING 2. "BUILDING A MODEL ON PILLARS"

One might expect that a continuous spectrum of distributions results for the end states of simulation runs, contrarily it appears that there is a limited number of typical end states. These typical final distributions can be qualitatively described as follows:

### Regime and niches

Similar to the regime and niches initial distribution. One constellation several factors greater in power than the other constellations that are pretty much equal in power and more or less randomly distributed over the system. See figure 2.7.

### Regime, niche-regime(s) and niches

One constellation significantly more powerful than the rest, although one or a couple 'niche-regimes' are present, approximately half as powerful. The rest are flatly distributed niches. Power was typically distributed something like 60-28-4-4-4%. See figure 2.8.

### Niches

Like the flat initial distribution. A number of niches of approximately equal power and functioning. See figure 2.9.

Some notes on this typical end distributions. The 'niches' and 'regime and niches' end distributions resemble some of the used initial distributions, this is however coincidence in as much as these end distributions could be obtained from all initial distributions under the right conditions, more on this later. Moreover when, say, a niches end state resulted from a niches initial state, it was still a very much changed system. As can be clearly seen in figure 2.9 where all niches are very close together in terms of functioning as well, resulting in a 100% proportion of power for the entire system, this was common for the niches end distribution.

This brings up an interesting characteristic of the transition dynamics as exhibited by this model. It so appears that the dynamics and consequently the end distribution of the societal system is determined much more by the *conditions* that is, the present tension, stress and pressure, than the composition. This is a rather central result of this modelling exercise and it deserves some more explication.

For any composition it turned out that applying no tension and no stress, which is easily done by setting the landscape vector to zero and the support vector equal to the vector of the total system, the end distribution was that of the niches. If tension was the dominant condition this led mostly to the regime, niche-regime and niches distribution and letting stress dominate drove the system to the regime and niches distribution. Running a simulation, with the same conditions several times yielded different end states in the quantitative sense, but in almost all cases, say more than 75%, to the same qualitative end distribution. A relation between the quantitative end state and the conditions nor the initial composition could be ascertained. That is, it appeared impossible to predict how much the entire system would shift in functioning as a function of the initial distribution nor the conditions. It must also be noted that although the conditions drive the dynamics of the system and apparently largely determine its end state they are not necessarily smaller if a stable distribution is reached.

## 2.4. CONCLUSIONS AND DISCUSSION

The influence of the composition of the system could be appreciated mostly in terms of the pressure it produced. Pressure being the influence of the other constellation on the most powerful one (or any one, in case of flat distributions), then can be said to drive the system to the niches distribution. Dominance of a certain condition can be stated in quantitative terms, a condition started dominating another already if it was greater by a factor of two. If two conditions were of approximately equally dominant, the end states also fluctuated between the corresponding conditions. Note: the end states were not in between as in average, but rather one run chose one state the next the other.

A transition was a radical change of functioning of the societal system. In transition studies this is either equated with a change of regime, which signifies fundamental change in functioning since it is the most powerful constellation, or with a change in functioning of the regime or other constellations. As far as the transition dynamics in this model is concerned, both can be observed and even within single simulation runs.

## 2.4 Conclusions and discussion

The introduction to this chapter spoke of a threefold aim for the chapter and the modelling exercise it describes. The first aim was to ascertain if the theory which is at the basis of this all, pillar theory, lives up to its ambition to be just that, a basis for a modelling approach. Apparently it is, casting the theoretically proposed mechanisms in the form of a computer programme working with and on the concepts of the theory resulted in a working simulation model. Moreover the model proved no need of a barrage of parameters to fine tune the dynamics. All mechanisms are the most simple and straightforward implementations possible.

This being said the next aim becomes all the more relevant, does the model produce plausible transition dynamics. This is impossible to answer without leaving some ambiguity. This has to do with the level of abstraction at which the theory operates and which consequently leads to an equal or higher level of abstraction in a formal form like a model. In this bare form the model could however already produce curves of changing functioning over time (like 2.5 and 2.6) that are very much akin to the s-shaped curves often referred to when speaking of transition dynamics. Regime changes were also easily observed.

The third aim puts this in perspective. Exploring how the dynamics of this model responded to the parameters resulted in the observation that a limited number of qualitatively different end states are typically produced by the model. Which end distribution was obtained could be clearly linked to the applied conditions, which is satisfactory in a modelling sense as well as in a theoretical sense.

Although speculative this does invite some societal interpretation, since the typical end distributions as shown in figures 2.7, 2.8 and 2.9 can be seen as stylised examples of existing societal systems. The regime-niches distribution for example is clear example of how transition studies view a societal system: one regime dictating the societal business as usual and several small scale niches filling in the societal gaps. Think for example of how fossil-fuels dominate the energy supply systems in most of the Western world. The regime, niche-regime and niches distribution also is also archetypical and probably best illustrated by the example of a mobility system where the regime is the constellation around

## BUILDING 2. "BUILDING A MODEL ON PILLARS"

car-based individual mobility and the niche-regime is public transportation, relevant and relatively powerful but not dominant, this with niches like slow modes of transport and green solutions. The niches distribution would correspond to societal systems with properly functioning free markets, think for instance of (tele)communication with many solutions ranging from mobile telephony to internet, postal services, etc. functioning next to each other without one clearly dominating. Although the niches end distributions feature niches that have almost similar functioning, something that requires further investigation to see if this phenomenon resolves when more than two facets are used to define functioning vectors. The observation that tension, stress and pressure do not necessarily lessen in the course of a transition might lead to more sombre societal interpretations if the conditions are interpreted as a 'negative' system state.

This kind of interpretation is also a key to further work on modelling in this fashion. Since it is the interplay between theory development, empirical research and modelling approaches that is the leading thought behind this model and the pillars on which is built, further work must further establish this link. For the model this means that more exploration is necessary between theoretically proposed mechanisms and dynamics and the modelling counterparts thereof. One thinks for example of investigating whether it is possible to reproduce and better understand transition paths like proposed by de Haan & Rotmans (forthcoming in 2010) and Geels & Schot (2007). It also suggests making the model more suitable for dealing with empirical data, be it for historical or ongoing transitions. This would entail increasing the number of facets used to represent functioning from two to whatever value necessary to adequately represent a societal system of choice. The model is written for functioning represented by vectors with an arbitrary number of facets but the behaviour has only been studied for two. Also it becomes necessary to have a more clear interpretation of time, since it is now simply unclear what the time span is that a cycle represents, nor even if each cycle represents an equal amount of time.

Other directions for further research could be to employ the model to investigate how interventions in societal systems impact the societal dynamics, like transition management (e.g. Loorbach, 2007), strategic niche management (for example Kemp *et al.*, 1998) or other more traditional policy approaches. In the light of modelling as a policy tool a further developed version could be used in scenario exercises for transition processes which are also seen as instrumental in transition management (Sondeijker *et al.*, 2006).

## 2.5 Plots and graphs

### 2.5.1 Initial distributions

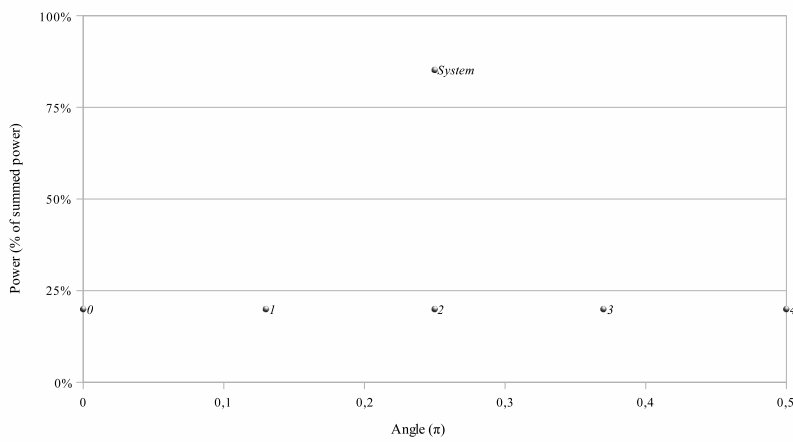


Figure 2.1: Flat initial distribution

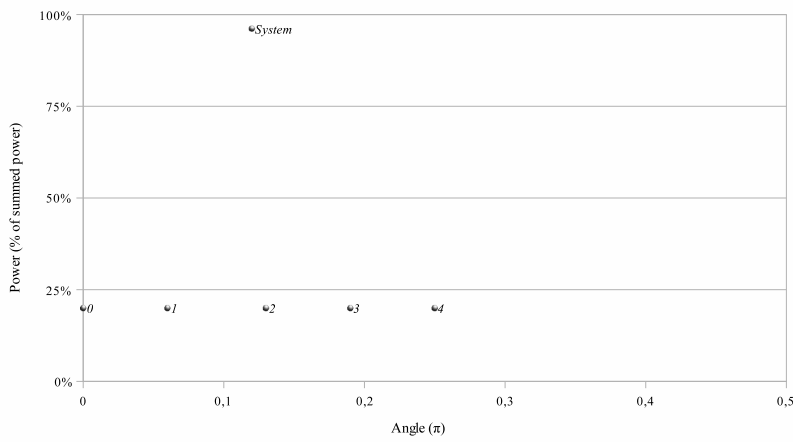


Figure 2.2: Skew initial distribution

BUILDING 2. "BUILDING A MODEL ON PILLARS"

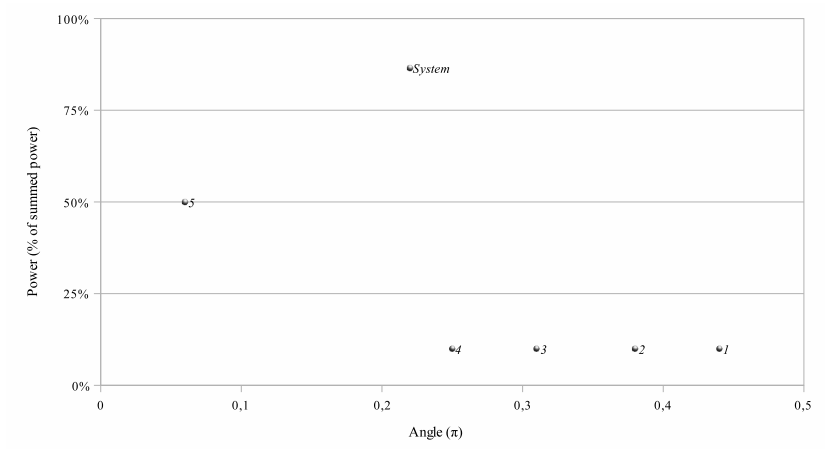


Figure 2.3: Regime-niches initial distribution

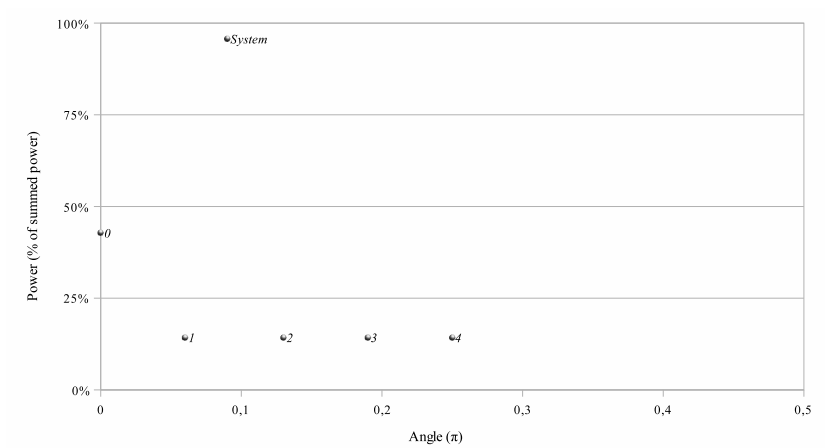


Figure 2.4: Regime-niches skew initial distribution

### 2.5.2 Typical runs

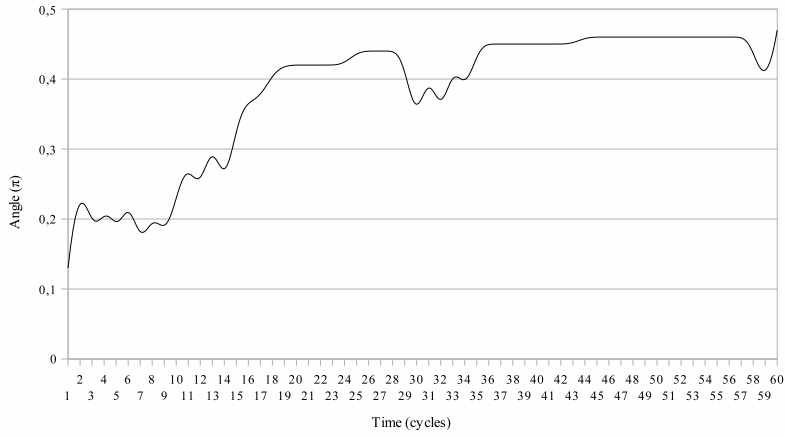


Figure 2.5: A typical run

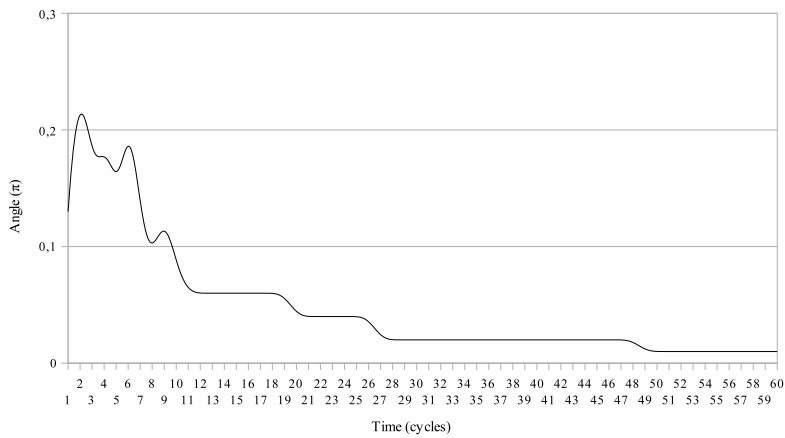


Figure 2.6: Another typical run

BUILDING 2. "BUILDING A MODEL ON PILLARS"

2.5.3 Typical end distributions

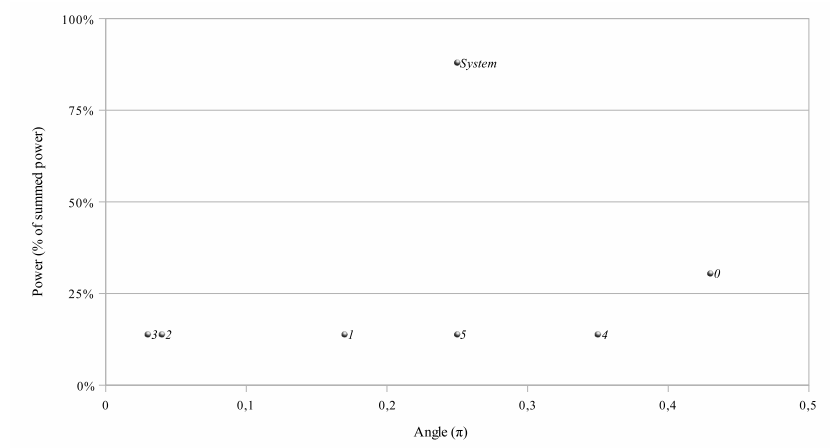


Figure 2.7: A regime and niches end distribution

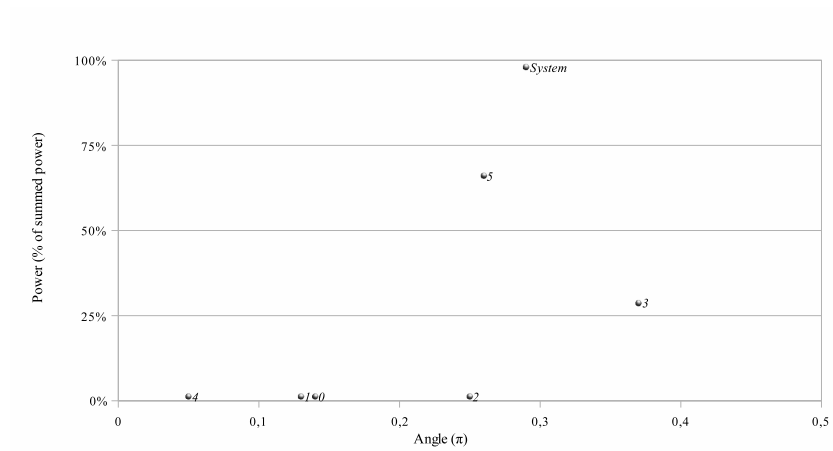


Figure 2.8: A regime, niche-regime and niches end distribution

## 2.5. PLOTS AND GRAPHS

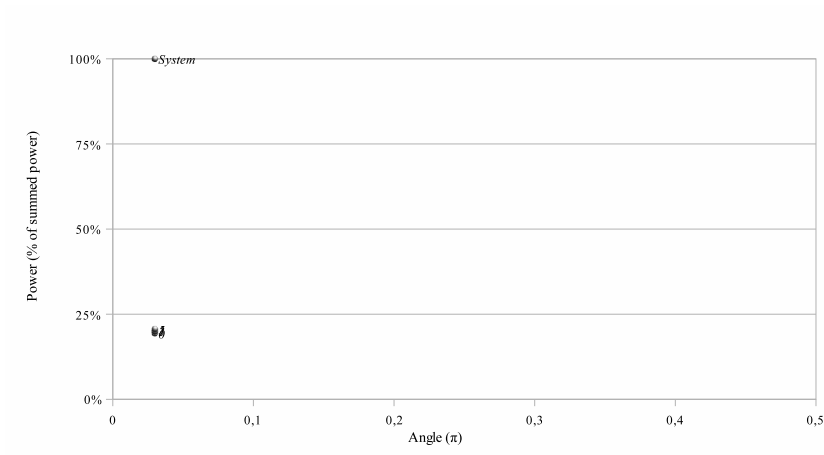


Figure 2.9: A niches end distribution



*BUILDING 2. "BUILDING A MODEL ON PILLARS"*

# References

- DOSI, G. 1982. Technological paradigms and technological trajectories: A suggested interpretation of the determinants and directions of technical change. *Research Policy*, 11: 147–162.
- GEELS, F. W. & J. SCHOT 2007. Typology of sociotechnical transition pathways. *Research Policy*, 36: 399–417.
- DE HAAN, J. & J. ROTMANS forthcoming in 2010. Patterns in transitions: Understanding complex chains of change. *Submitted to Technological Forecasting and Social Change*.
- KEMP, R. 1994. Technology and the transition to environmental sustainability. the problem of technological regime shifts. *Futures*, 26(10): 1023–1046.
- KEMP, R., J. SCHOT, & R. HOOGMA 1998. Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management. *Technology Analysis and Strategic Management*, 10(2): 175–195.
- LOORBACH, D. 2007. *Transition Management: New Mode of Governance for Sustainable Development*. International Books, Utrecht.
- NELSON, R. R. & S. G. WINTER 1977. In search of a useful theory of innovation. *Research Policy*, 6: 36–76.
- RIP, A. & R. KEMP 1998. *Technological Change*, volume Volume 2, pages 327–399. Battelle Press, Columbus, Ohio.
- SONDEIJKER, S., J. GEURTS, J. ROTMANS, & A. TUKKER 2006. Imagining sustainability: The added value of transition scenarios in transition management. *Futures*, 8(5): 15–30.

*REFERENCES*

## Building 3

# Mathematical model: “The Dynamics of Functioning”<sup>1</sup>

In this chapter a mathematical framework is introduced and explored for the study of processes in societal transitions. A transition is conceptualised as a fundamental shift in the functioning of a societal system. The framework views functioning as a real-valued field defined upon a real variable. The initial status quo prior to a transition is captured in a field called the regime and the alternative that possibly takes over is represented in a field called a niche. Think for example of a transition in an energy supply system, where the regime could be centrally produced, fossil fuel based energy supply and a niche decentralised renewable energy production. The chapter then proceeds to translate theoretical notions on the interactions and dynamics of regimes and niches from transition literature into the language of this framework. This is subsequently elaborated in some simple models and studied analytically or by means of computer simulation.

---

<sup>1</sup>This chapter is based on an article by the author, published as de Haan (2008) “The Dynamics of Functioning: Investigating Societal Transitions With Partial Differential Equations” in *Computational and Mathematical Organization Theory*, volume 14, pages 302–319.

### 3.1 Introduction

This chapter builds directly from the language part, that is, chapter 2 of book II. Rather than taking the dynamics of an entire societal system as object of study, here the dynamics of the functioning of constellations is studied, internally as well as in interaction with another constellation or societal support. In transition studies it is not uncommon to describe this in terms of a regime change. The new regime is then thought of as growing out of a niche, a nucleus of innovative or otherwise different functioning. How and under what conditions such a niche scales up and replaces the incumbent regime is therefore an important matter in a societal transition.

Since the regime by definition dominates the functioning of the societal system, and as such controls the infrastructures and determines the current discourse, it is a constellation with power. Compared to the regime, niches function on smaller scales, presumably with novel approaches and experimental technology. The lack of power of a niche can be partially compensated by societal support. For instance, financial support might be injected in the form of subsidies, political support could be given in the form of benign legislation or there might simply be a small market for the niche already.

The aim of this chapter is to make it possible to investigate — in a mathematical way — questions like: How does support influence the up scaling of niches? Or: Does the functioning or regimes and niches change if there is competition on one of the several dimensions? Is a niche more likely to scale up if it functions in a way similar to the regime, or just the opposite? The reason to investigate those questions mathematically is that it allows one to rigorously track the consequences of certain hypotheses. Although the elaborations might be complicated, the assumptions can be directly tested for their consequences. Another advantage of a mathematical approach is that any problem, once cast in equations becomes comparable to other problems that share the same mathematical form, although the original questions had nothing in common.

So now it becomes necessary to represent regimes and niches in some mathematical form. Apparently regimes and niches are constellation with properties in possibly various dimensions, like for instance ‘greenness’, ‘efficiency’ or simply ‘price’. Let such dimensions be parameterised by some real parameter  $\varphi$ , and for ease of exposition let the number of dimensions be limited to one. For example a regime or niche can be thought of as taking a position on the  $\varphi$ -axis running from fossil fuel only to renewable or bio fuel only. Again taking the initial example the regime would be distributed on the ‘fossil side’ of the  $\varphi$ -axis and a bio fuel niche on the other, ‘renewable’ side. This is of course a caricature, but in several dimensions such a characterisation is not unthinkable.

This has similarities with the spatial approaches in political science, e.g. (Downs, 1957; Kollman *et al.*, 1992; Laver, 2005). In those approaches a political party has coordinates along several political axes, like left-right, conservative-liberal, etc. In this approach for societal transitions one can take this a step further. Instead of assigning a regime or niche to a point on the  $\varphi$ -axis, the constellation becomes a distribution over that axis. In other words, regimes and niches become functions of  $\varphi$ , or in yet other words, they become fields defined on a functioning space. This has the extra advantage of allowing the surface area under these functions to represent the power of the regime or niche. To illustrate this again with the energy example, the regime would be a large

blob on the fossil side and a bio fuel niche a small blob on the renewable side.

Somewhat more on this interpretation of power maybe. If the surface area under a regime or niche function represents its power, then the value of its function at a certain point also has an interpretation, namely it is the specific power of the constellation over a specific form of functioning. One can easily imagine that this interpretation becomes relevant when a niche is in competition with a regime. Still more on power. Apparently the sum total of the surface areas under the regime and niches is the total power present in the societal system. If one considers the representation of the regimes and niches as complete, that is all the necessary dimensions are taken into account, then one can defend that the total power in the societal system is constant and merely redistributed over the several constellations. This implies a conservation law which can be quite instrumental<sup>2</sup>.

The chapter will now continue to expose the formalism and how hypotheses on transition dynamics are cast in this formalism. Clearly, when all the concepts are cast in the form of distributions evolving and interacting, the mathematics will be that of coupled partial differential equations (pde's). For clarity and mathematical tractability the number of dimensions in this chapter is limited to one, although there is in principle no restriction there. The merits and challenges of the framework will be discussed at length at the end of this chapter, suffice it to say that methodologically the approach is akin to that of pattern formation in physics<sup>3</sup>, chemistry, biology and ecology, but also to the classical field theories of physics (Landau & Lifshitz, 1951), which opens a wide body of knowledge for the emerging field of transition science. The remainder of the chapter will elaborate some simple consequences of the framework and illustrate its usefulness in several simple model settings.

## 3.2 Formalism

The introduction already suggested that the form the regime and niches will take would be that of functions of  $\varphi$ . The interpretation of surface area as power is easier if one demands that the functions be positive, which also avoids difficult interpretation of what would be negative functioning. Therefore, the regime, niches and possible other constellations are represented by fields defined upon the specific functioning  $\varphi$  and time  $t$  with values in  $\mathbb{R}^+$ . Where  $\varphi$  is a real parameter on some domain  $\Phi \subset \mathbb{R}$  and  $t \in [0, \infty)$ . This invites the following notation for the regime and niches:

The regime:

$$R(\varphi, t) \rightarrow \mathbb{R}^+. \quad (3.1)$$

A niche:

$$n(\varphi, t) \rightarrow \mathbb{R}^+. \quad (3.2)$$

It will sometimes prove to be convenient to speak of an arbitrary or general constellation (regime, niche)  $i$ , which will then be denoted as  $c_i(\varphi, t)$ . The

---

<sup>2</sup>This is a so-called zero-sum interpretation of power, described and criticised by Parsons already. For this and other debates on power see e.g. (Haugaard, 2002). For a more complete review of power in the context of transition studies see (Avelino & Rotmans, 2009; Avelino, 2009)

<sup>3</sup>The field is vast and ever growing, some of the authors favorites are (Cross & Hohenberg, 1993; van Saarloos, 2003)

### BUILDING 3. “THE DYNAMICS OF FUNCTIONING”

functioning of the societal system as a whole was already defined as the sum of the functioning of all its constellations, which then is obviously  $\sum_i c_i(\varphi, t)$ .

The fraction of the functioning of the societal system that a certain constellation contributes to the whole is an interpretation of its power. This leads to the following definition of the power  $\pi_i$  of a constellation  $i$ :

$$\pi_i = \int_{\Phi} d\varphi c_i(\varphi, t) \quad (3.3)$$

Now that power is defined in terms of functioning the it is possible to postulate the conservation of power as follows<sup>4</sup>:

$$\partial_t \int_{\Phi} d\varphi \sum_i c_i(\varphi, t) = \partial_t \sum_i \pi_i = \partial_t \Pi = 0 \quad (3.4)$$

implicitly defining  $\Pi$  as the total power.

It is obvious that the more two constellations are alike in functioning, the more their functioning will overlap. The power that is represented in this overlap is therefore a convenient measure of the likeness or closeness of niches and niches and the regime. The power overlap of two constellations  $c_i$  and  $c_j$  is

$$\pi_{ij} = \int_{\Phi} d\varphi c_i \theta(c_j - c_i) + c_j \theta(c_i - c_j), \quad (3.5)$$

where  $\theta(x)$  is the Heaviside function, unity for positive arguments and zero for negative arguments.

After these preliminary matters of definition and notation it is now possible to discuss how regimes and niches change in time. In other words, what is the form of the evolution equations for the constellations? One realises that a regime or niche, when left ‘alone’ will still have a time evolution caused by its internal dynamics. Constellations are after all societal constructs comprised of actor, institutions and such. Apart from the internal dynamics regimes and niches are *interacting*. Interaction can be the result of competition, political influencing, using the same resources or infrastructures and much more. Interaction of course results in the coupling of the evolution equations. In any case, these two influences on the dynamics of constellations will be separated in the equations. So a general form for an arbitrary constellation  $i$  takes the form of:

$$\partial_t c_i(\varphi, t) = \mathcal{F}_i [c_i] + \mathcal{I}_i [c_i, c_j], \quad (3.6)$$

where  $\mathcal{F}_i$  is a differential operator representing the internal dynamics of constellation  $c_i$  itself,  $\mathcal{I}_i$  with  $j \neq i$  is the interaction term that in general will depend on some or all of the other constellations.

The conservation of power leads to the idea that what one constellation gains in power must be lost by the others. This demand can be a direct consequence of the forms of  $\mathcal{F}_i$  and  $\mathcal{I}_i$ . If this is not the case balancing terms  $\mathcal{B}_i$  need to be introduced that in turn will depend on  $\mathcal{F}_i$  and  $\mathcal{I}_i$ , like in

$$\partial_t c_i(\varphi, t) = \mathcal{F}_i [c_i] + \mathcal{I}_i [c_i, c_j] - \mathcal{B}_i [c_i, \mathcal{F}_j, \mathcal{I}_j]. \quad (3.7)$$

If no direct interactions are present the  $\mathcal{B}_i$  will then still provide an indirect coupling.

---

<sup>4</sup>For ease of notation the partial derivative with respect to time is denoted as  $\partial_t$ .

For example, one niche and one regime with interactions that need balancing give rise to a set of coupled pde's of the following form:

$$\begin{cases} \partial_t R(\varphi, t) = \mathcal{F}_R [R] + \mathcal{I}_R [R, n] - \mathcal{B}_R [R, \mathcal{F}_n, \mathcal{I}_n] \\ \partial_t n(\varphi, t) = \mathcal{F}_n [n] + \mathcal{I}_n [n, R] - \mathcal{B}_n [n, \mathcal{F}_R, \mathcal{I}_R]. \end{cases} \quad (3.8)$$

One serious limitation of this approach is immediately clear from equations (3.8) is that a constellation is stuck to its  $\mathcal{F}$ . This means that niches can scale up but never become regimes within the frame of the equations and neither can a regime be dethroned in this sense. This *a priori* limits the approach to the initial stages of transitions or the analysis needs to be done in several phases.

### 3.2.1 Internal constellation dynamics

In times where little societal change is occurring the functioning of the societal system still evolves. Thus, each constellation, regime or niche, when left on its own will still evolve over time. Describing this evolution entails translating some theoretical knowledge and assumptions in mathematical terms.

One realisation is that all constellations, be they regime or niches, are expected to smoothen out the different aspects of their functioning. It is, for example, intuitively clear that a certain niche will next to its 'core-business' (say hydrogen energy storage) have an interest in related functioning (say solar energy capture). Moreover a constellation will try to integrate the various aspects of its functioning in a smooth, self-consistent whole. Much literature on regime and niche development elaborate on this development of such constellations to e.g. a self-consistent set of rules or dominant designs like (Nelson & Winter, 1977; Kemp *et al.*, 1998). That structuration and striving for self-consistency is also a more general realisation for soci(et)al systems is discussed at length in (Giddens, 1984; Luhmann, 1984), where in the abstract is discussed how action and structures are recursively shaping each other and how social systems develop a tendency to reproduce themselves.

This kind of dynamics appears to correspond to a diffusive behaviour of constellations, that is the  $\mathcal{F}_i$  are expected to have a form like

$$\mathcal{F}_i = \beta_i \partial_\varphi^2 c_i. \quad (3.9)$$

It is no coincidence that innovation literature also speaks of 'diffusion of innovation' (Rogers, 2003), which is readily generalised from technological to societal innovation. Another realisation is that innovation is likely to be found at the 'edges' of the constellations. In other words if constellations attempt to make their functioning self-consistent and smooth, innovation is to be expected on the edges. This is often argued as a reason for the incremental character of innovation (Levinthal, 1998) and the necessity for (societal) innovation to emerge in niches<sup>5</sup> (Rotmans, 2005; Kemp *et al.*, 1998). This also is a mathematical characteristic of diffusive behaviour if the initial form of the constellation is sufficiently localised as any text on pde's can testify e.g. (Farlow, 1982).

Although both the regime and niches apparently exhibit diffusive behaviour in the evolution of their functioning they have quite different characters. One expects the regime to be a stable constellation. The regime for various reasons

<sup>5</sup>For a review of the niche concept in technological change literature see (Schot & Geels, 2007).



### BUILDING 3. “THE DYNAMICS OF FUNCTIONING”

already alluded to above — vested interests, technological lock-in, etc. — is expected to stick to, and optimise its current functioning. This is also recognised in other uses of the regime concept, like policy regimes (Wilson, 2000) and urban regime theory (Mossberger & Stoker, 2001). From niches, conversely, a more dynamical behaviour is expected. Since they are the loci of novel or deviant functioning, and therefore have larger tendency for experimentation, their functioning is expected to be more mobile and prone to spread out. This spreading out of the functioning of niches is referred to as ‘broadening’ by (van den Bosch & Taanman, 2006).

To distinguish these characteristics mathematically the diffusion constant becomes the relevant parameter. The diffusion constant, in this context referred to as the ‘broadening parameter’, is assumed to be small for a regime and relatively large for niches, in any case

$$\beta_R < \beta_n. \quad (3.10)$$

Different values of the parameter  $\beta$  can also be interpreted as the strategies of a constellation, similar to the strategies in spatial elections of (Laver, 2005). A small  $\beta$  would correspond to an ‘aggregator’ strategy and a large  $\beta$  to a ‘hunter’ strategy.

#### 3.2.2 Constellation interactions

The interaction terms are a different matter altogether, several views are possible of how constellations interact and how regimes are taken over in the course of transitions. Most theorising however does not explicitly speak of interaction between niches and regimes, except perhaps pillar theory (de Haan, 2007) and discuss the intrinsic growth of niches and regimes, again see e.g. (Geels, 2002; Kemp *et al.*, 1998). The transition in such views take place when the niche has become the new most powerful constellation, for instance in terms of market share.

In the following the distinction will be made between *direct* and *indirect* interaction. Direct interaction will have  $\mathcal{F}_i$ ’s featuring constellations other than  $i$ , whereas with indirect interaction everything is mediated by the balancing terms  $\mathcal{B}_i$ . One of the consequences of this is that the local properties of a regime or niche, that is the form of their functions, is hardly relevant in the indirect interaction cases.

##### Indirect interaction

A mathematical representation of indirect interaction is possible through the demand of conservation of total power. The regime has an intrinsic tendency to grow but stay close to its current forms of functioning. The latter was reflected in its small value of  $\beta_R$  the former translates to a growth term proportional to itself. That is, added to  $\mathcal{F}_R$  is a term proportional to  $R$  resulting in:

$$\mathcal{F}_R = \beta_R \partial_\varphi^2 R + \sigma_R R. \quad (3.11)$$

Where  $\sigma_R$  is some positive constant.

For a niche a similar reasoning holds. Its greater tendency for exploration was already captured by a larger value of  $\beta_n$ . A desire to grow can however not

### 3.2. FORMALISM

be captured by a simple self-proportional growth term like in the regime case. Especially strategic niche management but also pillar theory point out that niches are vulnerable and often need societal support. This support can take various forms, be they political in the form of helpful legislation or financial in the form of subsidies. For this it is convenient to introduce the support function (*soil* in (de Haan, 2007))

$$S(\varphi), \quad (3.12)$$

depending on  $\varphi$  to reflect that support for specific forms of functioning is possible. For a niche  $\mathcal{F}_n$  will according to the above have the following form:

$$\mathcal{F}_n = \beta_n \partial_\varphi^2 n + \sigma_n n S. \quad (3.13)$$

The term  $\sigma_n$  is the intrinsic potential for a niche to grow and the less a niche needs support the larger  $\sigma_n$  is expected to be. The growing of a niche will be called ‘scaling up’ following (van den Bosch & Taanman, 2006) and the  $\sigma_i$ ’s will be denoted as ‘upscaling parameters’.

An example societal system with one regime and a niche now takes the form of the following set of equations:

$$\begin{cases} \partial_t R(\varphi, t) = \beta_R \partial_\varphi^2 R + \sigma_R R - \mathcal{B}_R \\ \partial_t n(\varphi, t) = \beta_n \partial_\varphi^2 n + \sigma_n n S - \mathcal{B}_n. \end{cases} \quad (3.14)$$

The balancing terms  $\mathcal{B}_i$  necessary for the conservation of power and thus the implicit interaction can still take several forms. The demand  $\partial_t \Pi = 0$  is equivalent to

$$\int_{\Phi} d\varphi \left( \sum_i \mathcal{F}_i - \mathcal{B}_i \right) = 0 \quad (3.15)$$

because of equation (3.4) and with the  $\mathcal{I}_i c_i$  equal to zero. This leads to

$$\sum_i \left( \int_{\Phi} d\varphi \mathcal{F}_i - \int_{\Phi} d\varphi \mathcal{B}_i \right) = 0. \quad (3.16)$$

and

$$\sum_i \int_{\Phi} d\varphi \mathcal{B}_i = \sum_i \int_{\Phi} d\varphi \mathcal{F}_i. \quad (3.17)$$

For the simple system of equations (3.14) the following  $\mathcal{B}_i$  keep the total power conserved:

$$\mathcal{B}_1 = \mathcal{F}_2, \quad (3.18)$$

$$\mathcal{B}_1 = \frac{c_1}{\int_{\Phi} d\varphi c_1} \int_{\Phi} d\varphi \mathcal{F}_2, \quad (3.19)$$

Other terms are possible as well, obviously. The first form of the balancing term simply takes the power from the one constellation at the specific functioning where the other gains it, the second form collects the gain of the one and takes it from the other in proportion to its functioning. A disadvantage is that these  $\mathcal{B}_i$  do not necessarily keep the constellations positive.

**Direct interaction**

Perhaps a more realistic view on societal dynamics is that the constellations interact directly. This view entails a sort of ‘competition for power over specific functioning’. When a constellation expands its functioning in this picture there will be competition in the regions where it overlaps with another. The difference between niche and regime, in that the former depends partially on support and the latter can grow independently, is retained.

This results in interaction terms that are reminiscent of those of spatial predator-prey, bacterial growth and even some vortex-fluids interaction models (Baggio *et al.*, 2004). For instance, for the regime an  $\mathcal{I}_R$  is proposed of the form:

$$\mathcal{I}_R R = \sigma_R R n. \tag{3.20}$$

And for niches of the form:

$$\mathcal{I}_n n = \sigma_n n R (S - R). \tag{3.21}$$

Where the  $S - R$  term reflects the notion raised in pillar theory (de Haan, 2007) that niches emerge where the power exercised by the regime is low and support is present.

Since the interaction is direct and local, the obvious choice for the  $\mathcal{B}_i$  is the form (3.18). An example system of a regime and one niche now becomes:

$$\begin{cases} \partial_t R(\varphi, t) = \beta_R \partial_\varphi^2 R + \sigma_R R n - \sigma_n n R (S - R) \\ \partial_t n(\varphi, t) = \beta_n \partial_\varphi^2 n + \sigma_n n R (S - R) - \sigma_R R n. \end{cases} \tag{3.22}$$

It is also possible to subsume the  $\mathcal{B}_i$  in the interaction terms. If the support is interpreted as giving a niche — locally — more power, as is the case for instance with subsidies and legislative measures. An expanding niche then competes with the regime with an effective specific power  $nS$ . The straightforward implementation of this would be that the niche gains if  $nS > R$  and the regime if  $nS < R$ . The interaction term would look like  $\gamma n R (nS - R)$ . A simple example system in this case becomes:

$$\begin{cases} \partial_t R(\varphi, t) = \beta_R \partial_\varphi^2 R - \gamma n R (nS - R) \\ \partial_t n(\varphi, t) = \beta_n \partial_\varphi^2 n + \gamma n R (nS - R). \end{cases} \tag{3.23}$$

Where  $\gamma$  is a ‘phenomenological’ interaction parameter. It is more difficult to speak of upscaling parameters here, although  $\gamma S$  can be interpreted as an effective upscaling parameter.

### 3.3 Elaborations

#### 3.3.1 Simple consequences of conservation of power

The conservation of power in itself already has its consequences for the dynamics. Take for instance a very simple system consisting of two constellations, a regime and a niche with some support. The dynamical system becomes

$$\begin{aligned}\partial_t R(\varphi, t) &= \beta_R \partial_\varphi^2 R + \sigma_R R - \frac{R}{\int_{\Phi} d\varphi R} \int_{\Phi} d\varphi (\beta_n \partial_\varphi^2 n + \sigma_n n S), \\ \partial_t n(\varphi, t) &= \beta_n \partial_\varphi^2 n + \sigma_n n S - \frac{n}{\int_{\Phi} d\varphi n} \int_{\Phi} d\varphi (\beta_R \partial_\varphi^2 R + \sigma_R R).\end{aligned}\quad (3.24)$$

To explore the consequences of power conservation the explicit dependency on the specific functioning becomes irrelevant and if one integrates out the dependence on  $\varphi$  an expression in terms of the power of the constellation is obtained:

$$\begin{aligned}\partial_t \pi_R(t) &= \sigma_R \pi_R - \sigma_n \bar{s} \pi_n, \\ \partial_t \pi_n(t) &= \sigma_n \bar{s} \pi_n - \sigma_R \pi_R.\end{aligned}\quad (3.25)$$

It is assumed that the diffusive term integrates to zero because of the boundary conditions and furthermore the contribution of the support function is captured in the effective support  $\bar{s}$ , which is equivalent to assuming  $S(\varphi)$  independent of  $\varphi$ .

Equations (3.25) are a simple linear system that can be solved exactly. In matrix notation they become

$$\partial_t \vec{\pi} = \begin{pmatrix} \sigma_R & -\sigma_n \bar{s} \\ -\sigma_R & \sigma_n \bar{s} \end{pmatrix} \vec{\pi}, \quad (3.26)$$

with

$$\vec{\pi} \equiv \begin{pmatrix} \pi_R(t) \\ \pi_n(t) \end{pmatrix}. \quad (3.27)$$

The general solution is

$$\vec{\pi}(t) = \vec{v}_1 k_1 e^{\lambda_1 t} + \vec{v}_2 k_2 e^{\lambda_2 t}, \quad (3.28)$$

where the  $v_i$  are the eigenvectors corresponding to the eigenvalues  $\lambda_i$ . The  $k_i$  are constants to be determined through the initial conditions. After some linear algebra the following is obtained for the time dependence of the power of the regime and the niche:

$$\begin{aligned}\pi_R(t) &= \frac{\sigma_n \bar{s}}{\sigma_R} k_1 - k_2 e^{(\sigma_R + \sigma_n \bar{s})t} \\ \pi_n(t) &= k_1 + k_2 e^{(\sigma_R + \sigma_n \bar{s})t}.\end{aligned}\quad (3.29)$$

To see the implications it is necessary to supply some initial conditions. In a typical system with an incumbent regime and an upcoming niche the power of the regime is large in comparison to that of the niche. It is then interesting to investigate how the initial power of the niche influences the evolution. To model this one assigns a value to the regime at  $t = 0$ , which can without loss of generality be set to unity, and leaves the initial power of the niche as a variable, like

$$\begin{aligned}\pi_R(0) &= 1 \\ \pi_n(0) &= \epsilon.\end{aligned}\quad (3.30)$$

### BUILDING 3. “THE DYNAMICS OF FUNCTIONING”

Typically one assumes  $\epsilon \ll 1$ . For the  $k_i$  one now finds

$$k_1 = \frac{1 + \epsilon}{1 + \left(\frac{\sigma_n \bar{s}}{\sigma_R}\right)} \quad \text{and} \quad k_2 = \frac{-1 + \left(\frac{\sigma_n \bar{s}}{\sigma_R}\right) \epsilon}{1 + \left(\frac{\sigma_n \bar{s}}{\sigma_R}\right)}. \quad (3.31)$$

A quick glance at equations (3.29) learns that for the niche to successfully scale up to a level where its power will exceed that of the regime,  $k_2$  needs to be positive. Since all upscaling parameters are positive this implies the *upscaling condition*

$$(\sigma_n \bar{s}) \epsilon > \sigma_R. \quad (3.32)$$

The interpretation of this condition is readily found. No matter how powerful the regime is in the beginning — that is, no matter how small  $\epsilon$  — if its *upscaling potential* is large enough it will become more powerful than the regime at a certain point in time. Of course, like alluded to earlier, before this point is reached the modelling assumptions become inconsistent since the niche is assumed to adopt regime-like behaviour when it becomes a serious competitor for the incumbent regime.

#### 3.3.2 The effect of functioning on transition paths

From the previous it is clear that in the cases with indirect interactions the functioning might influence the course of a transition, but not the eventual outcome. This is because it is possible to integrate out the dependence on  $\varphi$ . In these cases the end state depends on the initial conditions in a simple way, namely via an upscaling condition. Therefore, to appreciate how the functioning of constellations impacts the course of a transition it is insightful to employ direct interactions.

In the following an attempt is made to explore how the initial form and distribution of constellations lead to different transitions paths and outcomes. For this a simple system of one niche in interaction with a regime is used with a support function that is independent of  $\varphi$ . The influence of support on transition paths will be explored later on. The example system (3.23) will be studied here:

$$\begin{cases} \partial_t R(\varphi, t) = \beta_R \partial_\varphi^2 R - \gamma n R (nS - R) \\ \partial_t n(\varphi, t) = \beta_n \partial_\varphi^2 n + \gamma n R (nS - R). \end{cases} \quad (3.33)$$

with  $S$  taken independent of  $\varphi$ .

For the initial distributions of the regime and the niche Gaussian functions will be used, broad ones for the regime, more sharply peaked ones for the niche. To reflect the typical power distribution before a transition the regime will initially have approximately four to five times as much power as the niche<sup>6</sup>.

In systems with direct interactions, competition produces relevant dynamics. Therefore it is important to know how much power ‘overlaps’ initially, so to speak. The measure for this was already given by (3.5), which gives the power overlap  $\pi_{ij}$  of two constellations  $c_i$  and  $c_j$ .

---

<sup>6</sup>Of course this is disputable, since there is no agreed quantitative measure of power in general. However, within the framework sketched in the beginning of this chapter such a measure exists and then it is somewhat reasonable to say a factor of four to five is a significant difference.

### 3.3. ELABORATIONS

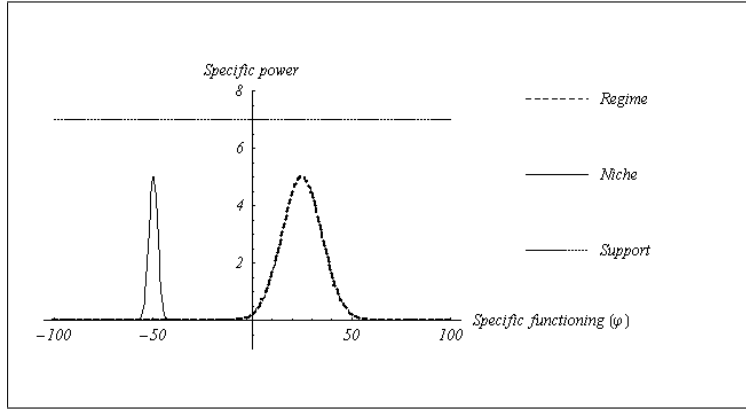


Figure 3.1: Typical initial conditions

The behaviour of the equations (3.23) was studied with Mathematica 5.1. The initial position of the regime was the same throughout and the initial position of the niche varied. This was repeated for various values for the support function. The boundary conditions were Neumann's, that is the derivatives were held zero. A spatial domain running from  $-100$  to  $100$  was used. As said, the regime's initial condition was constant and of the following form:

$$R(\varphi, 0) = 5e^{-\frac{(\varphi-25)^2}{200}}. \quad (3.34)$$

Resulting in a power  $\pi_R$  of 125.3. The form of the niche function was

$$n(\varphi, 0) = 5e^{-\frac{(\varphi-\varphi_0)^2}{10}}, \quad (3.35)$$

that is, less wide and less powerful with  $\pi_n = 28.0$ , see figure (3.1). To reflect the theoretical notion that the regime is more rigid than a niche a factor of five was initially chosen between their broadening parameters, specifically  $\beta_R = 1 \times 10^{-2}$  and  $\beta_n = 5 \times 10^{-2}$ .

As far as the influence of the support function is concerned the results were straightforward, more support increases the possibility for the niche to grow at the expense of the regime. More specifically it was found that between  $S = 4$  and  $S = 3$  the chances of a niche turn. If it has less support than  $S = 3$ , the niche has no chance. Given enough support, however, it proved to improve the chances of a niche to rise to power, when it starts *close* to, e.g. has more overlap power with, the regime.

This raises the following question: What exactly is to be considered a successful transition in this representation of a societal system? Is it a niche becoming the new most powerful constellation or is it a significant shift in the entire functioning of the societal system? Both views are present in the literature. The 'niche becoming the new regime' picture more prevalent in the socio-technical transition literature and the 'significant shift' picture more natural for transitions to sustainability. This is important to keep in mind as the functioning of the niche in becoming the most powerful constellation changed profoundly. To state this more exact, the median of the niche constellation in the process of the transition shifted to almost the median of the regime, like for instance in figure (3.3.2).

### BUILDING 3. “THE DYNAMICS OF FUNCTIONING”

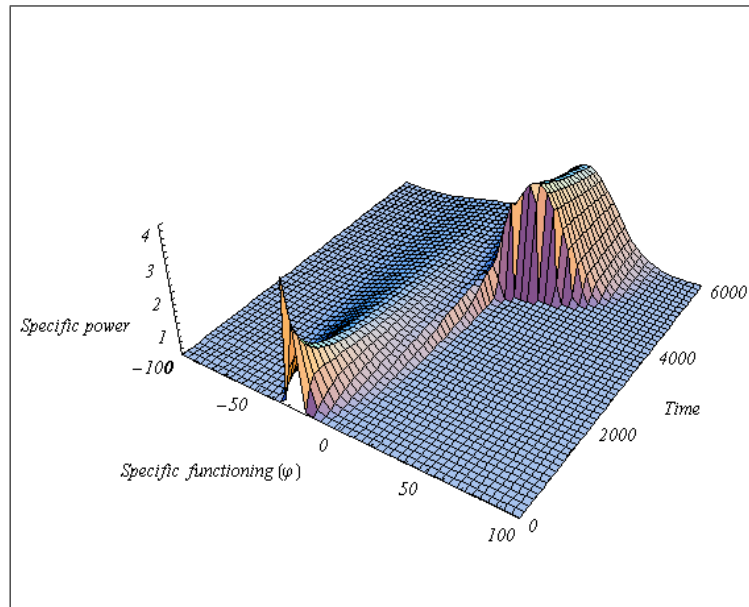


Figure 3.2: Evolution of a niche, gaining power by changing its functioning

#### 3.3.3 Behaviour and strategies of constellations

Or maybe this is because of the flexible behaviour of niche constellations? Maybe the larger value of the broadening parameter (it is a diffusion parameter, after all) is to blame. In any case it is interesting to now how the broadening parameter values influence the course of a transition. The value of  $\beta$  can also be interpreted as a strategy of the constellation, where low values correspond to rigid, but also robust, behaviour and higher values to a flexible behaviour that also makes it more vulnerable. The influence of the broadening parameter was therefore studied as well, keeping in mind the effects on the total functioning of the societal system.

To test this the system was prepared in a similar way as before with a niche functioning in a sufficiently different region than the regime (overlap of order  $10^{-8}$ , zero) and the broadening parameter was varied from 0.01 which is the regime value to 0.1. This again for different values of the support, in principle enough for a niche to become the most powerful constellation. In general it is observed that a high value of  $\beta_n$  makes the niche flexible but vulnerable. When  $\beta_n$  is decreased towards more rigidity nothing much happens for the end state until a certain, apparently critical, value is reached after which the niche takes over the dominant position in the societal system.

The peculiarity lies in decreasing the value of  $\beta_n$  even further. An extremely rigid — regime like — niche, does not change its functioning much until its slowly spreading out tail meets that of the regime. If this happens it can grow rapidly if given enough support. The effect however is that it grows in the region of the incumbent regime. An interpretation could be the following, if a niche is very loyal to its principles, or sticks to its core business it remains the same for a long time but when the opportunity comes forth it is seized, with change of ‘character’. This of course has consequences for how the total functioning

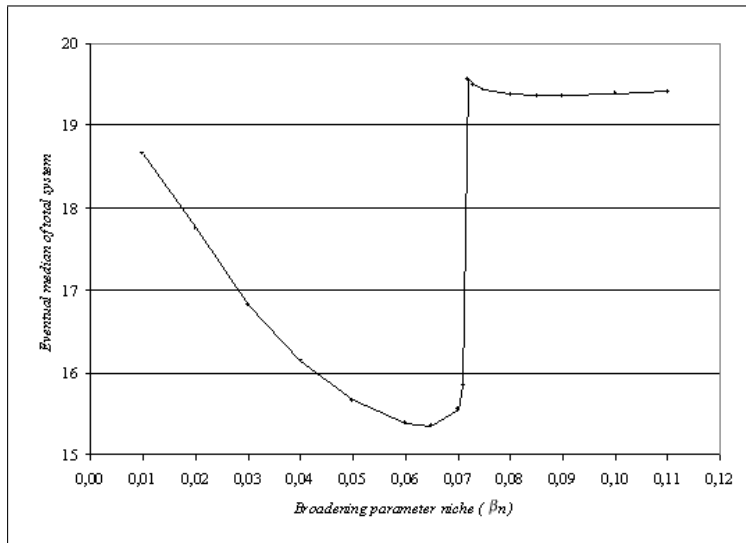


Figure 3.3: Median of total functioning as function of  $\beta_n$

changes as result of such a transition. Figure (3.3.3) shows how the eventual median functioning of the entire system varies as a function of the broadening parameter of the niche.

Note that even at its most extreme the median of the total functioning does not change dramatically. Given that the initial median functioning of the entire system was 22.2, and the niche had its median at  $-50$ , a new median at approximately 15 does not appeal to the idea of fundamental change. Nevertheless, an optimum broadening parameter appears to exist which is larger for larger values of the support. This implies that more support allows a niche to be more flexible, which intuitively makes sense.

### 3.3.4 The effect of societal support on transition paths

From the previous sections it already became clear that the societal support function was of influence on the course of the transition. The support function itself however was taken as independent of the specific functioning  $\varphi$ . The societal interpretation of this would be that a certain niche receives support, regardless of its functioning or future direction thereof. This sounds more ridiculous than it actually is, because subsidy schemes exist that do not discriminate on the sort of activities or research done by niche players but rather look at other criteria.

In the context of strategic niche management and transition management, however, it is relevant to know if it matters if support is given for regions of specific functioning. And if it is, in what way. For instance if the niche and the regime are distributed over the functional domain as in the previous example what are the consequences of a support function that only supports functioning for  $\varphi < 0$ ? Or a support function that supports only the functioning that the niche already performs? Here these two forms are examined that is a support



BUILDING 3. “THE DYNAMICS OF FUNCTIONING”

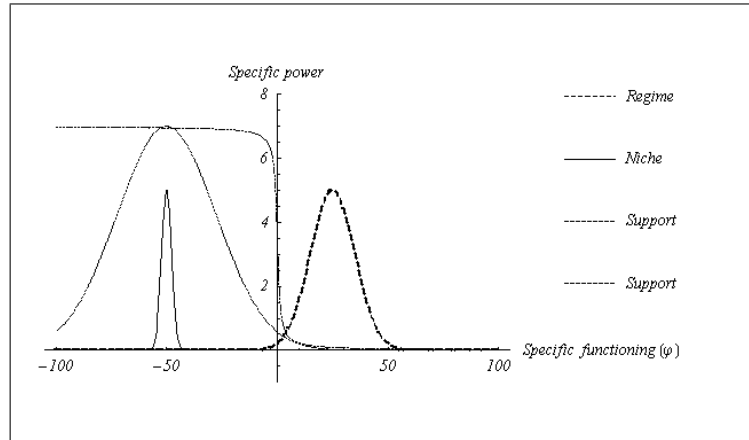


Figure 3.4: Initial conditions with specific support functions

function of the form

$$S(\varphi) = 7\left(\frac{1}{2} + \frac{1}{\pi} \arctan(\varphi)\right), \quad (3.36)$$

and one of the form

$$S(\varphi) = 7e^{-\frac{(\varphi-\varphi_0)^2}{100}}. \quad (3.37)$$

Especially the second form would correspond to what (van den Bosch & Taanman, 2006) call a strategy of ‘deepening’. This implies learning more and refining what the niche already does. See also figure (3.3.4). If the effect of these support functions is compared with the effect of the constant support with the same maximum value, 7, some counterintuitive things are to be noted. As in these cases support is given to a specific and therefore smaller region of functioning and with at most the same value, one would expect the niche to face more difficulties. This appears not the case. If the broadening parameter is varied from large (flexible) to small (rigid) the point at which the niche wins indeed is much further away. Stronger even, the point at which the niche took over in the case of constant support is close to the point where the niche merely survives in the specific support cases. This is made up for by the following, the niche can safely adopt a very rigid strategy. The niche then does not waste functioning in competition with the regime, since support is only given where the niche already functions and the regime is (almost) absent. For the transition as a whole this implies that when a niche adopts such a strategy and holds on the change in functioning is much more profound. As one can see in figure (3.3.3).

This works for the both forms of the support function with the best results with the form of equation (3.37).

### 3.4 Conclusions and discussion

Two types of conclusions can be drawn from this chapter. The first type concerns the contribution that this spatial approach to societal transitions is to this young field. In other words, how useful is the use of this approach in producing models based on current theoretical insights. The second is the converse, that is, what new insights for transition science have been gained by using this approach.

### 3.4. CONCLUSIONS AND DISCUSSION

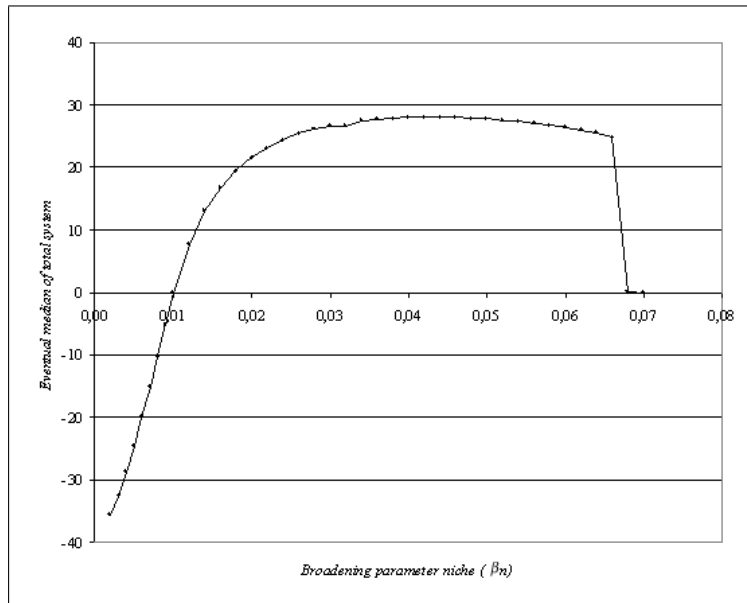


Figure 3.5: Transitions with a locally supported niche

To start with the first, by basing everything on the parameterisation of the concept of functioning the framework allowed for a mathematical description of some central transition theory concepts such as, regime and niches. On the methodological side of things it connects with the spatial approaches that have been proven insightful in political science and simultaneously with a rich field in applied mathematics. Transition theory concepts such as ‘broadening’, ‘deepening’ and ‘scaling up’ appeared naturally as parameters in the equations and even power entered in a straightforward way.

The second type of conclusions, the new insights, are best drawn with caution and modesty. The assumptions are crude and interpretation still shaky. However, the way for instance the influence of the broadening parameter and the form of the support function influence the course of the transitions in the model runs, could potentially be very useful and insightful in the context of strategic niche management and transition management, since both approaches to steering deal with niches that need to scale up and possibly change or take over the regime. Simple hypotheses like ‘it is good (or not) to support specific functioning for a transition to sustainability’ can be put to the test under various modelling assumptions. In the abstract some of these simple hypothesis have been evaluated in this chapter.

A pde approach like the one presented in this chapter has elegance in that all assumptions and hypotheses are compounded in the set of evolution equations and the initial and boundary conditions. It is in principle possible to glance at the equations and to know what the model is meant to model. But the drawback for this clarity is that the mathematical machinery is not that simple and to translate a hypothesis about how niches and regimes interact into a term in an equation demands some prior insight, or rather mathematical intuition.

Difficulties arise when one wants to increase the dimensionality. Through-

### *BUILDING 3. "THE DYNAMICS OF FUNCTIONING"*

out this chapter one dimension was used. For a realistic characterisation of the functioning of a societal system, however, probably several dimensions are necessary. This is no formal problem for the framework, but the tractability of the equations becomes difficult, first and foremost for analytical treatment, but simulation and numerical approaches as well suffer in more dimensions. Then again, in the political pendants in two dimensions already a lot of insight was gained. Expanding the model to feature more than just a regime and one niche is however rather easy.

Another difficulty is that although in this framework a niche can scale up to a point where it has gained more power than the regime it would still 'behave' as a niche. Simply because it is stuck in its equation so to say. This could in future approaches be solved by simulations in several parts, or a form of the equations where this niche-ness and regime-ness are parameters. This latter could be done by letting the diffusion coefficient, or in the terminology of this chapter the broadening parameter, depend on the power of the constellation.

In using this method some new questions arose as well, or in any case took on a new guise. An example of this is the relative importance of a niche become the new most powerful constellation in the system. In some cases the functioning of the entire system shifted significantly while the niche perished in the process. It is also, and maybe even in the first place, this kind of questions that justify the use of methods like these. That is for the rigorous exploration of theoretical assumptions and to serve as an conceptual laboratory for the transition scientist.

# References

- AVELINO, F. 2009. Empowerment and the challenge of applying transition management to ongoing projects. *Policy Sciences*, 42(4): 369–390.
- AVELINO, F. & J. ROTMANS 2009. Power in transition: An interdisciplinary framework to study power in relation to structural change. *European Journal of Social Theory*, 12(4): 543–569.
- BAGGIO, C., M. HOWARD, & W. V. SAARLOOS 2004. Dynamics and stability of vortex-antivortex fronts in type-II superconductors. *Physical Review E*, 70(026209).
- VAN DEN BOSCH, S. & M. TAANMAN 2006. How innovation impacts society. patterns and mechanisms through which innovation projects contribute to transitions.
- CROSS, M. C. & P. C. HOHENBERG 1993. Pattern formation outside of equilibrium. *Reviews of Modern Physics*, 65(3): 851–1112.
- DOWNES, A. 1957. *An Economic Theory of Democracy*. Harper and Brothers, New York.
- FARLOW, S. J. 1982. *Partial Differential Equations for Scientists and Engineers*. Dover Publications, Inc., New York, 2nd edition.
- GEELS, F. 2002. *Understanding the Dynamics of Technological Transitions: A Co-evolutionary and socio-technical analysis*. Ph.d. thesis, Twente University.
- GIDDENS, A. 1984. *The Constitution of Society*. University of California Press, Berkeley and Los Angeles.
- DE HAAN, J. 2007. Pillars of change: A theoretical framework for transition models, presented at the esee 2007 conference: “integrating natural and social sciences for sustainability”.
- HAUGAARD, M. 2002. *Power: A Reader*. Manchester University Press, Manchester and New York.
- KEMP, R., J. SCHOT, & R. HOOGMA 1998. Regime shifts to sustainability through processes of niche formation: The approach of strategic niche management. *Technology Analysis and Strategic Management*, 10(2): 175–195.
- KOLLMAN, K., J. H. MILLER, & S. E. PAGE 1992. Adaptive parties in spatial elections. *American Political Science Review*, 86(4): 8.

## REFERENCES

- LANDAU, L. & E. LIFSHITZ 1951. *The Classical Theory of Fields*, volume 2 of *Course of Theoretical Physics*. Elsevier, Ltd., fourth revised english edition (1975) edition.
- LAVER, M. 2005. Policy and the dynamics of political competition. *American Political Science Review*, 99(2): 18.
- LEVINTHAL, D. A. 1998. The slow pace of rapid technological change: Gradualism and punctuation in technological change. *Industrial and Corporate Change*, 7: 217–247.
- LUHMANN, N. 1984. *Social Systems*. Stanford University Press, Stanford, California.
- MOSSBERGER, K. & G. STOKER 2001. The evolution of urban regime theory: The challenge of conceptualization. *Urban Affairs Review*, 36(6): 810–834.
- NELSON, R. R. & S. G. WINTER 1977. In search of a useful theory of innovation. *Research Policy*, 6: 36–76.
- ROGERS, E. M. 2003. *Diffusion of Innovations*. Free Press, New York, 5th edition.
- ROTMANS, J. 2005. *Societal Innovation: between dream and reality lies complexity*. Inaugural Addresses Research in Management Series. ERIM, Erasmus Research Institute of Management, Rotterdam.
- VAN SAARLOOS, W. 2003. Front propagation into unstable states. *Physical Reports*, 386: 29–222.
- SCHOT, J. & F. W. GEELS 2007. Niches in evolutionary theories of technical change: A critical survey of the literature. *Journal of Evolutionary Economics*, 17(5): 605–622.
- WILSON, C. A. 2000. Policy regimes and policy change. *Journal of Public Policy*, 20(3): 247–274.





# Coda





# Biographical sketch

The author of this thesis was born in Rotterdam the tenth of September of the year 1977. He received both his primary education (basisschool de Sleutel) and secondary education (Chr. S.G. Johannes Calvijn, Atheneum) there. Discovering that working in any regular occupation entailed boredom and made inspiration flee, he began to study physics at Leiden University, graduating in theoretical physics under the supervision of Wim van Saarloos in the winter of 2003.

A master now, though not interested in pursuing more research in physics, a new search began. This led to the discovery of the newly founded Dutch Research Institute for Transitions at the Erasmus University in Rotterdam, which appeared just the place where a combination of intellectual endeavour and societal interest could be found. This was exactly the case, and on Valentine's day 2005 he started his Ph.D. research there under the supervision of Jan Rotmans. The assignment was not clearly demarcated, but if he would come up with a transition theory useful for narrative description, understanding, explaining, and modelling, the supervisor would be satisfied. Which was early 2010.

While the thesis was in the last stage of corrections and polishing, the author took a post-doctoral position at the faculty Technology, Policy and Management, at the Policy Analysis section headed by his committee member Wil Thissen.

Like the theory presented in this thesis is built upon three pillars, so is the life of its creator. Or at least the creative part thereof, which is all that matters anyway. Poetry, Music, Philosophy. Where the coming into being of the theory should be considered part of the last pillar.



# Samenvatting

In deze dissertatie wordt een theorie gepresenteerd over maatschappelijke transities, grootschalige — sectoren, industrieën — omwentelingen over termijnen van soms meerdere generaties. Uiteraard hebben vele vakgebieden zich hier al mee beziggehouden, economie, sociologie, geschiedenis, etc., maar een vakgebied dat deze fenomenen als zodanig bestudeert, met middelen en methoden uit vele disciplines, begint zich juist nu pas te vormen.

Derhalve bestond de behoefte om de grote, maar versplinterde, hoeveelheid theoretische kennis in combinatie met recente inzichten samen te brengen en uit te breiden tot een overkoepelende theorie. Niet om daarmee het laatste woord gezegd te kunnen hebben over maatschappelijke transities, maar veeleer om een basis te zijn voor verder onderzoek.

Een dergelijke theorie zou naast haar ‘talige’ verschijning, ook geschikt moeten zijn voor modellerbenaderingen, wiskundig, computationeel of anderszins. Dit stelt eisen aan de vorm, of liever, de structuur van de theorie. Daarnaast dient een theorie, om zich wetenschappelijk te durven noemen, zich kwetsbaar, namelijk toetsbaar, op te stellen. Tegelijkertijd is het wenselijk dat het bouwwerk niet met elk nieuw inzicht een beetje afbrokkeld, maar integendeel steviger komt te staan.

Om de stelling te onderbouwen dat *pillar theory* een dergelijke theorie is, heeft deze dissertatie de vorm van een drieluik. In het tweede boek wordt de eigenlijke theorie gepresenteerd, in haar talige verschijning en in een formele, pseudowiskundige formulering. In het eerste boek wordt een wetenschapsfilosofische onderbouwing van zowel de vorm als de inhoud van de theorie gegeven. Het derde boek probeert de bruikbaarheid van de theorie te demonstreren door haar op drie manieren toe te passen. Als eerste in een narratieve case study waarin transities in ruim anderhalve eeuw Nederlandse gezondheidszorg worden geanalyseerd. Daarna in een computermodel en een wiskundig model, die kunnen worden beschouwd als rigoureuze gedachtenexperimenten.